

DRUG SAFETY EVALUATION

DRUG SAFETY EVALUATION

3rd Edition

SHAYNE COX GAD

WILEY

Copyright © 2017 by John Wiley & Sons, Inc. All rights reserved

Published by John Wiley & Sons, Inc., Hoboken, New Jersey
Published simultaneously in Canada

No part of this publication may be reproduced, stored in a retrieval system, or transmitted in any form or by any means, electronic, mechanical, photocopying, recording, scanning, or otherwise, except as permitted under Section 107 or 108 of the 1976 United States Copyright Act, without either the prior written permission of the Publisher, or authorization through payment of the appropriate per-copy fee to the Copyright Clearance Center, Inc., 222 Rosewood Drive, Danvers, MA 01923, (978) 750-8400, fax (978) 750-4470, or on the web at www.copyright.com. Requests to the Publisher for permission should be addressed to the Permissions Department, John Wiley & Sons, Inc., 111 River Street, Hoboken, NJ 07030, (201) 748-6011, fax (201) 748-6008, or online at <http://www.wiley.com/go/permissions>.

Limit of Liability/Disclaimer of Warranty: While the publisher and author have used their best efforts in preparing this book, they make no representations or warranties with respect to the accuracy or completeness of the contents of this book and specifically disclaim any implied warranties of merchantability or fitness for a particular purpose. No warranty may be created or extended by sales representatives or written sales materials. The advice and strategies contained herein may not be suitable for your situation. You should consult with a professional where appropriate. Neither the publisher nor author shall be liable for any loss of profit or any other commercial damages, including but not limited to special, incidental, consequential, or other damages.

For general information on our other products and services or for technical support, please contact our Customer Care Department within the United States at (800) 762-2974, outside the United States at (317) 572-3993 or fax (317) 572-4002.

Wiley also publishes its books in a variety of electronic formats. Some content that appears in print may not be available in electronic formats. For more information about Wiley products, visit our web site at www.wiley.com.

Library of Congress Cataloging-in-Publication Data:

Names: Gad, Shayne C., 1948–, author.

Title: Drug safety evaluation / Shayne Cox Gad.

Description: 3rd edition. | Hoboken, New Jersey : John Wiley & Sons, Inc., [2017] | Includes bibliographical references and index.

Identifiers: LCCN 2016014287 | ISBN 9781119097396 (cloth) | ISBN 9781119097402 (Adobe pdf) | ISBN 9781119097419 (ePub)

Subjects: | MESH: Drug Evaluation, Preclinical | Drug Evaluation | Product Surveillance, Postmarketing

Classification: LCC RA1238 | NLM QV 771 | DDC 615/.19–dc23

LC record available at <http://lccn.loc.gov/2016014287>

Set in 10/12pt Times by SPi Global, Pondicherry, India

Printed in the United States of America

10 9 8 7 6 5 4 3 2 1

To the memory of my mother Norma Jean Cox Gad, who crossed over nine years ago, and my brother Scott Michael Gad who joined her six years ago. I hope that all your beloved little friends are there with you. I will see you both again.

—Shayne Cox Gad

ACKNOWLEDGEMENT

While all of my interns assisted in the preparation of this volume, its completion would not have occurred without the efforts of Leslie Eagle and indexing by my beloved wife, Novie Beth Ragan Gad.

CONTENTS

| | |
|--|--------------|
| PREFACE | xxv |
| ABOUT THE AUTHOR | xxvii |
| 1 The Drug Development Process and the Global Pharmaceutical Marketplace | 1 |
| 1.1 Introduction, 1 | |
| 1.2 The Marketplace, 1 | |
| 1.3 History of Modern Therapeutics, 4 | |
| 1.4 The Drug Development Process, 6 | |
| 1.5 Strategies for Development: Large Versus Small Company or the Short Versus Long Game, 7 | |
| 1.5.1 Do Only What You Must, 8 | |
| 1.5.2 Minimize the Risk of Subsequent Failure, 9 | |
| 1.6 Safety Assessment and the Evolution of Drug Safety, 11 | |
| 1.7 The Three Stages of Drug Safety Evaluation in the General Case, 11 | |
| References, 12 | |
| 2 Regulation of Human Pharmaceutical Safety: Routes to Human Use and Market | 13 |
| 2.1 Introduction, 13 | |
| 2.2 Brief History of US Pharmaceutical Law, 13 | |
| 2.2.1 1906: Pure Food and Drug Act, 13 | |
| 2.2.2 1938: Food, Drug, and Cosmetic Act, 15 | |
| 2.2.3 1962: Major Amendment, 17 | |
| 2.2.4 1992, 1997, 2002, 2007, and 2012: PDUFA and FDAMA, 18 | |
| 2.3 FDAMA Summary: Consequences and Other Regulations, 19 | |
| 2.4 Overview of US Regulations, 21 | |
| 2.4.1 Regulations: General Considerations, 21 | |
| 2.4.2 Regulations: Human Pharmaceuticals, 22 | |
| 2.4.3 Regulations: Environmental Impact, 23 | |
| 2.4.4 Regulations: Antibiotics, 23 | |
| 2.4.5 Regulations: Biologics, 24 | |
| 2.4.6 Regulations versus Law, 24 | |

- 2.5 Organizations Regulating Drug and Device Safety in the United States, 24
- 2.6 Process of Pharmaceutical Product Development and Approval, 25
- 2.7 Testing Guidelines, 28
 - 2.7.1 Toxicity Testing: Traditional Pharmaceuticals, 28
 - 2.7.2 General or Systematic Toxicity Assessment, 28
 - 2.7.3 Genetic Toxicity Assessment, 28
 - 2.7.4 Safety Pharmacology, 30
 - 2.7.5 Local Tissue Tolerance, 30
 - 2.7.6 Toxicity Testing: Biotechnology Products, 31
- 2.8 Toxicity/Safety Testing: Cellular and Gene Therapy Products, 33
 - 2.8.1 Cellular Therapies, 34
 - 2.8.2 Gene Therapies, 34
 - 2.8.3 *Ex Vivo*, 34
 - 2.8.4 *In Vivo*, 34
 - 2.8.5 Preclinical Safety Evaluation, 34
 - 2.8.6 Basic Principles for Preclinical Safety Evaluation of Cellular and Gene Therapies, 35
 - 2.8.7 Additional Considerations for Cellular Therapies, 35
 - 2.8.8 Additional Considerations for Gene Therapies, 35
- 2.9 Toxicity Testing: Special Cases, 35
 - 2.9.1 Oral Contraceptives, 35
 - 2.9.2 Life-Threatening Diseases (Compassionate Use), 35
 - 2.9.3 Optical Isomers, 36
 - 2.9.4 Special Populations: Pediatric and Geriatric Claims, 37
 - 2.9.5 Orphan Drugs, 38
 - 2.9.6 Botanical Drug Products, 41
 - 2.9.7 Types of New Drug Applications (NDAs), 41
- 2.10 International Pharmaceutical Regulation and Registration, 41
 - 2.10.1 International Conference on Harmonization, 41
 - 2.10.2 Other International Considerations, 45
 - 2.10.3 Safety Pharmacology, 50
- 2.11 Combination Products, 50
 - 2.11.1 Device Programs That CDER and CBRH Each Will Administer, 51
 - 2.11.2 Coordination, 51
 - 2.11.3 Submissions, 51
- 2.12 Conclusions, 55
- References, 55
- Further Reading, 57

3 Data Mining: Sources of Information for Consideration in Study and Program Design and in Safety Evaluation 59

- 3.1 Introduction, 59
 - 3.1.1 Claims, 59
 - 3.1.2 Time and Economies, 59
 - 3.1.3 Prior Knowledge, 59
 - 3.1.4 Miscellaneous Reference Sources, 60
 - 3.1.5 Search Procedure, 62
 - 3.1.6 Monitoring Published Literature and Other Research in Progress, 62
 - 3.1.7 Kinds of Information, 63
 - 3.1.8 Toxic Release Inventory (TRI), 63
 - 3.1.9 Material Safety Data Sheets (MSDS), 63
 - 3.1.10 Canadian Centre for Occupational Health and Safety (CCINFO), 64
 - 3.1.11 Pollution and Toxicology (POLTOX), 64
 - 3.1.12 MEDLINE, 64

- 3.2 PC-Based Information Products: Laser DISC, 65
 - 3.2.1 International Veterinary Pathology Slide Bank (IVPSB), 65
- 3.3 Conclusions, 65
- References, 65

4 Screens in Safety and Hazard Assessment 67

- 4.1 Introduction, 67
- 4.2 Characteristics of Screens, 68
- 4.3 Uses of Screens, 70
- 4.4 Types of Screens, 71
 - 4.4.1 Single Stage, 71
 - 4.4.2 Sequential, 71
 - 4.4.3 Tier (or Multistage), 71
- 4.5 Criterion: Development and Use, 71
- 4.6 Analysis of Screening Data, 73
- 4.7 Univariate Data, 73
 - 4.7.1 Control Charts, 73
 - 4.7.2 Central Tendency Plots, 74
 - 4.7.3 Multivariate Data, 75
 - 4.7.4 The Analog Plot, 75

References, 76

5 Formulations, Routes, and Dosage Regimens 79

- 5.1 Mechanisms, 81
 - 5.1.1 Local Effects, 81
 - 5.1.2 Absorption and Distribution, 81
 - 5.1.3 Metabolism, 82
- 5.2 Common Routes, 83
 - 5.2.1 Dermal Route, 83
 - 5.2.2 Parenteral Route, 84
 - 5.2.3 Bolus versus Infusion, 85
 - 5.2.4 Oral Route, 86
 - 5.2.5 Minor Routes, 94
 - 5.2.6 Route Comparisons and Contrasts, 96
- 5.3 Formulation of Test Materials, 96
 - 5.3.1 Preformulation, 97
 - 5.3.2 Dermal Formulations, 100
 - 5.3.3 Interactions between Skin, Vehicle, and Test Chemical, 102
 - 5.3.4 Oral Formulations, 103
 - 5.3.5 Parenteral Formulations, 104
- 5.4 Dosing Calculations, 105
- 5.5 Calculating Material Requirements, 105
- 5.6 Excipients, 106
 - 5.6.1 Regulation of Excipients, 106

References, 111

6 Nonclinical Manifestations, Mechanisms, and End Points of Drug Toxicity 115

- 6.1 Manifestations, 115
- 6.2 Mechanisms of Toxicity, 116
- 6.3 End Points Measured in General Toxicity Studies, 116
 - 6.3.1 Clinical Observations, 116
 - 6.3.2 Body Weights, 116
 - 6.3.3 Food and Water Consumption, 116

- 6.3.4 Clinical Signs, 117
- 6.3.5 Clinical Chemistry and Pathology, 117
- 6.3.6 Hematology, 124
- 6.3.7 Gross Necropsy and Organ Weights, 124
- 6.3.8 Histopathology, 125
- 6.3.9 Ophthalmology, 125
- 6.3.10 Cardiovascular Function, 125
- 6.3.11 Neurotoxicology, 125
- 6.3.12 Immunotoxicology, 125
- 6.4 Complications, 126
- References, 126

7 Pilot Toxicity Testing in Drug Safety Evaluation: MTD and DRF 129

- 7.1 Introduction, 129
- 7.2 Range-Finding Studies, 130
 - 7.2.1 Lethality Testing, 130
 - 7.2.2 Using Range-Finding Lethality Data in Drug Development: The Minimum Lethal Dose, 136
- 7.3 Acute Systemic Toxicity Characterization, 138
 - 7.3.1 Minimal Acute Toxicity Test, 139
 - 7.3.2 Complete Acute Toxicity Testing, 142
 - 7.3.3 Acute Toxicity Testing with Nonrodent Species, 146
 - 7.3.4 Factors that Can Affect Acute Tests, 148
 - 7.3.5 Selection of Dosages, 149
- 7.4 Screens, 150
 - 7.4.1 General Toxicity Screens, 151
 - 7.4.2 Specific Toxicity Screening, 153
- 7.5 PILOT and DRF Studies, 154
- References, 156

8 Repeat-Dose Toxicity Studies 159

- 8.1 Objectives, 159
- 8.2 Regulatory Considerations, 161
 - 8.2.1 Good Laboratory Practices, 161
 - 8.2.2 Animal Welfare Act, 161
 - 8.2.3 Regulatory Requirements for Study Design, 162
- 8.3 Study Design and Conduct, 162
 - 8.3.1 Animals, 162
 - 8.3.2 Routes and Setting Doses, 163
 - 8.3.3 Parameters to Measure, 164
 - 8.3.4 Study Designs, 164
- 8.4 Study Interpretation and Reporting, 165
- References, 166

9 Genotoxicity 169

- 9.1 ICH Test Profile, 169
- 9.2 DNA Structure, 169
 - 9.2.1 Transcription, 171
 - 9.2.2 Translation, 171
 - 9.2.3 Gene Regulation, 171
 - 9.2.4 DNA Repair, 171
 - 9.2.5 Error-Prone Repair, 172
 - 9.2.6 Mismatch Repair, 172
 - 9.2.7 The Adaptive Repair Pathway, 172
 - 9.2.8 Plasmids, 172
 - 9.2.9 Plasmids and DNA Repair, 173

- 9.2.10 Nature of Point Mutations, 173
- 9.2.11 Suppressor Mutations, 173
- 9.2.12 Adduct Formation, 173
- 9.2.13 Mutations Due to Insertion Sequences, 174
- 9.2.14 The Link between Mutation and Cancer, 174
- 9.2.15 Genotoxic versus Nongenotoxic Mechanisms of Carcinogenesis, 174
- 9.2.16 Genetic Damage and Heritable Defects, 175
- 9.2.17 Reproductive Effects, 176
- 9.3 Cytogenetics, 176
 - 9.3.1 Cytogenetic Damage and Its Consequences, 176
 - 9.3.2 Individual Chromosomal Damage, 176
 - 9.3.3 Chromosome Set Damage, 177
 - 9.3.4 Test Systems, 177
 - 9.3.5 *In Vitro* Test Systems, 178
 - 9.3.6 Bacterial Mutation Tests, 180
 - 9.3.7 Controls, 182
 - 9.3.8 Plate Incorporation Assay, 184
 - 9.3.9 Eukaryotic Mutation Tests, 185
 - 9.3.10 *In Vitro* Tests for the Detection of Mammalian Mutation, 185
 - 9.3.11 *In Vivo* Mammalian Mutation Tests, 193
- 9.4 *In Vitro* Cytogenetic Assays, 193
 - 9.4.1 Cell Types, 194
 - 9.4.2 Chinese Hamster Cell Lines, 194
 - 9.4.3 Human Peripheral Blood Lymphocytes, 194
 - 9.4.4 Positive and Negative Controls, 194
 - 9.4.5 Treatment of Cells, 195
 - 9.4.6 Scoring Procedures, 195
 - 9.4.7 Data Recording, 196
 - 9.4.8 Presentation of Results, 196
- 9.5 *In Vivo* Cytogenetic Assays, 196
 - 9.5.1 Somatic Cell Assays, 196
 - 9.5.2 Germ Cell Assays, 197
 - 9.5.3 Heritable Chromosome Assays, 197
 - 9.5.4 Germ Cell Cytogenetic Assays, 197
- 9.6 Sister Chromatid Exchange Assays, 197
 - 9.6.1 Relevance of SCE in Terms of Genotoxicity, 198
 - 9.6.2 Experimental Design, 198
- References, 199

10 QSAR Tools for Drug Safety

209

- 10.1 Structure–Activity Relationships, 209
 - 10.1.1 Basic Assumptions, 210
 - 10.1.2 Molecular Parameters of Interest, 210
- 10.2 SAR Modeling Methods, 210
- 10.3 Applications in Toxicology, 212
 - 10.3.1 Metabolism, 213
 - 10.3.2 Reproductive, 213
 - 10.3.3 Eye Irritation, 213
 - 10.3.4 Lethality, 214
 - 10.3.5 Carcinogenicity, 214
- 10.4 Genotoxicity, 215
 - 10.4.1 QSAR for Mutagenicity, 215
- 10.5 Comparison of Available Models/Applications, 216
 - 10.5.1 QSAR of Metabolism, 216

- 10.5.2 Meteor, 216
- 10.5.3 Derek, 218
- 10.5.4 Leadscope, 219
- 10.5.5 VEGA, 219
- 10.5.6 Derek versus Leadscope, 222

References, 222

11 Immunotoxicology in Drug Development 225

- 11.1 Introduction, 225
- 11.2 Overview of the Immune System, 227
- 11.3 Immunotoxic Effects, 229
- 11.4 Immunosuppression, 231
 - 11.4.1 Immunosuppressive Drugs, 232
- 11.5 Immunostimulation, 235
 - 11.5.1 Hypersensitivity (or Allergenicity), 235
 - 11.5.2 Photosensitization, 238
 - 11.5.3 Autoimmunity, 238
- 11.6 Regulatory Positions, 240
 - 11.6.1 CDER Guidance for Investigational New Drugs, 242
- 11.7 Evaluation of the Immune System, 245
 - 11.7.1 Immunopathologic Assessments, 246
 - 11.7.2 Humoral (Innate) Immune Response and Possible Entry Points for Immunotoxic Actions, 246
 - 11.7.3 Cell-Mediated Immunity, 250
- 11.8 Nonspecific Immunity Function Assay, 251
 - 11.8.1 Natural Killer Cell Assays, 251
 - 11.8.2 Macrophage Function, 251
 - 11.8.3 Mast Cell/Basophil Function, 252
- 11.9 T-Cell-Dependent Antibody Response (TDAR), 253
 - 11.9.1 Treatment, 253
 - 11.9.2 Hypersensitivity, 253
 - 11.9.3 Local Lymph Node Assay (LLNA), 255
 - 11.9.4 Photosensitization, 258
- 11.10 Approaches to Compound Evaluation, 259
 - 11.10.1 Use of *In Vivo* Tests, 260
 - 11.10.2 Use of *In Vitro* Tests, 261
 - 11.10.3 Assessment of Immunotoxicity and Immunogenicity/Allergenicity of Biotechnology-Derived Drugs, 261
 - 11.10.4 Suggested Approaches to Evaluation of Results, 262
- 11.11 Problems and Future Directions, 263
 - 11.11.1 Data Interpretation, 263
 - 11.11.2 Appropriate Animal Models, 263
 - 11.11.3 Indirect Immunotoxic Effects, 263
 - 11.11.4 Hypersensitivity Tests, 263
 - 11.11.5 Autoimmunity, 264
 - 11.11.6 Functional Reserve Capacity, 264
 - 11.11.7 Significance of Minor Perturbations, 264
 - 11.11.8 Biotechnology Products, 264

References, 264

12 Nonrodent Animal Studies 269

- 12.1 Introduction, 269
- 12.2 Comparison Between Rodent and Nonrodent Experimental Design, 269
 - 12.2.1 Number of Animals, 269

| | |
|---------|--|
| 12.3 | Differences in Study Activities, 270 |
| 12.3.1 | Blood Collection, 270 |
| 12.3.2 | Dosing, 270 |
| 12.3.3 | Handling of Animals, 270 |
| 12.3.4 | Behavioral Evaluation, 270 |
| 12.4 | Nonrodent Models, 270 |
| 12.5 | Dog, 270 |
| 12.5.1 | Environmental and Dietary Requirements, 270 |
| 12.5.2 | Common Study Protocols, 271 |
| 12.5.3 | General Study Activities, 272 |
| 12.5.4 | Advantages and Disadvantages, 272 |
| 12.6 | The Ferret, 273 |
| 12.6.1 | Environmental and Dietary Requirements, 273 |
| 12.6.2 | Study Protocols, 273 |
| 12.6.3 | General Study Activities, 274 |
| 12.6.4 | Advantages and Disadvantages, 275 |
| 12.7 | The Pig, 275 |
| 12.7.1 | Background, 275 |
| 12.7.2 | Clinical Laboratory, 276 |
| 12.7.3 | Xenobiotic Metabolism, 277 |
| 12.7.4 | Dermal Toxicity, 278 |
| 12.7.5 | Cardiovascular Toxicity, 279 |
| 12.7.6 | Advantages and Disadvantages, 279 |
| 12.8 | Nonhuman Primates, 279 |
| 12.8.1 | Environmental and Dietary Requirements, 281 |
| 12.8.2 | Common Study Protocols, 281 |
| 12.8.3 | General Study Activities, 281 |
| 12.8.4 | Advantages and Disadvantages, 283 |
| 12.9 | Statistics in Large Animal Studies, 283 |
| 12.9.1 | Reasons for Small Sample Sizes in Large Animal Toxicology, 284 |
| 12.9.2 | Cross-Sectional or Longitudinal Analysis?, 284 |
| 12.9.3 | Repeated Measures: Advantages, 284 |
| 12.9.4 | Repeated Measures: Disadvantages, 284 |
| 12.9.5 | Common Practices in Large Animal Toxicology, 284 |
| 12.9.6 | Univariate (Repeated Measures) Techniques: Advantages, 285 |
| 12.9.7 | Univariate (Repeated Measures) Techniques: Disadvantages, 285 |
| 12.9.8 | Multivariate Techniques: Advantages, 285 |
| 12.9.9 | Multivariate Techniques: Disadvantages, 285 |
| 12.9.10 | Some Other Design Factors to Be Considered in Analysis, 285 |
| 12.9.11 | Covariates: Advantages, 285 |
| 12.9.12 | Covariates: Disadvantages, 285 |
| 12.9.13 | Missing Values, 288 |
| 12.10 | Summary, 288 |
| | References, 288 |

13 Developmental and Reproductive Toxicity Testing **291**

| | |
|--------|--|
| 13.1 | Introduction, 291 |
| 13.2 | ICH Study Designs, 293 |
| 13.2.1 | Male and Female Fertility and Early Embryonic Development to Implantation, 294 |
| 13.2.2 | Embryo–Fetal Development, 295 |
| 13.2.3 | Adverse Effects, 295 |

| | |
|--------|---|
| 13.2.4 | Pre- and Postnatal Development, 295 |
| 13.2.5 | Single-Study and Two-Study Designs for Rodents, 296 |
| 13.2.6 | Preliminary Studies, 297 |
| 13.2.7 | Toxicokinetics, 297 |
| 13.2.8 | Timing of Studies, 297 |
| 13.3 | Methodological Issues, 298 |
| 13.3.1 | Control of Bias, 298 |
| 13.3.2 | Diet, 298 |
| 13.3.3 | Clinical Pathology, 299 |
| 13.3.4 | Gravid Uterine Weights, 299 |
| 13.3.5 | Implant Counts and Determination of Pregnancy, 301 |
| 13.3.6 | Fetal Examinations, 301 |
| 13.3.7 | Developmental Signs, 302 |
| 13.3.8 | Behavioral Tests, 303 |
| 13.3.9 | Detecting Effects on Male Reproduction, 303 |
| 13.4 | Developmental Studies in Primates, 303 |
| 13.5 | Data Interpretation, 304 |
| 13.5.1 | Use of Statistical Analyses, 304 |
| 13.5.2 | Potential Hazard Categories of Developmental Toxins, 307 |
| 13.5.3 | Associations between Developmental and Maternal Toxicity, 308 |
| 13.5.4 | Assessment of Human Risk, 308 |
| 13.6 | Juvenile and Pediatric Toxicology, 310 |
| 13.7 | <i>In Vitro</i> Tests for Developmental Toxicity, 312 |
| 13.8 | Appraisal of Current Approaches for Determining Developmental and Reproductive Hazards, 316 |
| | References, 317 |

14 Carcinogenicity Studies

321

| | |
|---------|--|
| 14.1 | Introduction, 321 |
| 14.1.1 | History of Xenobiotic Carcinogenesis, 322 |
| 14.2 | Mechanisms and Classes of Carcinogens, 322 |
| 14.3 | Genotoxic Carcinogens, 322 |
| 14.4 | Epigenetic Carcinogens, 325 |
| 14.5 | Regulatory Requirements and Timing, 328 |
| 14.6 | Species and Strain, 328 |
| 14.7 | Animal Husbandry, 330 |
| 14.8 | Dose Selection, 330 |
| 14.8.1 | Number of Dose Levels, 330 |
| 14.8.2 | Number of Control Groups, 330 |
| 14.8.3 | Criteria for Dose Selection, 331 |
| 14.9 | Group Size, 331 |
| 14.10 | Route of Administration, 332 |
| 14.11 | Study Duration, 332 |
| 14.12 | Survival, 332 |
| 14.13 | End Points Measured, 333 |
| 14.14 | Transgenic Mouse Models, 335 |
| 14.14.1 | The Tg.AC Mouse Model, 335 |
| 14.14.2 | The Tg.rasH2 Mouse Model, 336 |
| 14.14.3 | The P53 ^{+/-} Mouse Model, 336 |
| 14.14.4 | The XPA ^{-/-} Mouse Model, 337 |
| 14.15 | Interpretation of Results: Criteria for a Positive Result, 338 |

- 14.16 Statistical Analysis, 338
 - 14.16.1 Exact Tests, 339
 - 14.16.2 Trend Tests, 340
 - 14.16.3 Life Table and Survival Analysis, 341
 - 14.16.4 Peto Analysis, 341
 - 14.16.5 Methods to Be Avoided, 342
 - 14.16.6 Use of Historical Controls, 342
 - 14.16.7 Relevance to Humans, 342
- 14.17 Weight-of-Evidence Factors for Consideration in a Carcinogenicity Assessment Document (CAD), 344
- 14.18 Conclusions, 345
- References, 345

15 Histopathology in Nonclinical Pharmaceutical Safety Assessment 351

- 15.1 Introduction, 351
 - 15.1.1 Pathological Techniques, 354
 - 15.1.2 Organ Weights, 354
- 15.2 Clinical Pathology, 355
 - 15.2.1 Clinical Chemistry, 355
 - 15.2.2 Target Organ Toxicity Biomarkers, 355
- References, 356

16 Irritation and Local Tissue Tolerance in Pharmaceutical Safety Assessment 359

- 16.1 Introduction, 359
- 16.2 Factors Affecting Irritation Responses and Test Outcome, 359
- 16.3 Primary Dermal Irritation (PDI) Test, 360
- 16.4 Other Nonparenteral Route Irritation Tests, 362
- 16.5 Ocular Irritation Testing, 362
- 16.6 Vaginal Irritation, 364
- 16.7 Acute Primary Vaginal Irritation Study in the Female Rabbit, 365
 - 16.7.1 Repeated-Dose Vaginal Irritation in the Female Rabbit, 365
 - 16.7.2 Repeated-Dose Vaginal Irritation in the Ovariectomized Rats, 366
- 16.8 Parenteral Irritation/Tolerance, 367
 - 16.8.1 Parenteral Routes, 367
 - 16.8.2 Test Systems for Parenteral Irritation, 368
- 16.9 Problems in Testing (and Their Resolutions), 370
 - 16.9.1 Alternatives to *In Vivo* Parenteral Tests, 371
- 16.10 Phototoxicity, 371
 - 16.10.1 Theory and Mechanisms, 371
 - 16.10.2 Factors Influencing Phototoxicity/Photosensitization, 372
 - 16.10.3 Predictive Tests for Phototoxicity, 373
 - 16.10.4 3T3 *In Vitro* Test, 373
 - 16.10.5 Rabbit Phototoxicity Test, 373
 - 16.10.6 Guinea Pig, 374
 - 16.10.7 Pyrogenicity, 376
- 16.11 Hemocompatibility, 377
- References, 378

17 Pharmacokinetics and Toxicokinetics in Drug Safety Evaluation 381

- 17.1 Introduction, 381
- 17.2 Regulations, 382

- 17.3 Principles, 382
 - 17.3.1 Preliminary Work, 382
 - 17.3.2 Absorption, 384
 - 17.3.3 Distribution, 388
 - 17.3.4 Metabolism/Biotransformation, 389
 - 17.3.5 Excretion, 394
- 17.4 Pharmacokinetics, 395
- 17.5 Laboratory Methods, 395
 - 17.5.1 Analytical Methods, 395
- 17.6 Sampling Methods and Intervals, 397
 - 17.6.1 Blood, 397
 - 17.6.2 Excreta, 398
 - 17.6.3 Bile, 398
 - 17.6.4 Expired Air, 398
 - 17.6.5 Milk, 398
- 17.7 Study Types, 400
 - 17.7.1 Whole-Body Autoradiography, 401
 - 17.7.2 Mass Balance Studies, 402
- 17.8 Analysis of Data, 402
 - 17.8.1 Use of Data from Metabolism and Pharmacokinetic Studies, 404
- 17.9 Physiologically Based Pharmacokinetic (PBPK) Modeling, 404
- 17.10 Points to Consider, 405
- 17.11 Biologically Derived Materials, 406
 - 17.11.1 Immunoassay Methods, 407
- 17.12 Points to Consider, 410
- References, 410

18 Safety Pharmacology 413

- 18.1 Regulatory Requirements, 414
- 18.2 Study Designs and Principles, 415
- 18.3 Organ System-Specific Tests, 416
 - 18.3.1 General Considerations in Selection and Design of Safety Pharmacology Studies, 416
 - 18.3.2 Studies on Metabolites, Isomers, and Finished Products, 416
- 18.4 Cardiovascular, 416
 - 18.4.1 Hemodynamics, ECG, and Respiration in Anesthetized Dogs or Primates, 417
 - 18.4.2 Cardiac Conduction Studies, 417
 - 18.4.3 Conscious Dog, Primate, or Minipig Telemetry Studies, 417
 - 18.4.4 Six-Lead ECG Measurement in the Conscious Dog and Minipig, 417
 - 18.4.5 Systems for Recording Cardiac Action Potentials, 418
 - 18.4.6 Special Case (and Concern): QT Prolongation, 418
 - 18.4.7 Some Specific Techniques Which Can Be Employed, 419
 - 18.4.8 Relevance of hERG to QT Prolongation, 419
- 18.5 Central Nervous System, 419
 - 18.5.1 Isolated Tissue Assays, 420
 - 18.5.2 Electrophysiology Methods, 421
 - 18.5.3 CNS Function: Electroencephalography, 421
 - 18.5.4 Neurochemical and Biochemical Assays, 421
- 18.6 Respiratory/Pulmonary System, 422
 - 18.6.1 Design of Respiratory Function Safety Studies, 425
 - 18.6.2 Capnography, 426

- 18.7 Secondary Organ System, 427
 - 18.7.1 Gastric Emptying Rate and Gastric pH Changes: A New Model, 427
- 18.8 Renal Function Tests, 428
- 18.9 Summary, 428
- References, 428

19 Special Concerns for the Preclinical Evaluation of Biotechnology Products 433

- 19.1 Regulation, 436
- 19.2 Preclinical Safety Assessment, 437
- 19.3 Recombinant DNA Technology, 439
 - 19.3.1 General Safety Issues, 440
 - 19.3.2 Specific Toxicological Concerns, 440
- 19.4 Immunogenicity/Allergenicity, 440
- 19.5 Monoclonal Antibody Technology, 441
 - 19.5.1 Toxicological Concerns with Monoclonal Antibodies, 442
- 19.6 Bioprocess Technology, 446
- 19.7 Gene Therapy Products, 446
 - 19.7.1 Vectors, 447
 - 19.7.2 Studies to Support the First Dose in Man, 447
 - 19.7.3 Distribution of the Gene and Gene Product, 447
 - 19.7.4 Studies to Support Multiple Doses in Humans, 447
 - 19.7.5 Unnecessary Studies, 448
 - 19.7.6 *Ex Vivo* Procedures, 448
 - 19.7.7 Change of Gene or Vector, 448
 - 19.7.8 Change of Route, 448
 - 19.7.9 Insertional Mutagenesis, 448
- 19.8 Vaccines, 449
 - 19.8.1 Approaches to Vaccination, 449
 - 19.8.2 Genetic Engineering and Vaccine Development, 450
- 19.9 Special Challenges, 452
 - 19.9.1 Purity and Homology, 453
 - 19.9.2 Immunogenicity, 453
- 19.10 Planning a Safety Evaluation Program, 454
 - 19.10.1 The Producing System, 454
 - 19.10.2 The Process, 454
 - 19.10.3 The Product, 455
 - 19.10.4 Biology of Bioengineered Products, 455
 - 19.10.5 Animal Models, 455
 - 19.10.6 Study Design, 457
 - 19.10.7 Frequency and Route of Administration, 458
 - 19.10.8 Duration, 458
 - 19.10.9 Special Toxicity Testing, 458
 - 19.10.10 Program Design Considerations, 458
- 19.11 Challenges: Biosimilars, 458
- References, 459

20 Safety Assessment of Inhalant Drugs and Dermal Route Drugs 461

- 20.1 Inhaled Therapeutics, 461
- 20.2 The Pulmonary System, 461
- 20.3 Penetration and Absorption of Inhaled Gases and Vapors, 462
- 20.4 Deposition of Inhaled Aerosols, 463
- 20.5 Absorption and Clearance of Inhaled Aerosols, 464
- 20.6 Pharmacokinetics and Pharmacodynamics of Inhaled Aerosols, 464

| | | |
|-----------|---|------------|
| 20.7 | Methods for Safety Assessment of Inhaled Therapeutics, 465 | |
| 20.8 | Parameters of Toxicity Evaluation, 467 | |
| 20.8.1 | The Inhaled “Dose”, 467 | |
| 20.8.2 | The Dose–Response Relationship, 468 | |
| 20.8.3 | Exposure Concentration versus Response, 469 | |
| 20.8.4 | Product of Concentration and Duration (<i>Ct</i>) versus Responses, 469 | |
| 20.8.5 | Units for Exposure Concentration, 469 | |
| 20.9 | Inhalation Exposure Techniques, 470 | |
| 20.10 | The Utility of Toxicity Data, 473 | |
| 20.11 | Formulation and Potential Mucosal Damage, 473 | |
| 20.11.1 | Methods to Assess Irritancy and Damage, 473 | |
| 20.12 | Therapeutic Drug Delivery by the Dermal Route, 474 | |
| | References, 476 | |
| 21 | Special Case Products: Imaging Agents | 483 |
| 21.1 | Introduction, 483 | |
| 21.2 | Imaging Agents, 483 | |
| 21.2.1 | Contrast Agents, 484 | |
| 21.2.2 | Diagnostic Radiopharmaceuticals, 484 | |
| 21.2.3 | Medical Imaging Agent Characteristics Relevant to Safety, 485 | |
| 21.2.4 | Performance of Nonclinical Safety Assessments, 485 | |
| | References, 487 | |
| 22 | Special Case Products: Drugs for Treatment of Cancer | 489 |
| 22.1 | Introduction, 489 | |
| 22.1.1 | Dose Conversions: Perspective, 493 | |
| 22.1.2 | The Use of the mg/m ² Dose Unit, 493 | |
| | References, 493 | |
| 23 | Pediatric Product Safety Assessment (2006 Guidance, Including Juvenile Toxicology) | 495 |
| 23.1 | Introduction, 495 | |
| 23.1.1 | Scope of Nonclinical Safety Evaluation, 497 | |
| 23.1.2 | Timing of Juvenile Animal Studies in Relation to Clinical Testing, 497 | |
| 23.2 | Issues to Consider Regarding Juvenile Animal Studies, 498 | |
| 23.2.1 | Developmental Stage of Intended Population, 498 | |
| 23.2.2 | Evaluating Data to Determine When Juvenile Animal Studies Should Be Used, 498 | |
| 23.2.3 | Considering Developmental Windows When Determining Duration of Clinical Use, 498 | |
| 23.2.4 | Timing of Exposure, 498 | |
| 23.2.5 | Selection of Study Models, 499 | |
| 23.3 | General Considerations in Designing Toxicity Studies in Juvenile Animals, 499 | |
| 23.4 | Study Designs and Considerations, 500 | |
| | References, 501 | |
| 24 | Use of Imaging, Imaging Agents, and Radiopharmaceuticals in Nonclinical Toxicology | 503 |
| 24.1 | Introduction, 503 | |
| 24.1.1 | Multimodality Imaging Techniques, 504 | |
| 24.1.2 | Dynamic Molecular Imaging Techniques, 504 | |

- 24.2 X-ray, 505
 - 24.2.1 Angiography, 505
- 24.3 Positron Emission Tomography (PET), 505
- 24.4 Single-photon Emission Computed Tomography (SPECT), 505
- 24.5 Computed Tomography (CT), 506
- 24.6 Magnetic Resonance Imaging (MRI), 506
- 24.7 Optical Imaging, 507
- 24.8 Ultrasound, 508
 - 24.8.1 Echocardiography, 508
- 24.9 Nanoparticle Contrast Agents, 509
- 24.10 Radiopharmaceuticals, 509
- 24.11 Applications of Preclinical Imaging in Laboratory Animals, 509
 - 24.11.1 Molecular Imaging as an ADME Platform in Drug Screen, 509
 - 24.11.2 Preclinical Imaging in Oncology, 510
 - 24.11.3 Preclinical Imaging of CNS Disease, 514
 - 24.11.4 Preclinical Imaging of Autoimmune Disease, 514
 - 24.11.5 Imaging Animal Model of Infectious Disease, 515
 - 24.11.6 Preclinical Imaging of Cardiac Disease, 515
- 24.12 Nonclinical Safety Assessment for Imaging Agents, 515
- 24.13 Radiopharmaceuticals, 517
- 24.14 Nonclinical Late Radiation Toxicity Studies, 519
 - 24.14.1 Study Goals, 519
- 24.15 Study Design, 519
 - 24.15.1 Good Laboratory Practices, 519
 - 24.15.2 Species Selection, 519
 - 24.15.3 Timing of Study, 519
 - 24.15.4 General Study Design, 519
 - 24.15.5 Dose Levels, 520
 - 24.15.6 Clinical Pathology, 520
 - 24.15.7 Necropsy and Histopathology, 520
- References, 520

25 Occupational Toxicology in the Pharmaceutical Industry 523

- 25.1 Introduction, 523
- 25.2 Occupational Toxicology versus Drug Safety Evaluation, 523
- 25.3 Regulatory Pressures in the United States
and the European Community, 525
- 25.4 Organizational Structure, 526
- 25.5 Activities, 527
 - 25.5.1 Data Evaluation and Dissemination, 527
 - 25.5.2 Data Development, 528
 - 25.5.3 Occupational Exposure Limits (OELs), 531
 - 25.5.4 Hazard Assessment, 531
 - 25.5.5 Employee Training, 532
- 25.6 Conclusion, 534
- References, 534

26 Strategy and Phasing for Nonclinical Drug Safety Evaluation
in the Discovery and Development of Pharmaceuticals 537

- 26.1 Introduction, 537
- 26.2 Regulatory Requirements, 539
- 26.3 Essential Elements of Project Management, 542
- 26.4 Screens: Their Use and Interpretation in Safety Assessment, 544
 - 26.4.1 Characteristics of Screens, 545

- 26.5 Strategy and Phasing, 546
- 26.6 Critical Considerations, 550
- 26.7 Special Cases in Safety Assessment, 551
- 26.8 Summary, 551
- References, 551

27 The Application of *In Vitro* Techniques in Drug Safety Assessment 553

- 27.1 Introduction, 553
- 27.2 *In Vitro* Testing in Pharmaceutical Safety Assessment, 555
- 27.3 Defining Testing Objective, 558
 - 27.3.1 Objectives behind Data Generation and Utilization, 558
- 27.4 Test Systems: Characteristics, Development, and Selection, 558
- 27.5 *In Vitro* Models, 559
- 27.6 Lethality, 560
 - 27.6.1 Ocular Irritation, 564
 - 27.6.2 Dermal Irritation, 564
 - 27.6.3 Irritation of Parenterally Administered Pharmaceuticals, 565
 - 27.6.4 Sensitization and Photosensitization, 566
 - 27.6.5 Phototoxicity and Photosensitization, 567
 - 27.6.6 Developmental Toxicity, 568
 - 27.6.7 Target Organ Toxicity Models, 568
- 27.7 *In Silico* Methods, 572
- 27.8 The Final Frontier and Barrier: Regulatory Acceptance, 573
- 27.9 Summary, 573
- References, 575
- Further Reading, 581

28 Evaluation of Human Tolerance and Safety in Clinical Trials: Phase I and Beyond 583

- 28.1 The Pharmaceutical Clinical Development Process and Safety, 583
 - 28.1.1 Pharmacokinetics, 589
 - 28.1.2 Safety of Clinical Trial Subjects, 591
- 28.2 Limitations on/of Clinical Trials, 598
- 28.3 The Clinical Trial Process, 598
 - 28.3.1 Development of an Application Unrelated to Original Approved Use, 601
- 28.4 Institutional Review Boards (IRBS)/Ethics Committees in the Clinical Trial Process, 602
 - 28.4.1 Legal Authority and Responsibilities for IRBs, 602
 - 28.4.2 Duties of IRBs, 603
 - 28.4.3 Informed Consent, 603
- 28.5 Drug Formulations and Excipients, 604
 - 28.5.1 Route of Administration, 605
- 28.6 Phase I Designs, 605
 - 28.6.1 First Administration: Single Dose Escalating (SDE), 606
 - 28.6.2 First Administration in Humans: Multiple Dose Escalating (MDE), 608
- 28.7 Clinical Trial Safety Indicators, 609
 - 28.7.1 Overall Approach to Assessing Safety, 609
 - 28.7.2 Precautions, 610
 - 28.7.3 Clinical Chemistry, 613
 - 28.7.4 Urinalysis, 614
 - 28.7.5 Urine Screens, 614

- 28.7.6 Identifying New Diagnostic Laboratory Tests, 614
- 28.7.7 Ophthalmological Examination, 614
- 28.7.8 Dermatological Examinations, 614
- 28.7.9 Cardiovascular Safety, 615
- 28.7.10 Deaths in Clinical Trials, 615
- 28.7.11 Behavioral Rating Scales, Performance, Personality, and Disability Tests, 616
- 28.7.12 Adult Behavioral Rating Scales, 616
- 28.7.13 Pediatric Behavioral Rating and Diagnostic Scales, 618
- 28.7.14 Psychometric and Performance Tests, 619
- 28.7.15 Personality Tests, 621
- 28.8 Assessment of Unwanted Drug Effects, 621
 - 28.8.1 Separation of Adverse Reactions from Placebo Reactions, 621
- References, 626

29 Postmarketing Safety Evaluation: Monitoring, Assessing, and Reporting of Adverse Drug Responses (ADRs) 629

- 29.1 Causes of Safety Withdrawals, 637
- 29.2 Regulatory Requirements, 638
 - 29.2.1 The 15-Day Report versus the US Periodic Report, 639
- 29.3 Management of ADR and ADE Data, 641
 - 29.3.1 Sources of Data, 641
 - 29.3.2 Clinical Trials, 641
 - 29.3.3 Postmarketing Surveillance Studies, 641
 - 29.3.4 Spontaneous Reports, 641
 - 29.3.5 Literature, 642
 - 29.3.6 Searching for ADRs in the Literature, 642
 - 29.3.7 Information Required for Reports, 642
 - 29.3.8 Adverse Drug Reaction Forms and Form Design, 642
 - 29.3.9 Computerization of Drug Safety Data: Data Collection and Input, 644
 - 29.3.10 Medical and Drug Terminology, 644
 - 29.3.11 Dictionaries, 645
 - 29.3.12 Medical Term Coding Dictionaries, 645
 - 29.3.13 Medical Dictionary for Regulatory Activities, 645
 - 29.3.14 Periodic Reports, 646
- 29.4 Causality Assessment, 647
 - 29.4.1 Aims of Causality Assessment, 647
- 29.5 Courses of Corrective Action, 647
- 29.6 Legal Consequences of Safety Withdrawal, 648
 - 29.6.1 FDA Tools for Risk Management, 648
 - 29.6.2 Tier 1: Mandatory Studies, 649
 - 29.6.3 Tier 2: Labeling and Assessment, 649
 - 29.6.4 Tier 3: Enhanced Communication, 650
 - 29.6.5 Tier 4: Safe Use Restriction Defined by Provider, 650
 - 29.6.6 Tier 5: Safe Use Restriction Defined by Patient, 651
- References, 651

30 Statistics in Pharmaceutical Safety Assessment 653

- 30.1 Introduction, 653
 - 30.1.1 Bias and Chance, 655
 - 30.1.2 Hypothesis Testing and Probability (p) Values, 655
 - 30.1.3 Multiple Comparisons, 656
 - 30.1.4 Estimating the Size of the Effect, 656

- 30.1.5 Functions of Statistics, 657
- 30.1.6 Descriptive Statistics, 658
- 30.2 Experimental Design, 659
 - 30.2.1 Choice of Species and Strain, 659
 - 30.2.2 Sampling, 659
 - 30.2.3 Dose Levels, 660
 - 30.2.4 Number of Animals, 660
 - 30.2.5 Duration of the Study, 660
 - 30.2.6 Stratification, 661
 - 30.2.7 Randomization, 661
 - 30.2.8 Adequacy of Control Group, 661
- 30.3 Data Recording, 664
- 30.4 Generalized Methodology Selection, 665
- 30.5 Statistical Analysis: General Considerations, 665
 - 30.5.1 Variables to Be Analyzed, 665
 - 30.5.2 Combination of Observations
(Such as Pathological Conditions), 667
 - 30.5.3 Taking Severity into Account, 668
 - 30.5.4 Using Simple Methods Which Avoid Complex
Assumptions, 668
 - 30.5.5 Using All the Data, 668
 - 30.5.6 Combining, Pooling, and Stratification, 668
 - 30.5.7 Trend Analysis, Low-Dose Extrapolation,
and NOEL Estimation, 669
 - 30.5.8 Need for Age Adjustment, 671
 - 30.5.9 Need to Take Context of Observation into Account, 672
 - 30.5.10 Experimental and Observational Units, 672
 - 30.5.11 Missing Data, 672
 - 30.5.12 Use of Historical Control Data, 673
 - 30.5.13 Methods for Data Examination and Preparation, 673
 - 30.5.14 Scattergram, 673
 - 30.5.15 Bartlett's Test for Homogeneity of Variance, 675
 - 30.5.16 Statistical Goodness-of-Fit Tests, 676
 - 30.5.17 Randomization, 677
 - 30.5.18 Transformations, 677
 - 30.5.19 Exploratory Data Analysis, 678
- 30.6 Hypothesis Testing of Categorical and Ranked Data, 679
 - 30.6.1 Fisher's Exact Test, 679
 - 30.6.2 2×2 Chi-Square, 680
 - 30.6.3 $R \times C$ Chi-Square, 680
 - 30.6.4 Wilcoxon Rank-Sum Test, 681
 - 30.6.5 Distribution-Free Multiple Comparison, 682
 - 30.6.6 Mann-Whitney U Test, 682
 - 30.6.7 Kruskal-Wallis Nonparametric ANOVA, 683
 - 30.6.8 Log-Rank Test, 683
- 30.7 Hypothesis Testing: Univariate Parametric Tests, 684
 - 30.7.1 Student's t -Test (Unpaired t -Test), 685
 - 30.7.2 Cochran t -Test, 685
 - 30.7.3 F -Test, 686
 - 30.7.4 Analysis of Variance (ANOVA), 686
 - 30.7.5 Post Hoc Tests, 687
 - 30.7.6 Duncan's Multiple Range Test, 687
 - 30.7.7 Groups with Equal Number of Data ($N_1 = N_2$), 687
 - 30.7.8 Groups with Unequal Number of Data ($N_1 \neq N_2$), 688

| | |
|---------|---|
| 30.7.9 | Scheffe's Multiple Comparisons, 688 |
| 30.7.10 | Dunnett's <i>t</i> -Test, 688 |
| 30.7.11 | Williams' <i>t</i> -Test, 689 |
| 30.7.12 | Analysis of Covariance, 689 |
| 30.7.13 | Modeling, 690 |
| 30.7.14 | Linear Regression, 691 |
| 30.7.15 | Probit/Log Transforms and Regression, 691 |
| 30.7.16 | Nonlinear Regression, 692 |
| 30.7.17 | Correlation Coefficient, 693 |
| 30.7.18 | Kendall's Coefficient of Rank Correlation, 694 |
| 30.7.19 | Trend Analysis, 694 |
| 30.8 | Methods for the Reduction of Dimensionality, 694 |
| 30.8.1 | Classification, 695 |
| 30.8.2 | Statistical Graphics, 696 |
| 30.8.3 | Multidimensional and Nonmetric Scaling, 697 |
| 30.8.4 | Cluster Analysis, 699 |
| 30.8.5 | Fourier or Time Analysis, 699 |
| 30.8.6 | Life Tables, 700 |
| 30.9 | Meta-Analysis, 701 |
| 30.9.1 | Selection of the Studies to Be Analyzed, 701 |
| 30.9.2 | Pooled (Quantitative) Analysis, 701 |
| 30.9.3 | Methodological (Qualitative) Analysis, 702 |
| 30.10 | Bayesian Inference, 702 |
| 30.10.1 | Bayes' Theorem and Evaluation of Safety Assessment Studies, 702 |
| 30.10.2 | Bayes' Theorem and Individual Animal Evaluation, 703 |
| 30.11 | Data Analysis Applications in Safety Assessment Studies, 704 |
| 30.11.1 | Body and Organ Weights, 705 |
| 30.11.2 | Clinical Chemistry, 706 |
| 30.11.3 | Hematology, 706 |
| 30.11.4 | Histopathological Lesion Incidence, 706 |
| 30.11.5 | Carcinogenesis, 707 |
| | References, 708 |

31 Combination Products: Drugs and Devices 711

| | |
|--------|----------------------------|
| 31.1 | Combination Products, 711 |
| 31.1.1 | Historical Background, 711 |
| 31.1.2 | Future Trends, 712 |

References, 720

32 Qualification of Impurities, Degradants, Residual Solvents, Metals, and Leachables in Pharmaceuticals 721

| | |
|------|-----------------------------------|
| 32.1 | Impurities, 721 |
| 32.2 | Residual Solvents, 726 |
| 32.3 | Extractables and Leachables, 727 |
| 32.4 | Residual Metals and Elements, 728 |

References, 730

33 Tissue, Cell, and Gene Therapy 731

| | |
|--------|---|
| 33.1 | Safety Assessment of Cell Therapy (CT) Products, 732 |
| 33.1.1 | Recommendations for General Preclinical Program Design, 732 |
| 33.1.2 | Model Species Selection, 732 |
| 33.1.3 | Selection of Animal Models of Disease/Injury, 732 |

| | | |
|---|--|------------|
| 33.1.4 | Information Describing Limitations of Potential Animal Model(s), | 733 |
| 33.1.5 | Information Supporting the Choice of Animal Model(s), | 733 |
| 33.1.6 | Proof-of-Concept (POC) Studies, | 733 |
| 33.1.7 | Toxicology Studies, | 734 |
| 33.1.8 | Product Delivery Considerations, | 735 |
| 33.1.9 | Study Designs, | 736 |
| 33.1.10 | CT Products with Implantable Scaffolds, | 738 |
| 33.2 | Nonclinical Safety Assessment of Gene Therapy Products (GTPS), | 738 |
| 33.2.1 | CBER, | 738 |
| 33.2.2 | NIH, | 738 |
| 33.2.3 | Study Designs, | 739 |
| 33.2.4 | <i>Ex Vivo</i> Genetically Modified Cells, | 740 |
| 33.2.5 | Biodistribution Considerations, | 740 |
| 33.3 | Definitions, | 741 |
| | References, | 742 |
| Appendix A: Selected Regulatory and Toxicological Acronyms | | 743 |
| Appendix B: Definition of Terms and Lexicon of “Clinical” Observations in Nonclinical (Animal) Studies | | 745 |
| Appendix C: Notable Regulatory Internet Addresses | | 749 |
| Appendix D: Glossary of Terms Used in the Clinical Evaluation of Therapeutic Agents | | 755 |
| Appendix E: Common Vehicles for the Nonclinical Evaluation of Therapeutic Agents | | 759 |
| Appendix F: Global Directory of Contract Pharmaceutical Toxicology Labs | | 857 |
| INDEX | | 879 |

PREFACE

The third edition of *Drug Safety Evaluation* is a complete revision of the second edition which maintains the central objective of presenting an all-inclusive practical guide for those who are responsible for ensuring the safety of drugs and biologics to patients and shepherding valuable candidates to market, healthcare providers, those involved in the manufacture of medicinal products, and all those who need to understand how the safety of these products is evaluated. The many changes in regulatory requirements, pharmaceutical development, and technology have required both extensive revision to every chapter and the addition of four new chapters.

This practical guide presents a road map for safety assessment as an integral part of the development of new drugs and therapeutics. Individual chapters also address specific approaches to evaluation hazards, including problems that are encountered and their solutions. Also covered are the scientific and philosophical bases for evaluation of specific concerns (e.g., carcinogenicity, development toxicity, etc.) to provide both understanding and guidance for approaching new problems. *Drug Safety Evaluation* is aimed specifically at the pharmaceutical and biotechnology industries. It not only addresses the general cases for safety evaluation of small and large molecules but also all of the significant major subcases: imaging agents, dermal and inhalation route drugs, vaccines, and gene therapy products. It is hoped that the approaches and methodologies presented here will show a utilitarian yet scientifically valid path to the everyday challenges of safety evaluation and the problem solving that is required in drug discovery and development.

Shayne Cox Gad
Raleigh, North Carolina

ABOUT THE AUTHOR

Shayne Cox Gad, B.S. (Whittier College, Chemistry and Biology, 1971) and Ph.D. in Pharmacology/Toxicology (Texas, 1977) DABT, ATS, is the principal of Gad Consulting Services, a 24-year-old consulting firm with seven employees and more than 450 clients (including 200 pharmaceutical companies in the United States and 50 overseas). Prior to this, he served in director-level and above positions at Searle, Synergen, and Becton Dickinson. He has published 48 books and more than 350 chapters, articles, and abstracts in the fields of toxicology, statistics, pharmacology, drug development, and safety assessment. He has more than 39 years of broad-based experience in toxicology, drug and device development, statistics, and risk assessment. He has specific expertise in neurotoxicology, *in vitro* methods, cardiovascular toxicology, inhalation toxicology, immunotoxicology, and genotoxicology. Past president of the American College of Toxicology, the Roundtable of Toxicology Consultants, and three of SOT's specialty sections. He has direct involvement in the preparation of INDs (110 successfully to date), NDA, PLA, ANDA, 501(k), IDE, CTD, clinical databases for phase 1 and 2 studies, and PMAs. He has consulted for FDA, EPA, and NIH and has trained reviewers and been an expert witness for FDA. He has also conducted the triennial toxicology salary survey as a service to the profession for the last 27 years.

Dr. Shayne Cox Gad is also a retired Navy line officer.

THE DRUG DEVELOPMENT PROCESS AND THE GLOBAL PHARMACEUTICAL MARKETPLACE

1.1 INTRODUCTION

Pharmaceuticals are a global industry, grossing \$839 billion (US dollars) in 2014. They are developed to benefit (and sell to) individuals and societies worldwide. Their effectiveness and costs affect, directly or indirectly, all of us.

This third edition focuses (as its predecessors did) on the assessment of the safety of new drugs. In the broadest sense, this means it must address not only the traditional “small molecules” that have dominated the field for the last century and the large therapeutic molecules derived from biotechnology sources but also vaccines, biologics such as blood and blood products, cell therapies, and excipients. The globalization of the regulation of the safety, efficacy, and manufacture of pharmaceutical products comes from the success of the International Conference on Harmonisation (ICH) process. But, as will be seen, the same globalization of the industry and continuous advances of science have also led to market diversification of the types and use of drugs, and with this, regulatory drug safety evaluation requirements continue to fragment, which has made things more complex rather than simpler (Alder and Zbinden, 1988; Gad, 2011).

1.2 THE MARKETPLACE

The world marketplace for drugs is large, although the majority of sales are in the three regions: in 2013 about 39% of the pharmaceutical market resided in the United States, 24% in Europe, 15% in Japan, and 22% in emerging markets. The balance of sales is spread across the globe. This does not mean, however, that marketing applicants can or should ignore the requirements of other countries, for example,

Indonesia. Approval processes in these countries can, at times, be as rigorous as in any other regulatory authority domain.

Pharmaceuticals in all their forms compete today as part of a global market, though one which serves (and is available to) different parts of the world’s population to varying extents.

The term “pharmaceuticals” is here used in the broadest sense of man-made therapeutics: small molecules, large protein moieties, vaccines, blood products, and, as must be, their attendant components (excipients, impurities, and all) to different degrees and in different types of products.

According to the IMS 2013 global pharmaceutical market and therapy forecast, the global market for regulated drugs (as differentiated from dietary supplements, herbal products, and nutraceuticals) is estimated to be some \$870 billion in 2014 (US dollars). In 2015, there were 109 individual products with annual sales in excess of \$1 billion (i.e., “blockbusters”) which have tended to be the focus of pharmaceutical development until recently and the impending demise of patents on which is changing the industry (Table 1.1).

This concentration of total sales in a limited number of products (e.g., there are currently more than 22 000 approved prescription drugs in the United States) is widely held to have distorted the therapeutic aspects of new drug development but is now starting to undergo change (back to) a paradigm that looks at a decreased emphasis on the billion dollar “blockbuster” drugs.

Widely misunderstood is the extent and diversity of the pharmaceutical R&D sector. While precise numbers are unavailable (and meaningless, as companies are continuously being started, merged, or going out of business, though the overall trend is to increased numbers), best estimates place the

TABLE 1.1 Top 20 Selling Pharmaceuticals (2013)

| Rank | Drug | Current Manufacturer | Total Sales (USD) | % Change from 2012 | Primary Disease/Medical Use | Route(s) |
|------|-----------------|---|-------------------|--------------------|--|------------------|
| 1 | Abilify | Otsuka Pharmaceutical Co. Ltd | 6 293 801 | +11 | Psychotic conditions, major depressive disorder | Oral, injection |
| 2 | Nexium | Astra Zeneca Pharmaceuticals, LP | 5 974 550 | +5.4 | GERD, Zollinger-Ellison syndrome, erosive esophagitis, other conditions associated with excessive stomach acid | Oral, parenteral |
| 3 | Humira | AbbVie, Inc. | 5 428 479 | +20.75 | Inflammation (arthritis, ankylosing spondylitis, plaque psoriasis, and hidradenitis suppurativa, Crohn's disease or ulcerative colitis after other methods fail) | Injection |
| 4 | Crestor | Astra Zeneca Pharmaceuticals, LP | 5 195 930 | +8.3 | Cholesterol | Oral |
| 5 | Cymbalta | Eli Lilly and Company | 5 083 111 | +12 | Depression, Anxiety | Oral |
| 6 | Advair Diskus | GlaxoSmithKline | 4 981 108 | +7.3 | Asthma | Inhalation |
| 7 | Enbrel | Amogen, Inc. | 4 585 701 | +12.9 | Arthritis, or ankylosing spondylitis, plaque psoriasis and polyarticular juvenile idiopathic arthritis | Injection |
| 8 | Remicade | Centocor Ortho Biotech, Inc. | 3 980 556 | +6.5 | Arthritis, ulcerative colitis, Crohn's disease, ankylosing spondylitis, plaque psoriasis | IV |
| 9 | Copaxone | Teva Pharmaceuticals | 3 603 958 | +7.5 | Multiple Sclerosis | Injection |
| 10 | Neulasta | Amogen, Inc. | 3 472 969 | +4.1 | Neutropenia caused by receiving chemotherapy | Injection |
| 11 | Rituxan | Genetech, Inc. (member of Roche group) | 3 208 525 | +2.5 | Non-Hodgkin's lymphoma or chronic lymphocytic leukemia | IV |
| 12 | Spiriva | Boehringer Ingelheim Pharmaceuticals, Inc | 2 943 778 | +8.5 | COPD, bronchitis, emphysema, asthma | Inhalation |
| 13 | Lantus Solostar | Sanofi (formerly Sanofi Aventis) | 2 926 949 | +29.5 | Diabetes | Injection |
| 14 | Atripla | Gilead Sciences, Inc. | 2 794 285 | +2.5 | HIV | Oral |
| 15 | Januvia | Merck & Co., Inc. | 2 770 995 | +9.8 | Type 2 Diabetes | Oral |
| 16 | Avastin | Genetech, Inc. (member of Roche group) | 2 617 373 | +2 | Brain tumor, certain types of cancers of the kidney, lung, colon, rectum, cervix, ovary, or fallopian tube. Cancer of the membrane lining the internal organs in the abdomen | IV |
| 17 | Lantus | Sanofi (formerly Sanofi Aventis) | 2 505 281 | +12 | Type 1 or type 2 diabetes | Injection |
| 18 | OxyContin | Purdue Pharma LP | 2 462 851 | -8.6 | Moderate to severe extended pain | Oral |
| 19 | Lyrica | Pfizer Inc. | 2 357 959 | +18.4 | Control of seizures, fibromyalgia, diabetic neuropathy, herpes zoster, post-herpetic neuralgia, or neuropathic pain associated with spinal cord injury. | Oral |
| 20 | Epogen | Amogen, Inc. | 2 206 624 | +5.5 | Anemia in patients with chronic kidney disease, HIV patients, and cancer patients receiving chemotherapy | Injection. IV |

Drugs.com (2014).

TABLE 1.2 Top 25 Drug Companies by sales (2014)

| Company | Pharma sales 2014 (\$ million) | % Change from 2013 |
|----------------------|-----------------------------------|-----------------------|
| Novartis | 47101 | -1 |
| Pfizer | 45708 | -5 |
| Roche | 39120 | 0 |
| Sanofi | 36437 | -2 |
| Merck & Co. | 36042 | -4 |
| Johnson & Johnson | 32313 | 15 |
| GlaxoSmithKline | 29580 | -11 |
| AstraZeneca | 26095 | 1 |
| Gilead Sciences | 24474 | 127 |
| Takeda | 20446 | 7 |
| AbbVie | 20207 | 8 |
| Amgen | 19327 | 6 |
| Teva | 18374 | 0 |
| Lilly | 17266 | -18 |
| Bristol-Myers Squibb | 15879 | -3 |
| Bayer | 15486 | 4 |
| Novo Nordisk | 15329 | 3 |
| Astellas | 14099 | 4 |
| Boehringer Ingelheim | 13830 | -12 |
| Actavis | 13062 | 51 |
| Otsuka | 11308 | 1 |
| Daiichi Sankyo | 10430 | -14 |
| Biogen Idec | 9398 | 41 |
| Baxter | 8831 | 6 |
| Merck KGaA | 7678 | -9 |

PMLive (2015).

number of companies directly involved in discovering and developing new drugs in the United States and Canada at about 3800, 10% of which are publicly traded. There are an equal number in Europe and significant numbers in many other parts of the world (Japan, China, Australia, India, and Israel, to name just a few other countries). While most of the public focuses on very large companies, such as those in Table 1.2, there are many more midsize and small companies.

Starting in 1984 with the Drug Price Competition and Patent Term Restoration Act (better known as the Hatch–Waxman Act), “doses” of small molecule drugs leaving the period of patent protection could be introduced into the marketplace by an ANDA-approved route—a much simpler and quicker route to market approval. Such generics constituted 86% of prescriptions in the United States by 2013, though their market share by sales (\$260 billion in 2012) is only 31% of revenues (Thayer, 2014).

One factor to consider in the regulatory requirements for early development of new therapeutic entities is the higher degree to which costs may present barriers to smaller, innovative companies. This is commonly overlooked by many who also do not recognize that such small companies (most of which fail) are the primary initial source of new therapeutics.

A second complicating factor in considering the “pharmaceutical” market sector is the diversity of products involved. The most basic expression of this is the division of drugs into “small molecules” (which currently constitute approximately two-thirds of both INDs—applications for clinical evaluation of a new drug in humans and 80% of current new drug approvals) and biotechnology products (which constitute the bulk of the remainder—biologics such as vaccines are increasing in importance). The challenges in both developing and assessing the safety of these are very different. As will also be seen, if one considers further division into therapeutic claim areas (oncology, anti-infectives, cardiovascular, CNS, etc.), the differences become even more marked. Most of what will be presented and discussed in this volume speaks to regulatory requirements for non-clinical safety assessment in the general case for either small molecules or protein therapeutics. It should be kept in mind that this general case development model never fully applies.

Additionally, there is now a significant hybrid area—combination products, which include both device and drug (small molecule or biologic) components. These will be addressed in a separate chapter of the book, though there is no single dedicated regulatory arm (such as a center within the FDA truly dedicated to only their regulation) in any major market country or such. For that reason, more exploration of regulatory considerations will be provided in the chapter on these products.

The extent of regulations and practices for drug approval causes pharmaceutical companies to spend an enormous amount of resources on developing applications, following different standards for preclinical and nonclinical programs for specific therapeutic areas, as well as time and resources to satisfy the regulatory processes for clinical trials. Because of the regulatory diversity that existed, representatives from the regulatory authorities and trade associations came together in the late 1980s and early 1990s to attempt at harmonizing the process for drug approvals. Clearly this was a daunting task. With time, however, the International Conference on Harmonisation of Technical Requirements for Registration of Pharmaceuticals for Human Use has become increasingly more effective. Fortunately, the abbreviation for this very long title is ICH. Japan, Europe, and the United States represent the major pharmaceutical market for the world, and these regions have the most influence on developments within ICH and tend to follow the guidance documents that are prepared. However, other countries (rest of the world (ROW)) follow the developments within ICH

and tend to follow the guidance offered by ICH. However, it remains important, when seeking for the registration of pharmaceuticals, to be aware of local country regulations. For example, China has become a major economic force in many aspects. Placement of pharmaceutical manufacturing facilities and the marketing of drugs in China may potentially represent a significant marketing advantage to companies. With this new market area in Asia, regulatory processes are being developed; sometimes it seems at the whim of the government. With time it is hoped that China will align itself more with the processes and guidance that have been developed by ICH, FDA, and other further developed countries.

1.3 HISTORY OF MODERN THERAPEUTICS

Although, prior to the nineteenth century, preventive medicine had made some spectacular advances, for example, through nutrition (scurvy), control of infectious diseases (such as small pox, polio, and tuberculosis) and public health through sanitation, and control of childbirth fever and surgical infections using antiseptic techniques, truly therapeutic medicine was virtually nonexistent until the end of the nineteenth century.

Oliver Wendell Holmes (a physician and US Supreme Court Justice) wrote in 1860: “.... I firmly believe that if the whole material medica, as now used, could be sunk to the bottom of the sea, it would be all the better for mankind—and the worse for the fishes.” While there were a few effective medicines—digitalis, extract of willow bark, and quinine, for example—on balance, Holmes was quite correct, medicines did more harm than good.

The first edition of the British Pharmacopoeia (1864), which listed 311 preparations, gives an idea of the state of therapeutics at the time. Of those listed, 187 preparations were plant-derived materials and only nine of which were purified substances. Most of the plant products—lemon juice, rose hips, yeasts, etc.—lacked any components we would now regard as therapeutically relevant, but some (digitalis, castor oil, ergot, colchicum) were pharmacologically active. Of the 311 preparations, 103 were truly synthetic inorganic chemicals such as iodine, ferrous sulfate, sodium bicarbonate, and toxic salts of bismuth, arsenic, lead, and mercury, with but a few synthetic chemicals (diethyl ether and chloroform). The remainders were miscellaneous materials and a few animal products, such as lard, cantharidin, and cochineal.

For the pharmaceutical industry, the transition to an actual industry and discipline occurred late in the nineteenth century when three essential technologies came together. These were the science of biomedicine (especially pharmacology), synthetic organic chemistry, and the development of a chemical industry in Europe, coupled with a medical supplies/products trade.

Science began to be applied wholeheartedly to medicine—as to almost every other aspect of life—only late in the nineteenth century. Among the most important milestones from the point of view of drug discovery was the elaboration in 1858 of cell theory. This tremendous reductionist leap of the cell theory gave biology—and the pharmaceutical industry—the fundamental scientific underpinning it required. It is only by thinking of living systems in terms of the function of their cells that one can begin to understand how molecules affect them.

A second milestone was the birth of pharmacology as a scientific discipline when the world’s first Pharmacological Institute was set up in 1874 at Dorpat (then in Germany—now in Estonia) by Rudolf Buchheim—literally by Buchheim himself, as the Institute was in his own house and funded by his estate. This was advanced by pioneers, such as Magendie and Claude Bernard, and linked to therapeutics.

Another vital spark on this road came with Louis Pasteur’s germ theory of disease, proposed in Paris in 1878. A chemist in training, Pasteur’s initial interest was in the process of fermentation of wine and beer and the souring of milk. He showed, famously, that airborne infection was the underlying cause and concluded that the air was actually alive with microorganisms. Particular types, he argued, were pathogenic to humans and accounted for many forms of disease including anthrax, cholera, and rabies. Pasteur successfully introduced several specific immunization procedures to give protection against infectious diseases. Robert Koch, Pasteur’s rival and near-contemporary, clinched the infection theory by observing anthrax and other bacilli in the blood of infected animals.

The founder of chemotherapy—some would say the father of molecular pharmacology—was Paul Ehrlich. He invented “vital staining”—staining by dyes injected into living animals—and described how the chemical properties of the dyes, particularly their acidity and lipid solubility, influenced the distribution of dye to particular tissues and cellular structures. Thence came the idea of specific binding of molecules to particular cellular components. This led not only to Ehrlich’s study of chemotherapeutic agents but also became the basis of pharmacological thinking to the present day. “Receptors” and “magic bullets” were Ehrlich’s terms, though he envisaged receptors as targets for toxins rather than physiological mediators. Working in Koch’s Institute, Ehrlich developed diphtheria antitoxin for clinical use, and put forward a theory of antibody action based on specific chemical recognition of microbial molecules, a work for which he won the 1908 Nobel Prize.

The first synthetic organic chemicals to be used for medical purposes were not therapeutic agents at all but rather anesthetics. Diethyl ether (“sweet oil of vitriol”) was first made and described in 1540. Early in the nineteenth century, it and nitrous oxide (prepared by Sir Humphrey Davy in

1799 and found—by self-experimentation—to have stupor-inducing properties) had their usefulness as surgical anesthetics demonstrated only in the 1840s, by which time chloroform had also made its appearance. Synthetic chemistry at the time could deal only with very simple molecules, made by recipe rather than rational understanding of the underlying chemistry reasons, as our understanding of chemical processes and molecular structure was still in its infancy. The first therapeutic drug to truly come from synthetic chemistry was amyl nitrite, prepared in 1859 by Guthrie and used in treating angina by Brunton in 1864. This was the first example of a drug born in a recognizably “modern” way through the application of synthetic chemistry, physiology, and clinical medicine. This was a landmark indeed, for it was nearly 40 years before synthetic chemistry made any further significant contribution to therapeutics and not until well into the twentieth century that physiological and pharmacological knowledge began to be applied to the invention of new drugs.

During the latter half of the nineteenth century, the foundations of synthetic organic chemistry were laid, the impetus coming from work on aniline, a copious by-product of the coal-tar industry, with the discovery of how to produce a purple dye. This discovery gave birth to the synthetic dyestuffs industry, which played a major part in establishing the commercial potential of synthetic organic chemistry—a technology which later became the underpinning of the evolving pharmaceutical industry for the next century. A systematic approach to organic synthesis went hand in hand with improved understanding of chemical structure.

Despite the limited efficacy of the pharmaceutical preparations that were available in the nineteenth century (“patent medicines”), the pharmacists trade flourished; then, as now, physicians felt themselves obligated to issue prescriptions to satisfy the expectations of their patients for some therapeutic action—or at least cause for hope. Early in the nineteenth century, a few enterprising chemists undertook the task of isolating the active substances from these plant extracts. The trend began with Friedrich Serturner, a junior apothecary in Westphalia, who in 1805 isolated and purified morphine, barely surviving a test of its potency on himself. This was the first “alkaloid,” so named because of its ability to neutralize acids and form salts. This discovery in turn led to the isolation of other plant alkaloids, including strychnine, caffeine, and quinine. The recognition that medicinal plants owed their properties to their individual chemical constituents, rather than to some intangible property associated with their living nature, marks a critical point in the history of the pharmaceutical industry which can be recognized as the point of origin of two of the three roads from which the industry grew—namely, the beginnings of the “industrialization” of the pharmaceutical trade. This revelation hinted at the future and the possibility of making drugs artificially.

The first local apothecary business to move into large-scale production and marketing of pharmaceuticals was the old-established Darmstadt firm Merck founded in 1668. This development, in 1827, was stimulated by the advances in purification of natural products. Merck was closely followed in this astute business move by other German- and Swiss-based apothecary businesses, giving rise to some which later also became giant pharmaceutical companies, such as Schering and Boehringer. The American pharmaceutical industry emerged in the middle of the nineteenth century; Squibb began in 1858 with ether as its main product. The move into pharmaceuticals was also followed by several chemical companies such as Bayer, Hoechst, Agfa, Sandoz, Geigy, and others which began as dyestuffs manufacturers. The dyestuffs industry at that time was also based largely on plant products, which had to be refined and were sold in relatively small quantities, so the commercial parallels with the pharmaceutical industry were plain.

After 1870, with the crucial discovery by Kekule of the structure of benzene, the dyestuffs industry turned increasingly to synthetic chemistry as a source of new compounds, starting with aniline-based dyes. A glance through any modern pharmacopeia will show the overwhelming preponderance of synthetic aromatic compounds, based on the benzene ring structure, among the list of useful drugs. Understanding the nature of aromaticity was critical.

Thus, the beginnings of the pharmaceutical industry as we now know it, at the latest, date from about third of the 1800s, with origins in the apothecaries and patent medicine trades on the one hand and the dyestuffs industry on the other. Unfortunately, these enterprises had rather few effective products to sell (mainly inorganic compounds of varying degrees of toxicity and others most charitably described as concoctions).

Entering the 1900s, synthetic drugs had been made and tested, including the “antipyretics” and various central nervous system depressants. Chemical developments based on chloroform had produced chloral hydrate, the first nonvolatile CNS depressant, which was in clinical use for many years as a hypnotic drug. Independently, various compounds based on urea were found to act similarly, and von Mering followed this lead to produce the first barbiturate, barbitone (since renamed barbital), which was introduced in 1903 by Bayer and gained widespread clinical use as a hypnotic, tranquilizer, and antiepileptic drug—the first blockbuster. Barbitone and procaine were triumphs for chemical ingenuity but owed little or nothing to physiology or indeed pharmacology. The physiological site or sites of action of barbiturates remain unclear to this day, and their mechanism of action at the molecular level was unknown until the 1980s.

The pattern of drug discovery driven by synthetic chemistry—with biology often struggling to keep up—became the established model in the early part of the twentieth century and prevailed for at least 50 years. The balance of research in the pharmaceutical industry up to the 1970s

placed chemistry clearly as the key discipline in drug discovery, the task of biologists being mainly to devise and perform assays capable of revealing possible useful therapeutic activity among the many anonymous white powders that arrived for testing. Research management in the industry was largely in the hands of chemists. This strategy produced many successes, including benzodiazepine tranquilizers, several antiepileptic drugs, antihypertensive drugs, antidepressants, and antipsychotic drugs. The surviving practice, of classifying many drugs on the basis of their chemical structure rather than on the more logical basis of their site or mode of action (therapeutic class), stems from this era.

We have mentioned the early days of pharmacology, with its focus on plant-derived materials, such as atropine, tubocurarine, strychnine, digitalis, and ergot alkaloids, which were almost the only drugs that existed until well into the twentieth century. Despite the rise of synthetic chemistry, natural products not only remain a significant source of new drugs, particularly in the field of chemotherapy, but also in other applications. Following the discovery of penicillin by Fleming in 1929, and its development as an antibiotic for clinical use by Chain and Florey in 1938, an intense search was undertaken for antibacterial compounds produced by fungi and other microorganisms, which yielded many useful antibiotics, including chloramphenicol (1947), tetracyclines (1948), streptomycin (1949), and others. The same fungal source that yielded streptomycin also produced actinomycin D used in cancer chemotherapy. Higher plants have continued to yield useful drugs, including vincristine and vinblastine (1958), paclitaxel (or taxol, 1971), and ixabepilone (2007). Demain and Vaishnav (2011) provide an excellent review of this from the perspective of cancer chemotherapy.

Outside the field of chemotherapy, successful drugs derived from natural products include ciclosporin (1972) and tacrolimus (1993), both of which come from fungi and are used to prevent transplant rejection. Soon after came mevastatin (1976), another fungal metabolite, which was the first of the “statin” series of cholesterol-lowering drugs which act by inhibiting the enzyme HMG-CoA reductase.

Overall, the pharmaceutical industry continues to have something of an on-again, off-again relationship with natural products. They often have weird and wonderful structures that cause hardened chemists to turn pale; they are often near-impossible to synthesize, troublesome to produce from natural sources, and “optimizing” such molecules to make them suitable for therapeutic use is prone to frequent failure. But nature continues to unexpectedly provide some of our most useful drugs, and most of its potential remains untapped.

Although chemistry was the preeminent discipline in drug discovery until at least the 1970s, the seeds of the biological revolution were sown long before. Starting foremost in the field of chemotherapy, where Ehrlich defined the principles of drug specificity in terms of a specific interaction between the drug molecule and a target molecule—the “receptor site”—in the organism, although we now take it for granted that in almost all cases a highly specific chemical

target molecule, as well as the “pharmacophore” or an outline portion of the drug molecule, determines what effects a therapeutic will yield, before Ehrlich no one had envisaged drug action in this way. By linking chemistry and biology, Ehrlich defined the parameters of modern drug discovery.

Despite these discoveries in Ehrlich’s field, chemotherapy remained empirical rather than target directed. That said, for many years, Ehrlich’s preoccupation with curing syphilis and the binding of chemical dyes, as exemplified by biological target-based drug development from the 1950s onwards, steadily shifted the industry’s focus from chemistry to biology (Hill and Rang, 2012). The history of successes in the field of chemotherapy prior to the antibiotic era (Table 1.3) demonstrates the diversity of sources of new therapeutic entities. The popular image of “magic bullets”—(a phrase first used by Ehrlich in 1905)—is the essence of today’s target-directed approaches to drug discovery.

More recently, as this book will show, all new categories of therapeutic entities (biotechnology-derived monoclonal antibodies, cell tissue therapies, and gene therapies) have entered use in medicine as “drugs.”

1.4 THE DRUG DEVELOPMENT PROCESS

While the processes for the discovery of new potential therapeutic drugs are very diverse (Gad, 2005; Choerghade, 2006; Mathieu, 2008), once the decision is made to move a candidate compound forward to (hopefully) market approval, the general process is well defined in the components of its regulatory requirements (though with significant variability and frequent change in its details). It has many components which are beyond the scope of safety assessment, and therefore of this volume (including chemical development, clinical evaluation, and a host of regulatory actions.)

The process generally proceeds by way of getting regulatory concurrences for entering clinical trials, then proceeding through three (not strictly defined) stages of clinical trials (Phase I, Phase II, and finally Phase 3), followed by submission of a full set of documents, data, and a proposed label seeking regulatory approval for a marketing application.

The metrics of this process as it now operates make cancer the most prevalent therapeutic target for new drugs, with perhaps as many as one-third of all new drug candidates being in this claim area. Heart diseases, CNS diseases, nervous system diseases, and immune system disorders follow in order of current popularity (Table 1.4).

According to www.pharmabiointredients.com, more than 16000 different drugs to be in development in 2006 were spread across the entire course of the development process (Table 1.5).

At the same time, the metrics of regulatory applications for the development of new drugs in the United States (where the best data is available) show a continued increase in the number of candidates entering the development process as indicated by the number of new (or original) INDs filed,

TABLE 1.3 Examples of Drugs from Different Sources

| Natural Products | Synthetic Chemistry ^a | Biopharmaceuticals Produced by Recombinant DNA Technology |
|---|----------------------------------|--|
| Antibiotics (penicillin, streptomycin, tetracyclines, cephalosporins, etc.) | Early successes include: | Human insulin (the first biotech product, registered 1982) |
| Anticancer drugs (doxorubicin, bleomycin, actinomycin, vincristine, vinblastine, taxol, etc.) | Antiepileptic drugs | Human growth hormone |
| Atropine, hyoscyne | Antimetabolites | α -interferon, γ -interferon |
| Ciclosporin | Barbiturates | Hepatitis B vaccine |
| Cocaine | Bronchodilators | Tissue plasminogen activator (t-PA) |
| Colchicine | Diuretics | Hirudin |
| Digitalis (digoxin) | Local anesthetics | Blood-clotting factors |
| Ephedrine | Sulfonamides | Erythropoietin |
| Heparin | | Granulocyte and granulocyte-monocyte colony-stimulating factor (G-CSF, GM-CSF) |
| Human growth hormone ^b | | |
| Insulin (porcine, bovine) ^b | | |
| Opium alkaloids (morphine, papaverine) | | |
| Physostigmine | | |
| Rauwolfia alkaloids (reserpine) | | |
| Statins | | |
| Streptokinase | | |
| Tubocurarine | | |
| Vaccines | | |

^aSince about 1950, synthetic chemistry has accounted for the great majority of new drugs.

^bNow largely or entirely replaced by material prepared by recombinant DNA technology.

TABLE 1.4 Potential New Drugs in US Clinical Trials by Primary Disease/Medical Use, 2005–2006

| Disease/Medical Use | # of Potential New Drugs in US Clinical Trials |
|--|--|
| Cancer | 5468 |
| Mental and behavioral disorders | 2397 |
| Heart disease | 2342 |
| Rare diseases | 5765 |
| Symptoms and general pathology | 4227 |
| Nervous system diseases | 2928 |
| Immune system disorders (not including HIV/AIDS) | 2578 |
| Urinary tract and sexual organs and pregnancy | 1756 |
| Skin and connective tissue diseases | 1727 |
| Blood and lymph conditions | 1654 |
| Bacterial and fungal diseases | 1591 |
| Respiratory tract diseases | 1548 |
| Digestive system diseases | 1527 |
| Nutritional and metabolic diseases | 1296 |
| Gland- and hormone-related diseases | 1216 |
| Viral diseases | 1168 |
| Diseases or abnormalities at or before birth | 1090 |
| Injuries, poisonings, and occupational diseases | 832 |
| Muscle, bone, and cartilage diseases | 699 |

TABLE 1.5 2006 Status of Drugs in Development

| Stage | Drugs |
|---|--------|
| New drug application (NDA)/biological license application (BLA) filed | 482 |
| Phase III | 1179 |
| Phase II | 2622 |
| Phase I/IND Filed | 2415 |
| Preclinical/discovery | 7569 |
| Recent product launches | 2002 |
| Total | 16 269 |

with the proportion of these that are commercial (or traditional INDs) continuing to increase (see Table 1.6).

Also, at the same time, the rate of approval of new molecular entities has only recently recovered to levels of 30 a year for the last 2 years. This preceding multiyear “drought” finally caused recognition that the traditional/existing system of development focused on blockbusters is irretrievably broken.

1.5 STRATEGIES FOR DEVELOPMENT: LARGE VERSUS SMALL COMPANY OR THE SHORT VERSUS LONG GAME

While harmonization and societal concern for safety are driving the changes in regulatory processes for device and drug development to become more confused, strategies for

TABLE 1.6 INDs Received and Active at CDER

| Calendar Year Received | Original INDs Received | Number of Active INDs at Years End | NDA's |
|------------------------|------------------------|------------------------------------|------------------------------------|
| 1998 | 2,419 | 12,723 | 121 |
| 1999 | 1,763 | 12,584 | 139 |
| 2000 | 1,812 | 11,838 | 115 |
| 2001 | 1,872 | 10,873 | 98 |
| 2002 | 2,374 | 11,544 | 105 |
| 2003 | 2,120 (426 commercial) | 12,661 (4,544 commercial) | 109 |
| 2004 | 1,837 (621 commercial) | 12,778 (4,827 commercial) | 115 |
| 2005 | 1,934 (637 commercial) | 13,360 (5,029 commercial) | 116 |
| 2006 | 1,863 (713 commercial) | 14,117 (5,445 commercial) | 123 |
| 2007 | 2,589 (779 commercial) | 14,566 (5,417 commercial) | 124 |
| 2008 | 2,039 (883 commercial) | 15,892 (5,962 commercial) | 128 |
| 2009 | 1,554 (730 commercial) | 9,299 (5,876 commercial) | 146 |
| 2010 | 1,330 (601 commercial) | 9,633 (5,838 commercial) | 103 |
| 2011 | 1,404 (644 commercial) | 9,883 (6,030 commercial) | 105 |
| 2012 | 1,284 (636 commercial) | 9,627 (5,966 commercial) | 33 (only recorded for 3 months) |
| 2013 | 1,429 (732 commercial) | 10,205 (6,115 commercial) | 133 |
| 2014 | 1,508 (782 commercial) | 10,802 (6,599 commercial) | 123 |
| 2015 | 1,564 (799 commercial) | 10,973 (6,894 commercial) | 146 |

product development and the associated nonclinical safety assessment can still be viewed in terms of broad trends.

The driving truths behind strategies in developing new drugs are:

1. Most molecules will fail. While the true success rate is certainly greater than the often quoted 1 in 10,000, it is clear that only 3–5% of those that enter initial clinical evaluation (i.e., for which an IND “opens”) become marketed drugs. This rate varies depending on therapeutic class (oncology drugs having a success rate as low as 1–2% and CNS therapeutics being only somewhat higher) (Pangalos et al., 2007).
2. The cost of developing drugs is high—while not the currently quoted “average” of \$1.4 billion, just getting to the point of an IND opening will cost a minimum of \$2 million. One can spread out the rate of expenditure over time or shorten the required time by spending money more rapidly. But there are fixed minimums for cost and time.

Costs of development go up sharply with time/progress—subsequent to a plain vanilla first-in-man (FIM) trial, outlays come to be spoken of first in tens of millions, and (frequently) before a marketing approval filing in the hundreds

of millions. Once the decision is made to develop a molecule into a drug, the process takes years. Again, one can dispute how many (from 5 to 16 years about covers the extreme range) and at no point up to the end is success (achieving marketing approval and economically successful therapeutic use) assured.

These truths conspire to produce the principal general goals behind drug development strategy:

1. Kill the losers as early as possible before too much money is spent on them.
2. Do all you can to minimize the time spent in developing a drug.

These principles produce a spectrum of strategies in the nonclinical safety assessment of drugs, best illustrated by looking at the two extreme cases.

1.5.1 Do Only What You Must

Driven by financial limitations and the plan that, at an optimal point in development (most commonly after either FIM/Phase I trials or a “proof of concept” Phase II trial), the candidate therapeutic will be licensed to or partnered with a large company, only the technical and regulatory steps necessary to

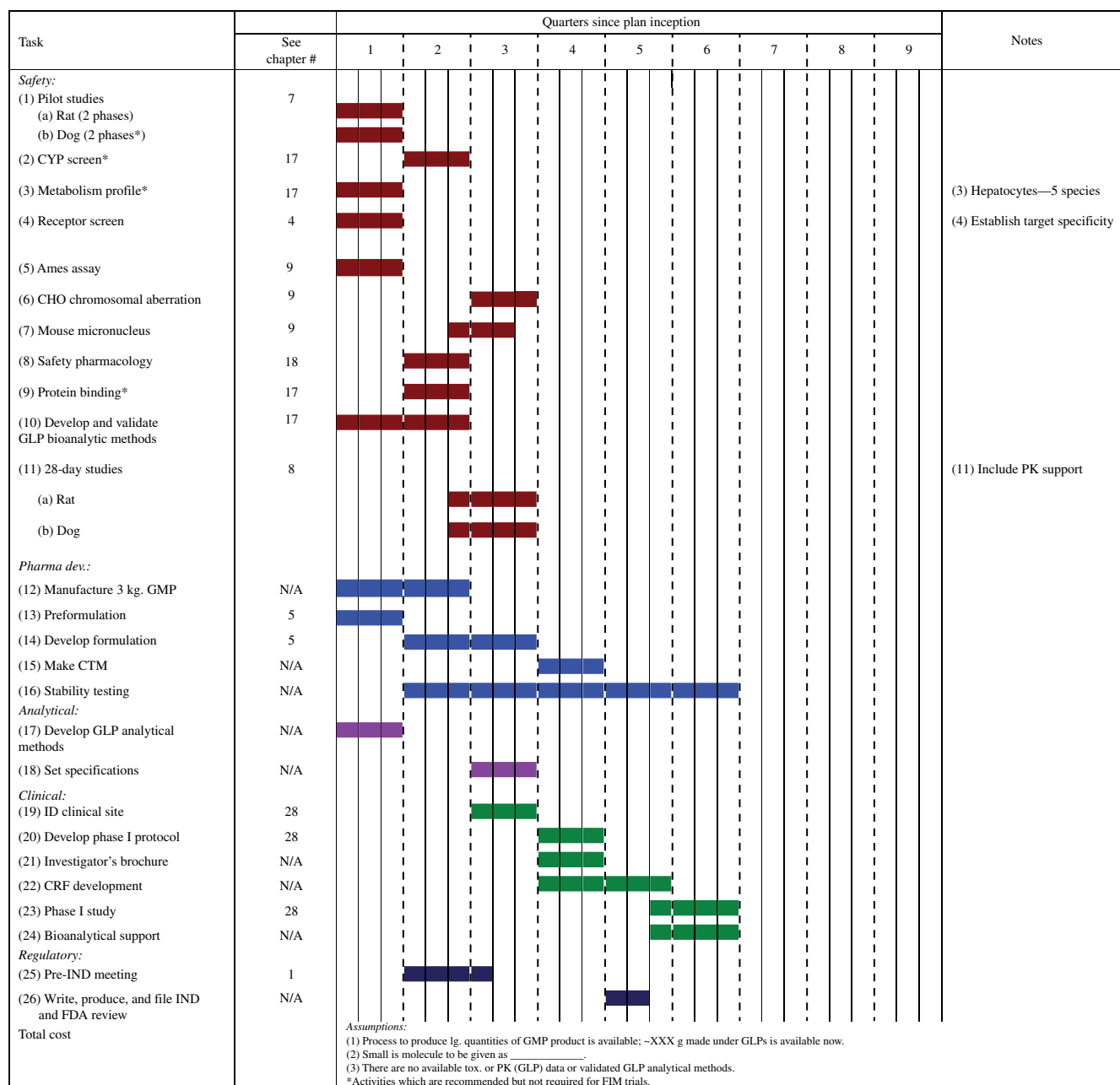


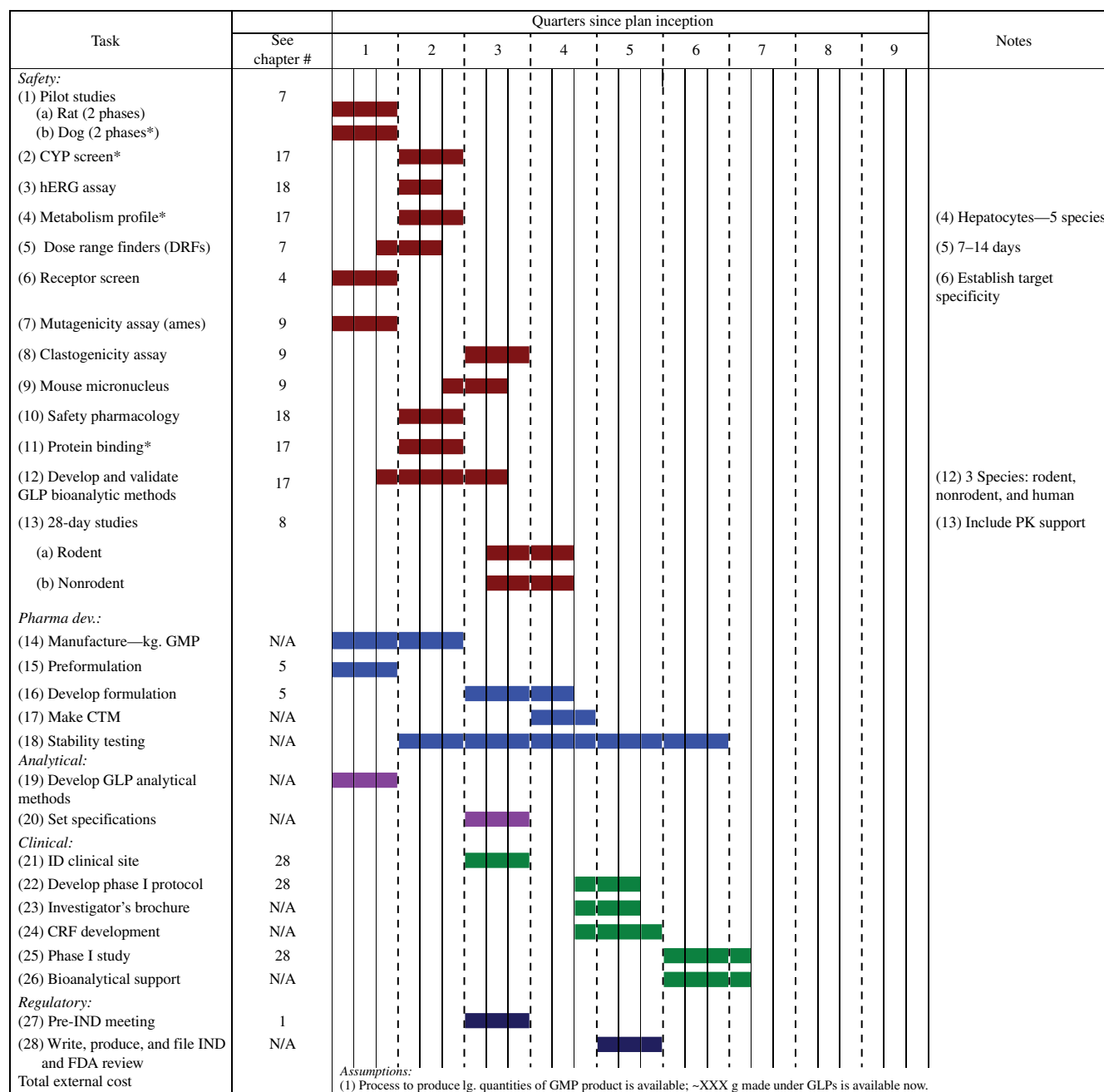
FIGURE 1.1 General case oral drug: lead through Phase I (do only what you must).

get a molecule to this point are to be performed. For those pursuing this case, the guidance provided by this book should prove essential (though not generally completely sufficient). This approach is summarized in Figure 1.1.

1.5.2 Minimize the Risk of Subsequent Failure

This is considered the traditional big company model. Studies and technical tasks are not limited to the minimum but rather are augmented by additional components. Development proceeds through a series of well-defined and carefully considered

“go-no-go” decision points. This approach is summarized in Figure 1.2. Many of the additional components are either limited, non-GLP forms of studies, which will be required later (such as Ames, acute toxicity, hERGs at only one concentration, and 7 days to 4 weeks repeat-dose studies), or studies which are inexpensive and could be done later (CYP inhibitors, induction, metabolic stability, and longer than required repeat-dose toxicity studies before proceeding into Phase II). Exactly which “extra” components are included vary from company to company and frequently reflect past experiences of the organization or individuals involved.



—: fill the number of Kg to be manufactured.

FIGURE 1.2 General case oral drug: lead through Phase I (minimize risk).

The studies performed to meet regulatory nonclinical safety assessment requirements (which must be considered to include all of the supportive toxicokinetic and metabolism activities and studies) can be thought of as belonging to three major categories:

- a. Those necessary to support the successful filing/opening of an IND, CTA or equivalent application, and of the subsequent FIM clinical studies.

- b. Those required to support continuation of clinical evaluation and development of a drug, up to and through successful Phase III studies.

- c. Those studies required to support a successful marketing approval application (NDA, BLA, or equivalent) but only required as such. This group is typically exemplified for carcinogenicity studies and the formal reproductive (as opposed to developmental) toxicity studies.

Which studies fit into what category is somewhat fluid and influenced by what patient population will be served (therapeutic claim) and the mechanism of action of the drug.

1.6 SAFETY ASSESSMENT AND THE EVOLUTION OF DRUG SAFETY

In the mid-nineteenth century, restrictions on the sale of poisonous substances were imposed in the United States and United Kingdom, but it was not until the early 1900s that any system of “prescription-only” medicines was introduced, requiring approval of purchase by a licensed medical practitioner. Soon afterwards, restrictions began to be imposed on what “cures” could be claimed in advertisements for pharmaceutical products and what information had to be given on the label; legislation evolved at a leisurely pace. Most of the concern was with controlling frankly poisonous or addictive substances or contaminants, not with the efficacy and possible harmful effects of new drugs.

In 1937, the use of diethylene glycol as a solvent for a sulfonamide preparation caused the deaths of 107 children in the United States, and a year later the 1906 Food and Drugs Act was revised, requiring safety to be demonstrated before new products could be marketed, as well as federal inspection of manufacturing facilities. The requirement for proven efficacy, as well as safety, was added in the Kefauver–Harris amendment in 1962 (said amendment being brought about largely by a safety issue—the thalidomide disaster in Europe).

In Europe, preoccupied with the political events in the first half of the century, matters of drug safety and efficacy were a minor concern, and it was not until the mid-1960s, in the wake of the thalidomide disaster—a disaster averted in the United States by an officer who used the provisions of the 1938 Food and Drugs Act to delay licensing approval—that the United Kingdom began to follow the United States’ lead in regulatory laws. Until then, the ability of drugs to do harm—short of being frankly poisonous or addictive—was not really appreciated, most of the concern having been about contaminants. In 1959, when thalidomide was first put on the market by the German company Chemie Grünenthal, it was up to the company to decide how much research was needed to satisfy itself that the drug was safe and effective. Grünenthal made a disastrously wrong judgment (see Sjöström and Nilsson (1972) for a full account), which resulted in an estimated 10 000 cases of severe congenital malformation following the company’s specific recommendation that the drug was suitable for use by pregnant women. This single event caused an urgent reappraisal on a global scale, leading to the introduction of much tighter government controls.

By the end of the 1960s, the primary planks in the regulatory platform—evidence of safety, efficacy, and chemical purity—were in place in most developed countries. Subsequently, the regulations have been adjusted in

various minor ways and adopted with local variations in most countries.

In 1988, Alder and Zbinden published *National and International Drug Safety Guidelines* which set forth the wide differences in safety assessment requirements between the different nations of the world, at the time global development of a drug required multiple safety assessment programs, with a great number of repetitions of studies and attendant extra costs and increased usage of test animals.

The solution to this was that ICH paradigm which, starting in the late 1980s, sought to have a harmonized set of global requirement for all aspects of drug development (not just assessment). The safety assessment aspects were embodied primarily in the S series ICH guidelines (M4 which sets forth the overall structure of nonclinical requirements being an exception). This did serve to largely standardize (“harmonize”) global requirements, with minor differences.

As the rest of this book will make clear, this system is now fraying a bit at the edges.

Recent additions of new guideline topic areas (e.g., immunotoxicology), revisions to existing guidelines (on genotoxicity and biotechnology), regional guideline responses to recent occurrences (the case in point being the failed TGN1412 FIM trial and the resulting two EMA special guidances issued in response to it), as well as differences in requirements for different therapeutic classes have reversed the harmonization trend.

Just as this book was being submitted for publication, reports have been released of a Phase I trial of BIA 10–2474, a fatty acid amide hydrolase (FAAH) inhibitor targeted at the body’s endocannabinoid system and intended to treat mood anxiety and movement coordination issues, going drastically wrong. Six males received repeat doses of the drug after 84 others had shown no marked effects. One was first pronounced brain dead but subsequently died, while three of the other five have also shown serious effects, perhaps irreversible.

The oral small molecule drug was made by the Portuguese company Bial, but clinical tests were performed in a commercial CRO in France (BioTrial). A meta-analysis of noncancer Phase I drug trials, published last year in *The British Medical Journal*, found serious adverse events in only 0.31% of participants and no deaths (Chan, 2016).

1.7 THE THREE STAGES OF DRUG SAFETY EVALUATION IN THE GENERAL CASE

Nonclinical safety assessment studies fall into three categories, as will be examined in detail in the remainder of this book. These are:

1. IND Enabling (“FIM”): the studies necessary to support the initiation of clinical trials in human beings. These are generally as specified in ICH M3, and this is the most common and numerous of all the three categories.

2. To support continued clinical development: as clinical development proceeds, longer repeat-dose drug studies must be performed, reproductive and developmental toxicology studies must be done, and other ancillary studies are required.
3. To support filing for marketing approval: the final studies generally required to support marketing of drugs—such as carcinogenicity.

Which studies fall into each of these categories, and exactly what studies must be done to support the development of a drug for a specific therapeutic claim, is extremely variable. The general case—much as specified in ICH M3(R2)—gives us a starting place for understanding what must be done.

At the same time, the image of the pharmaceutical industry in society is problematic (even more so in 2015 with well-publicized incidences of firms buying marketing rights to established small molecule drugs only to escalate prices 10–100-fold). The costs and economics of development are complex and not well understood, (Greider, 2003; Angell, 2004; Goozner, 2004; Petersen, 2008) while the role and abilities of regulatory agencies are equally misunderstood (Hawthorne, 2005).

But the general case really applies to the simplest oral drug intended for chronic use, and more often than not, doesn't apply. In fact, it may never fully apply.

REFERENCES

- Alder S, Zbinden G. (Eds.) (1988) *National and International Drug Safety Guidelines*. M.T.C. Verlag, Zollikon.
- Angell M. (2004) *The Truth about the Drug Companies: How They Deceive Us and What to Do About It*. Random House, New York.
- Chan S. (2016) 6 Hospitalized, one of them brain-dead, after drug trial in France. *The New York Times*, January 16, 2016: B2.
- Choerghade MS. (Ed.) (2006) *Drug Discovery and Development*. John Wiley & Sons, Inc., Hoboken, NJ.
- Demain AL, Vaishnav P. (2011) Natural products for cancer chemotherapy. *Microb. Biotechnol.* 4(6):687–699.
- Drugs.com (2014) *U.S. Pharmaceutical Sales—Q4 2013*. The Drugsite Trust, Auckland, New Zealand. Available at: <http://www.drugs.com/stats/top100/2013/sales> (accessed May 26, 2016).
- Gad SC. (Ed.) (2005) *Drug Discovery Handbook*. John Wiley & Sons, Inc., Hoboken, NJ.
- Gad SC. (2011) *Safety Evaluation of Pharmaceuticals and Medical Devices: International Regulatory Guidelines*, 1st Ed. Springer, New York.
- Goozner M. (2004) *The \$800 Million Pill: The Truth Behind the Cost of New Drugs*. University of California Press, Berkeley, CA.
- Greider K. (2003) *The Big Fix: How the Pharmaceutical Industry Rips Off American Consumers*. Public Affairs, New York.
- Hawthorne F. (2005) *Inside the FDA: The Business and Politics Behind the Drugs We Take and the Food We Eat*. John Wiley & Sons, Inc., Hoboken, NJ.
- Hill R, Rang HP. (Eds.) (2012) *Drug Discovery & Development: Technology in Transition*. Churchill Livingstone, New York.
- Mathieu M. (2008) *New Drug Development: A Regulatory Overview*, 8th Ed. Parexel International, Waltham, MA.
- Pangalos MN, Schecter LE, Hurko O. (2007) Drug development for CNS disorders: strategies for balancing risk and reducing attrition. *Nat. Rev. Drug Discov.* 6(7):521–532.
- Petersen M. (2008) *Our Daily Meds: How the Pharmaceutical Companies Transformed Themselves into Slick Marketing Machines and Hooked the Nation on Prescription Drugs*. Sarah Crichton Books, New York.
- PMLive (2015) *Top 25 Pharma Companies by Global Sales*. Accessible at: http://www.pmlive.com/top_pharma_list/global_revenues. Last accessed 8-9-2016.
- Sjöström H, Nilsson R. (1972) *Thalidomide and the Power of the Drug Companies*. Penguin, New York.
- Thayer AM. (2014) 30 years of generics: the door that legislation unlocked for generic drugs three decades ago has blown wide open. *Chem. Eng. News*. 92(39):8–16.

REGULATION OF HUMAN PHARMACEUTICAL SAFETY: ROUTES TO HUMAN USE AND MARKET

2.1 INTRODUCTION

The safety of pharmaceutical agents, medical devices, and food additives is the toxicology issue of the most obvious and longest-standing concern to the public. A common factor among the three is that any risk associated with a lack of safety of these agents is likely to affect a very broad part of the population, with those at risk having little or no option as to undertaking this risk. Modern drugs are essential for life in our modern society, yet there is a consistent high level of concern about their safety.

This chapter examines the regulations which establish how the safety of human pharmaceutical products is evaluated and established in the United States and the other major international markets. As a starting place, the history of this regulation will be reviewed, and the current organizational structure of the Food and Drug Administration (FDA) will be briefly reviewed, along with the other quasigovernmental bodies that also influence the regulatory processes. The current structure and context of the regulations in the United States and overseas will also be presented. From this point the general case of regulatory product development and approval will be presented. Nonclinical safety assessment study designs will be presented. The broad special case of biotechnology-derived therapeutic products and environmental concerns associated with the production of pharmaceuticals will be briefly addressed. The significant changes in regulation brought about by harmonization are also reflected.

As an aid to the reader, appendices are provided at the end of this book: a codex of acronyms that are used in this field, followed by a glossary which defines some key terms.

2.2 BRIEF HISTORY OF US PHARMACEUTICAL LAW

A synopsis of the history of US drug legislation is presented in Table 2.1. Here we will review the history of the three major legislative acts covering pharmaceuticals.

2.2.1 1906: Pure Food and Drug Act

As so eloquently discussed by Temin (1980), the history of health product legislation in the United States largely involves the passage of bills in Congress which were primarily in response to public demand. In 1902, for example, Congress passed the Biologics Act in response to a tragedy in St. Louis where 10 children had died after being given contaminated diphtheria toxins. Interestingly, the background that led to the passage of the first Pure Food and Drug Act in 1906 had more to do with food processing than drugs. The conversion from an agrarian to an urban society fostered the growth of a food-processing industry that was rife with poor practice. Tainted and adulterated food was commonly sold. Practices were sensationalized by the muckraking press, including books such as *The Jungle* by Upton Sinclair.

In the early debates in the US Congress on the Pure Food and Drug Act (passed in 1906), there was little mention of toxicity testing. When Harvey Wiley, chief of the Bureau of Chemistry, Department of Agriculture and driving force in the enactment of this early law, did his pioneering work (beginning in 1904) on the effects of various food preservatives on health, he did so using only human subjects and with no prior experiments in animals (Anderson, 1958). Ironically, work that led to the establishment of the FDA

TABLE 2.1 Important Dates in US Federal Drug Law

| Year | Event |
|---------------------|--|
| 1902 | Passage of the Virus Act, regulating therapeutic serums and antitoxins. Enforcement by the Hygienic Laboratory (later to become the National Institutes of Health (NIH)), Treasury Department |
| 1906 | Passage of Pure Food Act, including provisions for the regulations of drugs to prevent the sale of misbranded and adulterated products. Enforcement by the Chemistry Laboratory, Agriculture |
| 1912 | Passage of the Sherley Amendment. Specifically outlawed any false label claims as to curative effect |
| 1927 | Bureau of Chemistry renamed the Food, Drug, and Insecticide Administration |
| 1931 | Renamed again to Food and Drug Administration |
| 1938 | Passage of the Food, Drug, and Cosmetic Act. Superseded the law of 1906. Required evidence of safety, for example, studies in animals. Included coverage of cosmetics and medical devices. Specifically excluded biologics |
| 1944 | Administrative Procedures Act, codifying public health laws: included provision that for a biological license to be granted, a product must meet standards for safety, purity, and potency. NIH also given the responsibility for developing biologics not developed by the private sector |
| 1945 | Amendment to the 1936 Act requiring that the FDA examine and certify for release each batch of penicillin. Subsequently amended to include other antibiotics |
| 1949 | Publication of the first set of criteria for animal safety studies. Following several revisions, guidelines published in 1959 as Appraisals Handbook |
| 1951 | Passage of Durham–Humphrey Amendment. Provided the means for manufacturers to classify drugs as over the counter (not requiring prescription) |
| 1953 | Transfer of FDA to the Department of Health, Education, and Welfare (HEW) from Agriculture (now the Department of Health and Human Services) |
| 1962 | Passage of major amendments (the Kefauver bill) to the 1938 FDCA, which required proof of safety and effectiveness (efficacy) before granting approval of New Drug Applications. Required affirmative FDA approval |
| 1968 | FDA placed under the Public Health Service of HEW |
| 1970 | Controlled Substance Act and Controlled Substances Import and Export Act. Removed regulation of drug abuse from FDA (transferred to the Drug Enforcement Agency) and provided for stringent regulation of pharmaceuticals with abuse potential |
| 1972 | Transfer of authority to regulate biologics transferred from NIH to FDA. The NIH retained the responsibility of developing biologics |
| 1973 | Consumer Product Safety Act, leading to the formation of separate Consumer Product Safety Commission, which assumes responsibilities once handled by the FDA's Bureau of Product Safety |
| 1976 | Medical Device Amendment to the FDCA requiring for devices that not only effectiveness be proven but also safety |
| 1979 | Passage of the Good Laboratory Practices Act |
| 1983 | Passage of the first Orphan Drug Amendment to encourage development of drugs for small markets |
| 1984 | Drug Price Competition and Patent Term Restoration Act intended to allow companies to recover some of the useful patent life of a novel drug lost due to the time it takes the FDA to review and approve. Also permits the marketing of generic copies of approved drugs |
| 1985 | The “NDA rewrite” final rule. An administrative action streamlining and clarifying the New Drug Application process. Now embodied in 21 CFR 314 |
| 1986 | The US Drug Export Amendment Act of 1986. Permitted the export of drugs outside the United States prior to approval for the US market |
| 1987 | The “IND rewrite” final rule. “...to encourage innovation and drug development while continuing to assure the safety of (clinical) test subjects.” Federal Register 52:8798, 1987. Now embodied in 21 CFR 312 |
| 1992 | Prescription Drug User Fee Act. Established the payment of fees for the filing of applications (e.g., IND, NDA, PLA, etc.) |
| 1994 | Orphan Drug Amendment |
| 1997 | The Food and Drug Administration Modernization Act: to streamline the drug and device review and approval process |
| 2002, 2007 and 2012 | The Food and Drug Administration Modernization Act Amendments |

Note: Laws and amendments that have covered other aspects of FDA law, such as those governing food additives (e.g., FQPA), are not included in this table.

would probably not have been permitted under the current guidelines of the agency. Wiley's studies were not double blinded, so it is also doubtful that his conclusions would have been accepted by the present agency or the modern scientific community. Legislation in place in 1906 consisted strictly of a labeling law prohibiting the sale of processed food or drugs that were misbranded. No approval process was involved and enforcement relied on postmarketing criminal charges. Efficacy was not a consideration until 1911, when the Sherley Amendment outlawed fraudulent therapeutic claims.

2.2.2 1938: Food, Drug, and Cosmetic Act

The present regulations are largely shaped by the law passed in 1938. It will, therefore, be discussed in some detail. The story of the 1938 Food, Drug, and Cosmetic Act (FDCA) actually begins in 1933. Franklin D. Roosevelt had just won his first election and installed his first cabinet. Walter Campbell was the chief of the FDA, reporting to Rexford Tugwell, the Undersecretary of Agriculture. The country was in the depths of its greatest economic depression. This was before the therapeutic revolution wrought by antibiotics in the 1940s, and medicine and pharmacy as we know them in the 2010s were not practiced. Most medicines were, in fact, self-prescribed. Only a relatively small number of drugs were sold via physicians' prescription. The use of so-called patent (because the ingredients were kept secret) preparations was rife, as was fraudulent advertising. Today, for example, it is difficult to believe that in the early 1930s a preparation such as Radithor (nothing more than a solution of radium) was advertised for treatment of 160 diseases. It is in this environment that 1 day in the winter of 1933, Campbell delivered a memo to Tugwell on an action level of an insecticide (lead arsenite) used on fruits. Tugwell briskly asked why, if the chemical was so toxic, was it not banned outright. He was amazed to find out from Campbell that the agency had no power to do so.

The 1906 law was designed to control blatantly misbranded and/or adulterated foods and drugs and relied on post facto criminal charges for enforcement. Safety and efficacy were not an issue so long as the product was not misbranded with regard to content. Premarketing review of a drug was an unknown practice. Thus, attempts at rewriting the old 1906 law to include control of bogus therapeutic claims and dangerous preparations proved to be unsatisfactory. Paul Dunbar of the FDA suggested to Campbell that an entirely new law was needed. A committee of FDA professionals and outside academic consultants drafted a new bill, which immediately ran into trouble because no one in Congress was willing to sponsor it. After peddling the bill up and down the halls of Congress, Campbell and Tugwell convinced Senator Royal Copeland of New York to sponsor the bill. Unknowingly at the time, Copeland put himself in the eye of a hurricane that would last for 5 years.

The forces that swirled around Copeland and the Tugwell bill (Senate bill S.1944) were many. First was the immediate and fierce opposition from the patent medicine lobby. Flyers decried S.1944 as everything from a communist plot to un-American, stating it "would deny the sacred right of self-medication." In opposition to the patent trade organizations were two separate but unlikely allies: a variety of consumer advocacy and women's groups (such as the American Association of University Women, whose unfaltering support for the bill eventually proved critical to passage) and the mainline professional organizations. Interestingly, many of these organizations at first opposed the bill because it was not stringent enough. There were also the mainline professional pharmacy and medical organizations (such as the American Medical Association (AMA) and the American Association of Colleges of Pharmacy) whose support for the bill ranged from neutral to tepid, but did grow over the years from 1933 to 1938.

Secondly, there was the basic mistrust on the part of Congress toward Tugwell and other "New Dealers." At the same time, Roosevelt gave the measure only lukewarm support at best (legend has it that if it had not been for the First Lady, Eleanor, he would have given it no support at all) because of his political differences with Royal Copeland.

Thirdly, there was a considerable bureaucratic turf war over the control of pharmaceutical advertising. Finally, despite all efforts of the various lobbying groups, there was no popular interest or support for the bill. By the end of the congressional period, S.1944 had died for lack of passage.

The next 5 years would see the introduction of new bills, amendments, and competing measures, as well as committee meetings and hearings, lobbying, and House/Senate conferences. The details of this parliamentary infighting make for fascinating history but are outside the scope of this book. The reader is referred to an excellent history of this period, *Food and Drug Legislation in the New Deal* (Jackson, 1970).

The FDA was surprised by the force and depth of the opposition to the bill. The proposed law contained a then-novel idea that a drug was misbranded if its labeling made any therapeutic claim which was contrary to general medical practice and opinion. The definition of a drug was broadened to include devices used for medical purposes.¹ *Adulteration* was defined as any drug product dangerous to health when used according to label directions. The patent manufacturers charged that the new bill granted too much discretionary power to a federal agency and that no manufacturer could stay in business except by the grace of the Department of Agriculture, a charge that may have been correct. In response to the patent trade lobbying effort, the FDA launched its own educational drive consisting of radio spots, displays (such as

¹The use of a broad definition of what constitutes a drug for regulatory purposes is a precedent that remains in place today. For example, the computer software used in diagnostic systems is considered to be a pharmaceutical for purposes of regulation.

the sensationalized Chamber of Horrors exhibition, in which the toxicity of a variety of useless medicines was clearly displayed), mimeographed circulars, speaking engagements, posters, etc.

Ruth Lamb, FDA information officer at the time, was perhaps one of the hardest working and most quotable of the FDA staffers working the street at the time. For example, in reference to one of the counter bills that had language similar to the original Copeland bill, but with extremely complicated enforcement provisions, Ruth Lamb called it “an opus for the relief of indigent and unemployed lawyers.” She once described the Bailey amendment, which would have made proprietary drugs virtually immune to multiple seizures, as permitting the “sale of colored tap water as a cure for cancer...unless arsenic was added to each dose making [it] immediately dangerous.” After 1934, however, the educational efforts of the FDA were greatly attenuated by federal laws prohibiting lobbying by federal agencies.

With the autumn of 1937 came the beginnings of the oft-told elixir of sulfanilamide incident, which remains one of the nation’s worst drug tragedies. The Massengill Company was not one of the industry giants, but neither was it a “snake oil peddler.” The company’s chief chemist, Harold Watkins, was simply trying to develop a product and, in fact, did so in a manner consistent with the norms of the time. There was a perceived need for a liquid form of sulfanilamide, but it was difficult to dissolve. Then, Watkins hit upon diethylene glycol (at 72%) for use as a solvent. No toxicity tests were performed on the finished product, although the product did pass through the “control lab” where it was checked for appearance, fragrance, and consistency.

The first reports of human toxicity occurred in October 1937 when Dr. James Stevenson of Tulsa requested some information from the AMA because of six deaths in his area that were attributable to the elixir. At the time, no product of Massengill stood accepted by the Council on Pharmacy and Chemistry, and the Council recognized no solution of sulfanilamide. The AMA telegraphed Massengill, requesting samples of the preparation for testing. Massengill complied. The test revealed the diethylene glycol to be the toxic agent, and the AMA issued a general warning to the public on October 18, 1937. In the meantime, the FDA had become aware of the deaths and launched an investigation through its Kansas City station. By October 20, when at least 14 people had died, Massengill wired the AMA to request an antidote for their own product. By the end of October, at least 73 people had died, and another 20 suspicious deaths were linked to the drug. Had it not been for the response of the FDA, more deaths may have occurred. The agency put its full force of field investigators (239 members) on the problem and eventually recovered and accounted for 99.2% of the elixir produced. Massengill fully cooperated with the investigation and in November published a public letter expressing regret over the matter, but further stating that no

law had been broken. In fact, the company was eventually convicted on a long list of misbranding charges and fined a total of \$26,000 (the largest fine ever levied under the 1906 law).

The Massengill incident made the limits of the 1906 law quite clear. Because there were no provisions against dangerous drugs, the FDA could move only on the technicality of misbranding. The term *elixir* was defined by the US Pharmacopeia (USP) as “a preparation containing alcohol,” which elixir of sulfanilamide was not. It was only this technicality that permitted the FDA to declare the “elixir” misbranded, to seize the inventory, and to stop the sale of this preparation. If it had been called *solution of sulfanilamide*, no charges could have been brought.

The extensive press coverage of the disaster became part of the national dialogue. Letters poured into congressmen demanding action to prevent another such tragedy. Medical and pharmacy groups and journals insisted that a new law was required. Congress was in special session in November 1937 and did not need to be told about the tragedy. Copeland and Representative Chapman (of Kentucky) pressed resolutions calling for a report from the FDA on the tragedy. When issued, the FDA report stunned Congress, not only because of the human disaster but also because it made apparent that even had the bill then before Congress been law, the entire tragedy would still have occurred because there were no provisions for toxicity testing before new drugs entered the market. By December 1937 a new bill, S.3037, was introduced which stated that manufacturers seeking to place new drugs on the market would be required to supply records of testing, lists of components, descriptions of each manufacturing process, and sample labels. Drugs would require certification by the FDA before sale was permitted. A similar bill was introduced in the House by Chapman, although the issues of which agency was to control advertising of drugs still festered in the House. In January 1938, debate started on the Wheeler-Lea bill, which would ensure that all controls over drug advertising would remain with the Federal Trade Commission (FTC). Despite strong opposition by the FDA, the Wheeler-Lea bill was signed into law March 1938. While the loss of advertising control was a blow to the FDA, the Wheeler-Lea bill did facilitate the passage of the new food and drug law.

With the issue of advertising controls settled, the Copeland–Chapman bill faced one last hurdle. Section 701, which had been added in committee, provided for appeal suits that could be entered in any federal district court to enjoin the agency from enforcing new regulations promulgated as a result of the Act. Interestingly, this issue had more to do with foods than drugs, as its major focus was with acceptable tolerance limits for insecticides in food. The new bill defined an *adulterated food* as one containing any poison. However, because efforts to remove insecticides from fresh fruits and vegetables had never been completely

successful, the Secretary of Agriculture needed this power to set tolerance levels. Allies of food producers tried to introduce provisions in the new bill that provided methods for stalling a tolerance regulation with rounds of appeals. The bill passed the House despite such provisions (Section 701) and despite the resistance of consumer groups and the FDA, and went into joint committee. Roosevelt, in one of his rare efforts to support the FDA, made it clear that he would not accept the bill with such a cumbersome appeals process. The resulting compromise was an appeals process which limited the new evidence that could be introduced into one of the 10 circuit courts. Other provisions regarding labeling were also rectified in joint committee. In May 1938, S.3073 passed by unanimous vote. Both chambers ratified the joint committee report, and Franklin Delano Roosevelt signed the new law in June of 1938.

A historical note to this story was that Royal Copeland did not live to see his measure passed. In May 1938, he collapsed on the Senate floor. His death occurred 1 month before President Roosevelt signed his bill into law.

2.2.3 1962: Major Amendment

The 1938 law very much changed the manner in which Americans purchased pharmaceutical agents. In effect, it changed the pharmaceutical industry from a traditional consumer product industry to one in which purchases were made as directed by a third party (the physician). In 1929, ethical pharmaceuticals (prescription drugs) comprised only 32% of all medicines. By 1969 this was up to 83% (Temin, 1980). This led to a peculiar lack of competition in the ethical market. In 1959, Senator Estes Kefauver initiated his now-famous hearings on the drug industry. Interestingly, almost 30 years later, Senator Edward Kennedy had hearings on exactly the same matter. In 1961, Kefauver submitted a proposed legislation to amend the 1938 Act in such a way as to increase FDA oversight of the drug industry. The proposed amendment contained two novel propositions. The first was compulsory licensing, which would have required, for example, company "A" to license (with a royalty of no greater than 8% of sales) and company "B" to market a drug patented by company "A." Company "A" would have only 3 years' exclusivity with its patent. The second novel provision was that new drugs had to be not only "safe" but also "efficacious." There was not a ground swell of support for this legislation. When it was reported out of committee, it had been rewritten (including the removal of the licensing requirement) to the point that even Kefauver refused to support it. The Kennedy administration wanted new legislation but did not specifically support the Kefauver bill; rather it introduced its own legislation, sponsored by Representative Orrin Harris of Arkansas, and also with little support.

As in 1938, a tragic incident would again intercede in the legislative process: 1961 would see the development of the

thalidomide tragedy. An antianxiety agent marketed in Europe, thalidomide, was prescribed for pregnancy-related depression and nausea (or "morning sickness") and taken by countless women. At about the same time, phocomelia, a birth defect marked by the imperfect development of arms and legs, appeared in Europe. Thalidomide was eventually determined to be the causative teratogen in 1961 and was subsequently taken off the European market. The William S. Merrell Company had applied for a New Drug Application (NDA) for thalidomide in the United States in 1960. It was never approved because the FDA examiner, Dr. Frances Kelsey, had returned the application for lack of sufficient information. Eventually, the company withdrew the application. Senator Kefauver's staff had uncovered the thalidomide story as it was unfolding and had turned its findings over to the *Washington Post*. The *Post* reported the episode under the headline "Heroine of the FDA Keeps Bad Drug off the Market" in July 1962, 3 days after the Kefauver bill was reported out of committee. Needless to say, the news created public support for the bill, which was sent back to committee and reported out again with new language in August 1962. The Kefauver-Harris bill was signed into law in October 1962. It was demonstrated after the fact that thalidomide was teratogenic in the rabbit; out of the episode grew the current practice of testing new human pharmaceuticals for teratogenicity in two species, one generally being the rabbit.

The 1962 Drug Amendment made three major changes in the manner in which new drugs could be approved (Merrill, 1994). First, and perhaps the most important, was that it introduced the concept of effectiveness into the approval process. An NDA had to contain evidence that the drug was not only safe but also effective. The 1938 law contained no such specification. The effectiveness requirement necessitated that a drug company had to do more extensive clinical trials. The new law required that a company apply to the FDA for approval of its clinical testing plan under an Investigational New Drug Application (INDA). No response from the FDA was deemed to be acceptance. As each level of clinical testing came to require FDA review and approval, the new law made the FDA an active partner in the development of all drugs.

The second major change enacted under the 1962 law was the change in the approval process from premarket notification to a premarket approval system. Under the terms of the 1938 law, an NDA would take effect automatically if the FDA did not respond. For example, the only reason thalidomide was not approved was because Dr. Kelsey returned the application to the sponsor with a request for more information. In contrast, the 1962 law required affirmative FDA action before a drug could be put on the market. Under the terms of the 1962 amendments, the FDA was also empowered to withdraw NDA approval and remove a drug from the market for a variety of reasons, including new evidence that the product was unsafe or that the sponsor had

misrepresented or underreported data. The basic nonclinical safety testing regimen which currently applies was developed and adapted in that time frame (Goldenthal, 1968).

The third major change enlarged the FDA's authority over clinical testing of new drugs. Thus, not only was evidence of effectiveness required, but Section 505(d) of the Act specified the types of studies required. "Substantial evidence consisting of adequate and well-controlled investigations, including clinical investigations by a qualified expert." In meeting the statutory requirement for setting standards of clinical evidence, the FDA has become highly influential in the design of drug testing regimens (Merrill, 1994). Interestingly, discussed in detail by Hutt (1987), the FDA was initially quite unprepared for this new level of responsibility. It was not until 1973 that audited regulations on the determination of safety and effectiveness were put into place (these were, in fact, approved by the Supreme Court). While there have been several procedural changes (e.g., the 1985 Investigational New Drug (IND) rewrite) and additions (e.g., the 1988 IND procedures for life-threatening disease treatment), there have actually been no major changes in the law through 1992 with Prescription Drug User Fee Act (PDUFA) and 1997 with Food and Drug Administration Modernization Act (FDAMA) (amended in 2002, 2007, and 2012).

We must interject with an interesting historical aside at this point. Despite its reputation, thalidomide made a bit of a comeback in the 1990s (Blakeslee, 1998). Among other properties, thalidomide has been shown to have good anti-inflammatory properties, due to the fact that it apparently decreases the synthesis and/or release of tissue necrosis factor.

2.2.4 1992, 1997, 2002, 2007, and 2012: PDUFA and FDAMA

The history of pharmaceutical regulations has been dominated by two oft-opposing schools of thought: the need to provide the citizenry with effective medicaments and the need to protect the consumer from unsafe and misbranded products. The reader is referred to Peter B. Hutt's in-depth reviews (Hutt, 1983a, b) on the subject. For example, the very first federal drug legislation in the United States was the Vaccine Act of 1813, which mandated the provision of the smallpox vaccine to the general public. In the modern era, legislative debate could be further defined as the constant swing back and forth on these two issues (Hutt, 1983a, b), that is, safety versus development costs. In 1963, for example, Senator Hubert Humphrey presided over hearings on the FDA's implementation of the Drug Amendment of 1962. The FDA came under substantial criticism for failure to take strong action to protect the public from dangerous drugs. Eleven years later (1974), Senator Edward Kennedy conducted hearings addressing exactly the same issue.

Commissioner Schmidt pressed the point that the FDA is under constant scrutiny regarding the approval of "dangerous" drugs, but no hearing had ever been conducted (up to that time) on the failure of the FDA to approve an important new therapy.

The next decade and a half saw a proliferation of work that analyzed the impact of regulation on competitiveness and the introduction of new therapies (see Hutt (1983b) for a complete review). This included Grabowski and Vernon's work (1983), which concluded that regulation had significant adverse effect on pharmaceutical innovation. This examination of the cost of regulation continued into the 1990s. In a meticulous and well-researched study, DiMasi et al. (1994) reported that throughout the 1980s, the number of INDAs was decreasing and the new drug application success rate was also dropping, while the length of time between discovery and approval was increasing. Clearly this is a situation that could not go on forever. The reported cost of developing a new drug has risen from \$54 million (US) in 1976 to \$2.558 billion (US, with \$1.395 billion out of pocket and \$1.163 billion in time cost) in 2014 (DiMasi et al., 1991; Tufts, 2014). Members of the pharmaceutical industry and the biotechnology industry were becoming increasingly alarmed by the negative synergy caused by increased costs and increased time to market. In 1991, Dranove published an editorial examining the increased costs and decreased product flow that resulted from the 1962 amendment. He made the observation that European requirements are less stringent than those of the United States, yet the Europeans did not seem to be afflicted by a greater number of dangerous drugs (see Table 1.2). Yet, if one looks at an analysis of worldwide withdrawals for safety from 1960 to 1999 (Fung et al., 2001), one sees that of 121 products identified 42.1% were withdrawn from European markets alone, then 5% from North America, 3.3% from Asia Pacific, and 49.6% from multiple markets. The top five safety reasons for withdrawal were hepatic (26.2%), hematologic (10.5%), cardiovascular (8.7%), dermatologic (6.3%), and carcinogenic (6.3%) issue.

In an age of decreasing regulatory recourses, the FDA (as well as the Congress) was under increasing pressure to review and release drugs more quickly. In response, the Congress passed the 1992 PDUFA. Under the terms of this Act, companies would pay a fee to the agency to defray costs associated with application review. They would supposedly provide the FDA with the resources available to decrease application review time. In return, companies were guaranteed a more rapid review time. By all accounts, PDUFA has been successful. In 1992 (the year PDUFA was passed), 26 NDAs were approved, requiring on average 29.9 months for data review, while in 1996, 53 new drug (or biological) products were approved, each requiring an average of 17.8 months of review time. PDUFA was successful in decreasing review times, but has not really streamlined the procedures.

The acquired immune deficiency syndrome (AIDS) activist community was particularly vocal and effective in demanding more rapid approvals and increased access to therapies. There was also demand for FDA reform on a number of other fronts (e.g., medical devices, pediatric claims, women and minority considerations, manufacturing changes, etc.). In 1993 the House Commerce Committee on Oversight and Investigations, chaired by John Dingel (D-MI), released a comprehensive investigation and evaluation of the FDA entitled *Less than the Sum of its Parts*. The report was highly critical of the FDA and made a number of recommendations (Pilot and Waldmann, 1998). The mid-1990s also saw the reinventing government initiatives (RIGO) chaired by Vice President AL Gore. Under RIGO, the FDA sought to identify and implement administrative reform. The RIGO report issued was entitled *Reinventing Regulation of Drugs and Medical Devices*. The 104th Congress started hearings on FDA reform again in the winter of 1995. Two bills were introduced that provided the essential outline of what would become FDAMA. Senator Nancy Kassebaum (R-KS), chair of the Senate Committee on Labor and Human Resources, introduced S-1477. The second was H.R.3201, introduced by Rep. Joe Barton (R-TX). Other bills were introduced by Senator Paul Wellstone (D-MN) and Rep. Ron Weyden (D-OR), which focused more on medical devices but still paved the way for bipartisan support of FDA reform (Pilot and Waldmann, 1998). Eventually, the 105th Congress passed the FDAMA, which was signed into law by President Clinton in November 1997. The various sections of FDAMA are listed in Table 2.2. By any measure it was a very broad and complex, if not overly deep, piece of legislation. In 1998, Marwick (1998) observed, "a measure of the extent of the task is that implementation of the Act will require 42 new regulations, ... 23 new guidance notices, and 45 reports and other tasks." The FDA has identified these various tasks, regulations, and guidances necessary for the implementation of FDAMA. (FDA's FDAMA Implementation Chart is available at <http://www.fda.gov/RegulatoryInformation/egislation/SignificantAmendmentstotheFDCA/FDAMA/FDAMAImplementationChart/default.htm>, and the reader is urged to explore this site.) There is an FDAMA icon on the FDA home page, and both the Center for Biologics Evaluation and Research (CBER) and the Center for Drug Evaluation and Research (CDER) have issued various guidance documents. Some of the more interesting sections of the Act that may be of interest to toxicologists include the following:

- Two successive renewals of PDUFA for another 5 years.
- Fast track for breakthrough products.
- Changes in the fashion biologicals are regulated (elimination of the Establishment and Product licenses, both replaced with a Biologics License Application (BLA)).

- Changes in the fashion antibiotics are developed and regulated.
- Incentives for the development of pediatric claims.
- Companies will be permitted to disseminate information about approved uses for their products.
- FDAMA requires that the FDA establish a clinical trials database for drugs used to treat serious and life-threatening diseases, other than AIDS and cancers (databases for these diseases had already been established).

The full impact of FDAMA in the pharmaceutical industry in general and on toxicology within this industry in particular remains to be established.

This is a debate that has continued to the present and has been highlighted by demands for anti-HIV chemotherapeutic agents.

While it is not possible to review the history of regulations worldwide, it is possible to point out some differences. We will call attention to specific differences where appropriate throughout the remainder of the text.

The strength of the US regulatory system was emphasized at the BIO-Europe 1993 Conference. David Holtzman stated: "the main subject of the conference was regulation, and the U.S. was perceived to have the superior regulatory agency. It may be more difficult to satisfy but it is more predictable and scientifically based" (Holtzman, 1993). This predictability has not stultified growth in the biotechnology industry in the United States and has, in fact, made the United States a more inciting target for investment than Europe. It is also a system that, while not perfect, has permitted very few unsafe products on the market.

2.3 FDAMA SUMMARY: CONSEQUENCES AND OTHER REGULATIONS

In summary, federal regulation of the safety of drugs has had three major objectives:

1. Requiring testing to establish safety and efficacy
2. Establishing guidelines as to which tests are required and how they are designed
3. Promulgating requirements of data recording and reporting

The first of these objectives was served by the 1906 Act, which required that agents be labeled appropriately. This was amended in 1938, in response to the tragedies associated with elixir of sulfanilamide and Lash Lure, to require that drugs and marketed formulations of drugs be shown to be safe when used as intended. In the aftermath of the thalidomide tragedy, the 1962 Kefauver–Harris Amendment significantly tightened requirements for preclinical testing

TABLE 2.2 Summary of the Contents of the 1997 Food and Drug Administration Modernization Act

| Title/Subtitle | Section |
|--------------------------------------|---|
| I. Improving regulatory drugs | |
| A. Fees relating to drugs | 101. Findings |
| | 102. Definitions |
| | 103. Authority to assess and use drug fees |
| | 104. Annual reports |
| | 105. Savings |
| | 106. Effective date |
| | 107. Termination of effectiveness |
| B. Other improvements | 111. Pediatric studies of drugs |
| | 112. Expanding study and approval of fast-track drugs |
| | 113. Information program on trials for serious disease |
| | 114. Healthcare economic information |
| | 115. Manufacturing changes for drugs |
| | 116. Streamlining clinical research for drugs |
| | 118. Data requirements for drugs and biologics |
| | 119. Content and review of applications |
| | 120. Scientific advisory panels |
| | 121. Positron emission tomography |
| | 122. Requirements for radiopharmaceuticals |
| | 123. Modernization of regulation |
| | 124. Pilot- and small-scale manufacture |
| | 125. Insulin and antibiotics |
| | 126. Elimination of certain labeling requirements |
| | 127. Application of federal law to pharmacy compounding |
| | 128. Reauthorization of clinical pharmacology program |
| | 129. Regulation of sunscreen products |
| | 130. Report of postmarketing approval studies |
| | 131. Notification of discontinuance of a lifesaving product |
| II. Improving regulation of devices | 201. Investigational device exemptions |
| | 202. Special review for certain devices |
| | 203. Expanding humanitarian use of devices |
| | 204. Device standards |
| | 205. Collaborative determinations of device data requirements |
| | 206. Premarket notification |
| | 207. Evaluation of automatic class III designation |
| | 208. Classification panels |
| | 209. Certainty of review time frames |
| | 210. Accreditation of person for review of premarket notification reports |
| | 211. Device tracking |
| | 212. Postmarket notification |
| | 213. Reports |
| | 214. Practice of medicine |
| | 215. Noninvasive blood glucose meter |
| | 216. Data relating to premarket approval: product development protocol |
| | 217. Number of required clinical investigations for approval |
| III. Improving regulation of food | 301. Flexibility for regarding claims |
| | 302. Petitions for claims |
| | 303. Health claims for food products |
| | 304. Nutrient content claims |
| | 305. Referral statements |
| | 306. Disclosure of radiation |
| | 307. Irradiation petition |
| | 308. Glass and ceramic ware |
| | 309. Food contact substance |

TABLE 2.2 (Continued)

| Title/Subtitle | Section |
|------------------------|--|
| IV. General provisions | 401. Dissemination of information new uses 402. Expanded access of investigational therapies and diagnostics 403. Approval of supplemental applications for approved products 404. Dispute resolution 405. Informal agency statements 406. FDA mission and annual report 407. Information system 408. Education and training 409. Centers for education and research on therapeutics 410. Mutual recognition of agreements and global harmonization 411. Environmental impact review 412. National uniformity for nonprescription drugs and cosmetics 413. FDA study of mercury in drugs and foods 414. Interagency collaboration 415. Contracts for expert review 416. Product classification 417. Registration of foreign establishments 418. Clarification of seizure authority 419. Interstate commerce 420. Safety report disclaimers 421. Labeling and advertising compliance with statutory requirements 422. Rule of construction |
| V. Effective date | 501. Effective date |

(the INDAs) and premarket approval (the NDA) of new drugs. Regulations pertaining to INDAs and NDAs have been modified (most recently in 1988) but essentially remain the backbone of regulations of the toxicity evaluation of new human pharmaceutical agents.

The Good Laboratories Practice (GLP) Act, which specifies standards for study planning, personnel training, data recording, and reporting, came out in 1978 in response to perceived shoddy practices of the operations of laboratories involved in the conduct of preclinical safety studies. It was revised in 1985 and is discussed elsewhere in this book.

The final major regulatory initiative on preclinical evaluation for drug safety arose out of the AIDS crisis. To that point, the process of drug review and approval had very generally been perceived as slowing down, the FDA pursuing a conservative approach to requiring proof of safety and efficacy before allowing new drugs to become generally available. In response to AIDS, in 1988 the Expedited Delivery of Drugs for Life-Threatening Diseases Act established a basis for less rigorous standards (and more rapid drug development) in some limited cases.

In the United Kingdom, the Committee on Safety of Medicines (reporting to the minister of Health) regulates drug safety and development under the Medicines Act of 1968 (which has replaced the Therapeutic Substances Act of 1925). Details on differences in drug safety regulations in the international marketplace can be found in Alder and Zbinden (1988), but key points are presented in this chapter.

2.4 OVERVIEW OF US REGULATIONS

2.4.1 Regulations: General Considerations

The US federal regulations governing the testing, manufacture, and sale of pharmaceutical agents and medical devices are covered in Chapter 1, Title 21 of the Code of Federal Regulations (21 CFR). These comprise nine 6" × 8" (double-sided) volumes which stack 8" high. This title also covers foods, veterinary products, and cosmetics. As these topics will be discussed elsewhere in this book, in this chapter we will briefly review those parts of 21 CFR that are applicable to human health products and medicinal devices.

Of most interest to a toxicologist working in the pharmaceutical arena would be Chapter 1, Subchapter A (Parts 1–78), which cover general provisions, organization, etc. The GLPs are codified in 21 CFR 58.

General regulations that apply to drugs are in Subchapter C (Parts 200–299). This covers topics such as labeling, advertising, commercial registration, manufacture, and distribution. Of most interest to a toxicologist would be a section on labeling (Part 201, Subparts A–G, which covers Sections 201.1 through 201.317 of the regulations) as much of the toxicological research on a human prescription drug goes toward supporting a label claim. For example, specific requirements on content and format of labeling for human prescription drugs are covered in Section 201.57. Directions for what should be included under the “Precautions” section of a label are listed in 201.57(f). This includes 201.57(f)(6),

which covers categorization of pregnancy risk, and the reliance upon animal reproduction studies in making these categorizations is made quite clear. For example, a drug is given a pregnancy category B if “animal reproduction studies have failed to demonstrate a risk to the fetus.” The point here is not to give the impression that the law is most concerned with pregnancy risk. Rather, we wish to emphasize that much basic toxicological information must be summarized on the drug label (or package insert). This section of the law is quite detailed as to what information is to be presented as well as the format of presentation. Toxicologists working in the pharmaceutical arena should be familiar with this section of the CFR.

2.4.2 Regulations: Human Pharmaceuticals

The regulations specifically applicable to human drugs are covered in Subchapter D, Parts 300–399. The definition of a new drug is covered in Part 310(g):

A new drug substance means any substance that when used in the manufacture, processing or packaging of a drug causes that drug to be a new drug but does not include intermediates used in the synthesis of such substances.

The regulation then goes on to discuss “newness with regard to new formulations, indications, or in combinations.” For toxicologists, the meat of the regulations can be found in Section 312 (INDA) and Section 314 (applications for approval to market a new drug or antibiotic drug or NDA). The major focus for a toxicologist working in the pharmaceutical industry is on preparing the correct toxicology “packages” to be included to “support” these two types of applications. (The exact nature of these packages will be covered in the following.)

In a nutshell, the law requires solid scientific evidence of safety and efficacy before a new drug will be permitted into clinical trials or (later) onto the market. The INDA (covered in 21CFR 310) is for permission to proceed with clinical trials on human subjects. Once clinical trials have been completed, the manufacturer or “sponsor” can then proceed to file an NDA (covered in 21 CFR 314) for permission to market the new drug.

As stated in 321.21, “A sponsor shall submit an IND if the sponsor intends to conduct a clinical investigation with a new drug... [and] shall not begin a clinical investigation until... an IND... is in effect.” Similar procedures are in place in other major countries. In the United Kingdom, for example, a Clinical Trials Certificate (CTC) must be filed or a clinical trial exemption (CTX) obtained before clinical trials may proceed. Clinical trials are divided into three phases, as described in 312.21. Phase I trials are initial introductions into healthy volunteers primarily for the purposes of establishing tolerance (side effects), bioavailability, and

metabolism. Phase II clinical trials are “controlled studies... to evaluate effectiveness of the drug for a particular indication or disease.” The secondary objective is to determine common short-term side effects; hence the subjects are closely monitored. Phase III studies are expanded clinical trials. It is during this phase that definitive, large-scale, double-blind studies are performed.

The toxicologist’s main responsibilities in the IND process are to design, conduct, and interpret appropriate toxicology studies (or “packages”) to support the initial IND and then design the appropriate studies necessary to support each additional phase of investigation. Exactly what may constitute appropriate studies are covered elsewhere in this chapter. The toxicologist’s second responsibility is to prepare the toxicology summaries for the (clinical) investigator’s brochure (described in 312.23(a)(8)(ii)). This is an integrated summary of the toxicological effects of the drug in animals and *in vitro*. The FDA has prepared numerous guidance documents covering the content and format of INDs. It is of interest that in the Guidance for Industry (CDER and CBER, 1995), an in-depth description of the expected contents of the pharmacology and toxicology sections was presented. The document contains the following self-explanatory passage:

Therefore, if final, fully quality-assured individual study reports are not available at the time of IND submission, an integrated summary report of toxicological findings based on the unaudited draft toxicologic reports of the completed animal studies may be submitted.

If audited draft but not yet finalized reports are used in an initial IND, the finalized report must be submitted within 120 days of the start of the clinical trial. The sponsor must also prepare a document identifying any differences between the preliminary and final reports and the impact (if any) on interpretation.

Thus, while the submission of fully audited reports is preferable, the agency does allow for the use of incomplete reports.

Once an IND or CTC/CTX is opened, the toxicologists may have several additional responsibilities: First, to design, conduct, and report the additional tests necessary to support a new clinical protocol or an amendment to the current clinical protocol (Section 312.20). Secondly, to bring to the sponsor’s attention any finding in an ongoing toxicology study in animals “suggesting a significant risk to human subjects, including any finding of mutagenicity, teratogenicity or carcinogenicity,” as described in 21 CFR 312.32. The sponsor has a legal obligation to report such findings within 10 working days. Third, to prepare a “list of the preclinical studies ... completed or in progress during the past year” and a summary of the major preclinical findings. The sponsor is required (under Section 312.23) to file an annual report

TABLE 2.3 Composition of Standard Investigational New Drug Application (Traditional Format)

| |
|---|
| 1. IND cover sheets (Form FDA-1571) |
| 2. Table of contents |
| 3. Introductory statement |
| 4. General (clinical) investigation plan |
| 5. (Clinical) investigators brochure |
| 6. (Proposed) clinical protocol(s) |
| 7. Chemistry, manufacturing, and control information (CMC) |
| 8. Pharmacology and toxicology information (includes metabolism and pharmacokinetic assessments done in animals) |
| 9. Previous human experience with the investigational drug |
| 10. Additional information |
| 11. Other relevant information |

(within 60 days of the IND anniversary date) describing the progress of the investigation. INDs are never “approved” in the strict sense of the word. Once filed, an IND can be opened 30 days after submission, unless the FDA informs the sponsor otherwise. Complete and thorough reports on all pivotal toxicological studies must be provided with the application. The structure of an IND is outlined in Table 2.3.

If the clinical trials conducted under an IND are successful in demonstrating safety and effectiveness (often established at a pre-NDA meeting, described in 21 CFR 312.47(b)(2)), the sponsor can then submit an NDA. Unlike an IND, the NDA must be specifically approved by the agency. The toxicologist’s responsibility in the NDA/Marketing Authorization Application (MAA) process is to prepare an integrated summary of all the toxicology and/or safety studies performed and be in a position to present and review the toxicology findings to the FDA or its advisory bodies. The approval process can be exhausting, including many meetings, hearings, appeals, etc. The ground rules for all of these are described in Part A of the law. For example, all NDAs are reviewed by an “independent” (persons not connected with either the sponsor or the agency) scientific advisory panel which reviews the findings and makes recommendations as to approval. MAAs must be reviewed by and reported on by an expert recognized by the cognizant regulatory authority. Final statutory approval in the United States lies with the Commissioner of the FDA. It is hoped that few additional studies will be requested during the NDA review and approval process. When an NDA is approved, the agency will send the sponsor an approval letter and will issue a Summary Basis of Approval (SBA)(312.30), which is designed and intended to provide a public record on the agency’s reasoning for approving the NDA while not revealing any proprietary information. The SBA can be obtained through Freedom of Information and can provide insights into the precedents for which types of toxicology studies are used to support specific types of claims.

2.4.3 Regulations: Environmental Impact

Environmental impact statements, while once important only for animal drugs, must now accompany all MDAs. This assessment must also be included in the Drug Master File (DMF). The procedures, formats, and requirements are described in 21 CFR 2531. This requirement has grown in response to the National Environmental Policy Act, the heart of which required that federal agencies evaluate every major action that could affect the quality of the environment. In the INDs, this statement can be a relatively short section claiming that relatively small amounts will pose little risk to the environment. The EEC has similar requirements for drug entities in Europe, though data requirements are more strenuous. With NDAs, this statement must be more substantial, detailing any manufacturing and/or distribution process that may result in release into the environment. Environmental fate (e.g., photohydrolysis) and toxicity (e.g., fish, daphnia, and algae) studies will be required. While not mammalian toxicology in the tradition of pharmaceutical testing, preparing an environmental impact statement will clearly require toxicological input. The FDA has published a technical bulletin covering the tests it may require (FDA, 1987).

2.4.4 Regulations: Antibiotics

The NDA law (safety and effectiveness) applies to all drugs, but antibiotic drugs were treated differently until the passage of FDAMA in 1997. Antibiotic drugs had been treated differently by the FDA since the development of penicillin revolutionized medicine during World War II. The laws applicable to antibiotic drugs were covered in 21 CFR 430 and 431. Antibiotics such as penicillin or doxorubicin are drugs derived (in whole or in part) from natural sources (such as molds or plants) which have cytotoxic or cytostatic properties. They were treated differently from other drugs as the applicable laws required a batch-to-batch certification process. Originally passed into law in 1945 specifically for penicillin, this certification process was expanded by the 1962 amendment (under Section 507 of the FDCA) to require certification of all antibiotic drugs, meaning that the FDA would assay each lot of antibiotic for purity, potency, and safety. The actual regulations were covered in 21 CFR Subchapter D, Parts 430–460 (over 600 pages), which describes the standards and methods used for certification for all approved antibiotics. Section 507 was repealed by FDAMA (Section 125). As a result of the repeal of Sections 507, the FDA is no longer required to publish antibiotic monographs. In addition, the testing, filing, and reviewing of antibiotic applications are now handled under Section 505 of the Act like any other new therapeutic agent. The FDA has published a guidance document to which the reader is referred for more details (CDER, 1998).

2.4.5 Regulations: Biologics

Biological products are covered in Subchapter F, Parts 600–680. As described in 21 CFR 600.3(h), “biological product means any virus, therapeutic serum, toxin, antitoxin or analogous product applicable to the prevention, treatment or cure of diseases or injuries of man.” In other words, these are vaccines and other protein products derived from animal sources. Clearly the toxicological concerns with such products are vastly different than those involved with low molecular weight synthetic molecules. There is little rational basis, for example, for conducting a 1-year repeated-dose toxicity study with a vaccine or a human blood product. The FDA definition for safety with regard to these products is found in 21 CFR 603.1(p): “Relative freedom from harmful effect to persons affected, directly or indirectly, by a product when prudently administered.” Such safety consideration has more to do with purity, sterility, and adherence to good manufacturing standards than with the toxicity of the therapeutic molecule itself. The testing required to show safety is stated in licensing procedures 21 CFR 601.25(d)(1): “Proof of safety shall consist of adequate test methods reasonably applicable to show the biological product is safe under the prescribed conditions.” Once a license is granted, each batch or lot of biological product must be tested for safety, and the methods of doing so are written into the law. A general test for safety (i.e., required in addition to other safety tests) is prescribed using guinea pigs as described in 610.11. Additional tests are often applied to specific products. For example, 21 CFR 630.35 describes the safety tests required for measles vaccines, which includes tests in mice and *in vitro* assays with tissue culture. Many new therapeutic entities produced by biotechnology are seeking approval as biologics with the results being FDA approval of a Product License Application (PLA). Table 2.4 presents general guidance for the basis of deciding if an individual entity falls under CDER or CBER authority for review.

The International Conference on Harmonization (ICH) has published its document S6 Preclinical Safety Evaluation of Biotechnology-Derived Pharmaceuticals. The FDA (both CDER and CBER jointly) has published the document as a Guidance for Industry (FDA, 1997).

A current list of regulatory documents (including the most recent points to consider (PTCs)) can be found on the FDA website by accessing the FDA home page at www.fda.gov, scrolling down and finding the “Regulatory Information” tab with applicable links near the bottom left. The Regulatory Information site can also be directly accessed using the following web address: <http://www.fda.gov/RegulatoryInformation/default.htm>

2.4.6 Regulations versus Law

A note of caution must be inserted here. The law (document passed by Congress) and the regulations (documents written by regulatory authorities to enforce laws) are separate

TABLE 2.4 Product Class Review Responsibilities

| |
|---|
| <i>Center for Drug Evaluation and Research</i> |
| Natural products purified from plant or mineral sources |
| Products produced from solid tissue sources (excluding procoagulants, venoms, blood products, etc.) |
| Antibiotics, regardless of method of manufacture |
| Certain substances produced by fermentation |
| Disaccharidase inhibitors |
| HMG-CoA inhibitors |
| Synthetic chemicals |
| Traditional chemical synthesis |
| Synthesized mononuclear or polynuclear products including antisense chemicals |
| Hormone products |
| <i>Center for Biologics Evaluation and Research</i> |
| Vaccines, regardless of manufacturing method |
| <i>In vivo</i> diagnostic allergenic products |
| Human blood products |
| Protein, peptide, and/or carbohydrate products produced by cell culture (other than antibiotics and hormones) |
| Immunoglobulin products |
| Products containing intact cells or microorganisms |
| Proteins secreted into fluids by transgenic animals |
| Animal venoms |
| Synthetic allergens |
| Blood banking and infusion adjuncts |

documents. Sections in the law do not necessarily have numerical correspondence with those of the regulations. For example, the regulations on the NDA process are described in 21 CFR 312, but the law describing the requirement for an NDA process is in Section 505 of the FDCA. Because regulations rather than laws themselves have a greater impact on toxicological practice, greater emphasis is placed on regulation in this chapter. For a complete review of FDA law, the reader is referred to the monograph by Food and Drug Law Institute in 1984.

Laws authorize the activities and responsibilities of the various federal agencies. All proposed laws before the US Congress are referred to committees for review and approval. The committees responsible for FDA oversight are summarized in Table 2.5. This table also highlights that authorizations and appropriations (the funding necessary to execute authorizations) are handled by different committees.

2.5 ORGANIZATIONS REGULATING DRUG AND DEVICE SAFETY IN THE UNITED STATES

The agency formally charged with overseeing the safety of drugs in the United States is the FDA. The FDA is headed by a commissioner who reports to the Secretary of the Department of Health and Human Services (DHHS) and has a tremendous range of responsibilities. Drugs are

TABLE 2.5 Congressional Committees Responsible for FDA Oversight

| | |
|----------------------|---|
| <i>Authorization</i> | |
| Senate | All public health service agencies are under the jurisdiction of the Labor and Human Resources Committee |
| House | Most public health agencies are under the jurisdiction of the Health and the Environmental Subcommittee of the House Energy and Commerce Committee |
| <i>Appropriation</i> | |
| Senate | Unlike most other public health agencies, the FDA is under the jurisdiction of the Agriculture, Rural Development, and Related Agencies Subcommittee of the Senate Appropriations Committee |
| House | Under the jurisdiction of the Agriculture, Rural Development, and Related Agencies Subcommittee of the House Appropriations Committee |

overseen primarily by the CDER (though some therapeutic or healthcare entities are considered biologics and are overseen by the corresponding CBER). Figure 2.1 presents the organization of CDER, and that of CBER is shown in Figure 2.2.

Most of the regulatory interactions of toxicologists take place with these two offices of Drug Evaluation, which have under them a set of groups focused on areas of therapeutic claim (cardiorenal, neuropharmacological, gastrointestinal and coagulation, oncology and pulmonary, metabolism and endocrine, anti-infective and antiviral). Within each of these are chemists, pharmacologists/toxicologists, statisticians, and clinicians. When an IND is submitted to the offices of Drug Evaluation, it is assigned to one of the therapeutic groups based on its area of therapeutic claim. Generally, it will remain with that group throughout its regulatory approval “life.” INDs, when allowed, grant investigators the ability to go forward into clinical (human) trials with their drug candidate in a predefined manner, advancing through various steps of evaluation in human (and in additional preclinical or animal studies) until an NDA can be supported, developed, and submitted. Likewise for biological products, the PLA or other applications (INDA, IND) are handled by the offices of Biological Products Review within the CBER.

For drugs, there is at least one nongovernmental body which must review and approve various aspects—the USP (established in 1820)—which maintains (and revises) the compendia of the same name, as well as the National Formulary which sets drug composition standards (Ember, 2001). This volume sets forth standards for purity of products in which residues may be present and tests for determining various characteristics of drugs, devices, and biologics. The USP also contains significant “guidance” for the evaluation process (USP, 2015).

2.6 PROCESS OF PHARMACEUTICAL PRODUCT DEVELOPMENT AND APPROVAL

Except for a very few special cases (treatments for life-threatening diseases such as cancer or AIDS), the safety assessment of new drugs is mandated by regulations which seemingly proceed in a rather fixed manner. The IND is filed to support (or enable) clinical testing and development of the drug. An initial set of studies (typically, studies of appropriate length by the route intended for humans are performed in both a rodent (typically rat) and a nonrodent (usually a dog or a primate)) are required to support phase I clinical testing. Such phase I testing is intended to evaluate the safety (“tolerance” in clinical subjects), pharmacokinetics, and general biological effects of a new drug and is conducted in normal volunteers (almost always males).

Successful completion of phase I testing allows, with the approval of the FDA, progression into phase II clinical testing. Here, selected patients are enrolled to evaluate therapeutic efficacy, dose ranging, and more details about the pharmacokinetics and metabolism. Longer-term systemic toxicity studies must be in conformity with the guidelines that are presented in the next section. Once a sufficient understanding of the actions, therapeutic dose–response, and potential risk-to-benefit ratio of a drug is in hand (once again, with FDA approval), trials move into phase III testing.

Phase III tests are large, long, and expensive. They are conducted using large samples of selected patients and are intended to produce proof of safety and efficacy of a drug. Two studies providing statistically significant proof of the claimed therapeutic benefit must be provided. All resulting data from preclinical and clinical animal studies are organized in a specified format in the form of an NDA, which is then submitted to the FDA.

By the time phase III testing is completed, some additional preclinical safety tests must also generally be in hand. These include the three separate reproductive and developmental toxicity studies (segments I and III in the rat and segment II in the rat and rabbit) and carcinogenicity studies in both rats and mice (unless the period of therapeutic usage is intended to be very short). Some assessment of genetic toxicity will also be expected.

The ultimate product of the pharmaceutical toxicologist will thus generally be the toxicology summaries of the IND and NDA (or PLA). For medical devices, the equivalents are the Investigational Device Exemption (IDE) and Product Development Notification (PDN). Data required to support each of these documents is specified in a series of guidelines, as will be discussed in the following.

Acceptance of these applications is contingent not only upon adherence to guidelines and good science but also adherence to GLPs.

**Food and Drug Administration
Office of Medical Products and Tobacco
center of Drug Evaluation and Research**

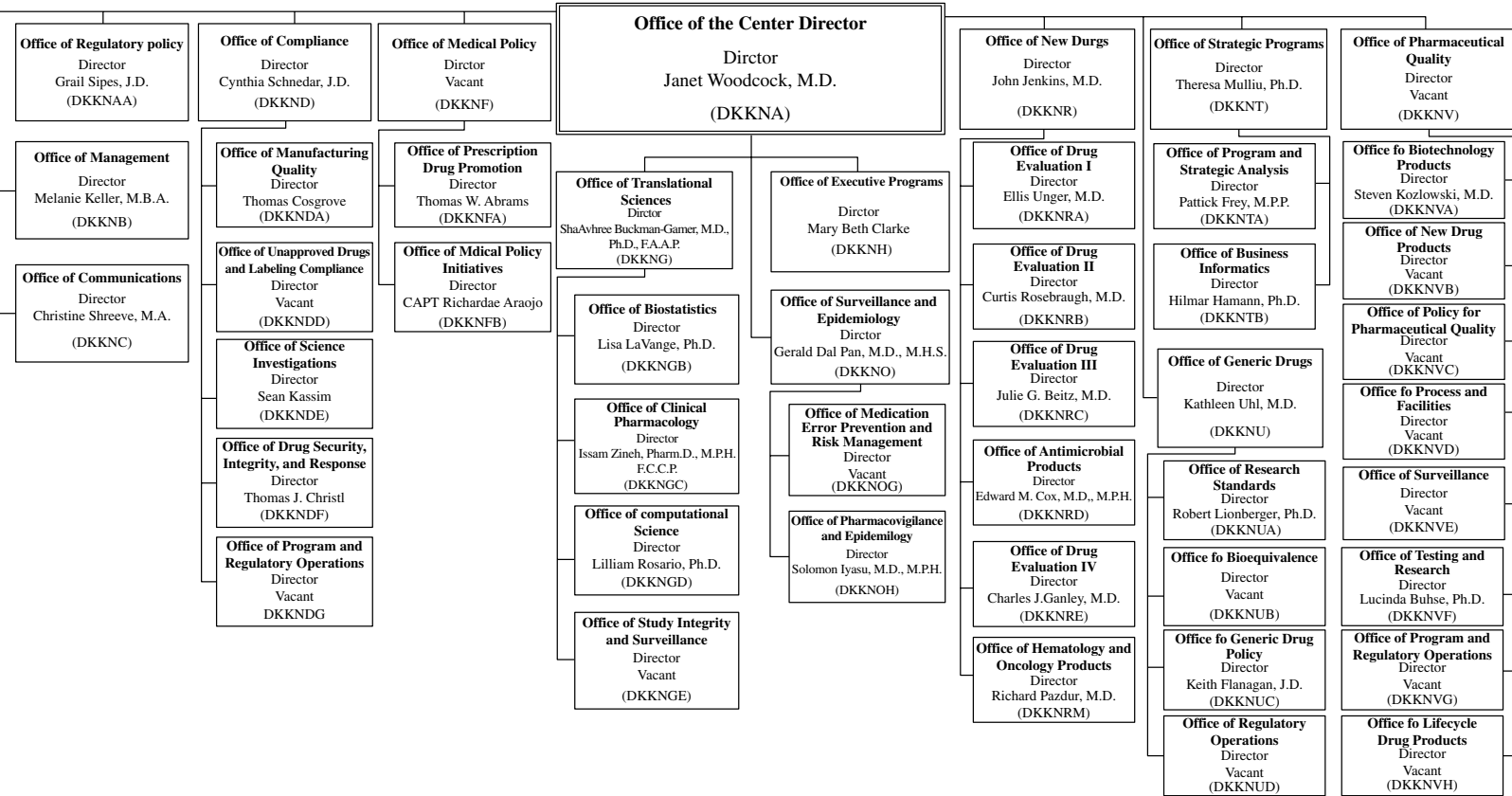


FIGURE 2.1 Center for Drug Evaluation and Research (CDER). *Source:* <http://www.fda.gov/downloads/AboutFDA/CentersOffices/OrganizationCharts/UCM439876.pdf>

**Food and Drug Administration
Office of Medical Products and Tobacco
Center for Biologics Evaluation and Research**

Director
Karen Midthun, M.D.
Deputy Director
Peter W. Marks, M.D., Ph.D.
Associate Director for Medicine
Barbara D. Buch, M.D.
Associate Director for Policy
Diane M. Maloney, J.D.
Associate Director for Quality Assurance
Sheryl L. Lard-Whiteford, Ph.D.
Associate Director for Research
Carolyn A. Wilson, Ph.D.
Associate Director for Review Management
Christopher Joneckis, Ph.D.
Senior Advisor for Counterterrorism/Medical
Countermeasures
Cynthia L. Kelley, M.S.
Senior Advisor for International Affairs
Joan Wilmarth Blair, M.A.
Senior Scientist for Emerging and Pandemic Threat
Preparedness
David S. Cho, Ph.D.
Executive Operations Staff
Anne E. Wilcox
(DKKBA)

Office of Management

Director
James M. Sigg
Deputy Director
Deirdre P. Hussey
Planning and Performance
Management Staff
Mary Pat Leary
(DKKBB)

**Office of Compliance and
Biologics Quality**

Director
Mary Anne Malarkey
Deputy Director
Mark Schwartz, J.D.
Associate Director for Policy
Anita F. Richardson
Associate Director for Regulatory
Policy
Diane Alexander
Associate Director for
Manufacturing Science
Edward Patten, M.S.
Associate Director for Safety and
Communication
Jean Makie
Associate Director for
Bioinformatics
Sujay Pandey, Ph.D.
(DKKBC)

**Office of Blood Research and
Review**

Director
Jay S. Epstein, M.D.
Deputy Director
Ginette Y. Michaud, M.D.
Associate Director for
Regulatory Affairs
Alan E. Williams, Ph.D.
Associate Director for
Research
Chintamani D. Atreya, Ph.D.
Associate Director for Medical
Affairs
Basil Golding, M.D. (Acting)
Associate Deputy Director for
Quality Assurance
Joseph Giglio, M.S.
Deputy Associate Director for
Bioinformatics
Vacant
Administrative Staff
Michelle Salvetti
Regulatory Project
Management Staff
Iliana Valencia, M.S.
(DKKBD)

**Office of Vaccines Research
and Review**

Director
Marion F. Gruber, Ph.D.
Deputy Director
Phillip R. Krause, M.D.
Associate Director for
Regulatory Policy
Theresa M. Finn, Ph.D.
Associate Director
for Research
Konstantin Chumakov, Ph.D.
Associate Director for
Scientific Affairs
Erik A. Henchal, Ph.D.
Associate Director for
Medical Policy and Vaccine
Safety
Karen Farizo, M.D. (Acting)
(DKKBF)

**Office of Biostatistics and
Epidemiology**

Director
Steven A. Anderson, Ph.D.,
M.P.P.
Deputy Director
David Martin, M.D.
Associate Director for Risk
Assessment
Mark O. Walderhaug, Ph.D.
Associate Director for
Research
Richard Forshee, Ph.D.
Associate Director for
Regulatory Policy
Jawahar Tiwari, Ph.D.
(DKKBJ)

**Office of Communication,
Outreach, and Development**

Director
Lorrie H. McNeill
Deputy Director
Joanne C. Binkley
(DKKBH)

**Office of Cellular, Tissue, and
Gene Therapies**

Director
Cecilia M. Witten, M.D., Ph.D.
Deputy Director
Stephanie L. Simek, Ph.D.
Associate Director for
Research
Suzanne L. Epstein, Ph.D.
Associate Director for Policy
Richard D. McFarland, M.D.
Associate Director for
Information Management
Theodore R. Stevens, M.S.
Associate Director for Policy—
New Legislation
Rachael F. Anatol, Ph.D.
Regulatory Management Staff
Patrick S. Riggins, Ph.D.
(DKKBL)

FIGURE 2.2 Center for Biologics Evaluation and Research (CBER). *Source:* <http://www.fda.gov/AboutFDA/CentersOffices/OrganizationCharts/ucm347874.htm>

2.7 TESTING GUIDELINES

2.7.1 Toxicity Testing: Traditional Pharmaceuticals

Although the 1938 Act required safety assessment studies, no consistent guidelines were available. Guidelines were first proposed in 1949 and published in the *Food, Drug, and Cosmetic Law Journal* that year (Burns, 1983). Following several revisions, these guidelines were issued as the Appraisal Handbook in 1959. While never formally called a guideline, it set the standard for preclinical toxicity test design for several years. The current basic guidelines for testing required for safety assessment in support of the phases of clinical development of drugs were first outlined by Goldenthal (1968) and later incorporated into a 1971 FDA publication entitled *FDA Introduction to Total Drug Quality*.

All general case pharmaceuticals need to address four major aspects of toxicology before going into humans. These are systemic toxicity, potential genetic toxicity, safety pharmacology, and (if any route of administration other than oral) local tissue tolerance issues.

2.7.2 General or Systematic Toxicity Assessment

Table 2.6 presents an overview of the current FDA toxicity testing guidelines for human drugs. Table 2.7 presents the parallel ICH guidance (ICH, 2009) which now largely supplants the FDA guidelines. They are misleading in their apparent simplicity, however. First, each of the systemic toxicity studies in these guidelines must be designed and executed in a satisfactory manner. Sufficient animals must be used to have confidence in finding and characterizing any adverse drug actions that may be present. In practice, as the duration of the study increases, small doses are administered, and larger numbers of animals must be employed per group. These two features—dosage level and group size—are critical to study designs. Table 2.8 presents general guidance on the number of animals to be used in systemic studies. These and other technical considerations for the safety assessment of pharmaceuticals are present in detail in this book.

The protocols discussed thus far have focused on general or systemic toxicity assessment. The agency and, indeed, the lay public have a special set of concerns with reproductive toxicity, fetal/embryo toxicity, and developmental toxicity (also called *teratogenicity*). Collectively, these concerns often go by the acronyms *DART* (developmental and reproductive toxicity) or *RTF* (reproduction, teratogenicity, fertility). Segment II studies are more designed to detect developmental toxicity. Only pregnant females are dosed during critical period of organogenesis. Generally, the first protocol DART test (exclusive of range-finding studies) is a segment I study of rats in fertility and general reproductive performance. This is generally done while the drug is in phase II clinical trials. Alternatively, many companies are now performing the segment II teratology study in rats before the segment I study

because the former is less time and resource intensive. One or both should be completed before including women of child-bearing potential in clinical trials. The FDA requires teratogenicity testing in two species—a rodent (rat or mouse) and the rabbit. Use of the rabbit was instituted as a result of the finding that thalidomide was a positive teratogen in the rabbit but not in the rat. On occasion, when a test article is not compatible with the rabbit, teratogenicity data in the mouse may be substituted. There are also some specific classes of therapeutics (e.g., quinolone antibiotics) where segment II studies in primates are effectively required prior to product approval. Both should be completed before entering phase III clinical trials. The most complicated of the DART protocols—segment III—is generally commenced during phase III trials and should be part of the NDA. There are differences in the different national guidelines (as discussed later with international considerations) regarding the conduct of these studies. The large multinational drug companies try to design their protocols to be in compliance with as many guidelines as possible to avoid duplication of testing while allowing the broadest possible approval and marketing of therapeutics.

2.7.3 Genetic Toxicity Assessment

Genetic toxicity testing generally focuses on the potential of a new drug to cause mutations (in single-cell systems) or other forms of genetic damage. The tests, generally short in duration, often rely on *in vitro* systems and generally have a single end point of effect (point mutations, chromosomal damage, etc.). For a complete review of protocols, technology, etc., the reader is referred to Brusick (1987). It is of interest that the FDA had no standard or statutory requirement for genetic toxicity testing but generally expects to see at least some such tests performed and will ask for them if the issue is not addressed. If one performs such a study, any data collected, of course, must be sent to the agency as part of any IND, PLA, or NDA. These studies have yet to gain favor with the FDA (or other national regulatory agencies) as substitutes for *in vivo* carcinogenicity testing. However, even with completed negative carcinogenicity tests, at least some genetic toxicity assays are generally required. Generally, pharmaceuticals in the United States are evaluated for mutagenic potential (e.g., the Ames assay) or for chromosomal damage (e.g., the *in vivo* mouse micronucleus test). In general, in the United States, pharmaceutical companies apply genetic toxicity testing in the following fashion:

- *As a screen* An agent that is positive in one or more genetic toxicity tests may be more likely than one that is negative to be carcinogenic and, therefore, may not warrant further development.
- *As an adjunct* An agent that is negative in carcinogenicity testing in two species and also negative in a genetic toxicity battery is more likely than not to be noncarcinogenic in human beings.

TABLE 2.6 Synopsis of General Guidelines for Animal Toxicity Studies for Drugs

| Category | Duration of Human Administration | Clinical Phase | Subacute or Chronic Toxicity | Special Studies |
|----------------------------------|----------------------------------|-----------------|--|--|
| Oral or parenteral | Several days | I, II, III, NDA | Two species; 2 weeks | For parentally administered drugs |
| | Up to 2 weeks | I | Two species; 4 weeks | |
| | | II | Two species; up to 4 weeks | |
| | | III, NDA | Two species; up to 3 months | Compatibility with blood where applicable |
| | Up to 3 months | I, II | Two species; 4 weeks | |
| | | III | Two species; 3 months | |
| | | NDA | Two species; up to 6 months | |
| | 6 months to unlimited | I, II | Two species; 3 months | |
| | | III | Two species; 6 months or longer | |
| | | NDA | Two species; 9 months (nonrodent) and 12 months (rodent) | |
| | | | +2 rodent species for CA; 18 months (mouse); 24 months (rat). Mouse may be replaced with an allowable transgenic mouse study | |
| Inhalation (general anesthetics) | | I, II, III, NDA | Four species; 5 days (3 h day ⁻¹) | |
| Dermal | Single application | I | One species; single 24 h exposure followed by 2-week observation | Sensitization |
| | Single or short-term application | II | One species; 20-day repeated exposure (intact and abraded skin) | |
| | Short-term application | III | As aforementioned | |
| | Unlimited application | NDA | As aforementioned, but intact skin study extended up to 6 months | |
| Ophthalmic | Single application | I | | Eye irritation tests with graded doses |
| | Multiple application | I, II, III | One species; 3-week daily applications, as in clinical use | |
| | | NDA | One species; duration commensurate with period of drug administration | |
| Vaginal or rectal | Single application | I | | Local and systematic toxicity after vaginal or rectal application in two species |
| | Multiple application | I, II, III, NDA | Two species; duration and number of applications determined by proposed use | |
| Drug combinations | | I, II, III, NDA | Two species; up to 3 months | Lethality by appropriate route, compared to components run concurrently in one species |

TABLE 2.7 Duration of Repeated-Dose Toxicity Studies to Support Clinical Trials and Marketing^a

| Duration of Clinical Trials | Minimum Duration of Repeated-Dose Toxicity Studies ^b | | Duration of Clinical Trials | Minimum Duration of Repeated-Dose Toxicity Studies ^c | |
|-----------------------------|---|-----------------------|-----------------------------|---|----------------------|
| | Rodents | Nonrodents | | Rodents | Nonrodents |
| Single dose | 2 weeks ^d | 2 weeks | Up to 2 weeks | 1 month | 1 month |
| Up to 2 weeks | 2 weeks ^d | 2 weeks | Up to 1 month | 3 months | 3 months |
| Up to 1 month | 1 month | 1 month | Up to 3 months | 6 months | 3 months |
| Up to 6 months | 6 months | 6 months ^e | >3 months | 6 months | Chronic ^d |
| >6 months | 6 months | Chronic ^e | | | |

^a In Japan, if there are no phase II clinical trials of equivalent duration to the planned phase III trials, conduct of longer duration toxicity studies is recommended as given earlier.

^b Data from 6 months of administration in nonrodents should be available before the initiation of clinical trials longer than 3 months. Alternatively, if applicable, data from a 9-month nonrodent study should be available before the treatment duration exceeds that which is supported by the available toxicity studies.

^c The table also reflects the marketing recommendations in the three regions except that a chronic nonrodent study is recommended for clinical use >1 month.

^d In the United States, as an alternative to 2-week studies, single-dose toxicity studies with extended examinations can support single-dose human trials (4).

^e To support phase I and II trials in the EU and phase I, II, and III trials in the United States and Japan.

TABLE 2.8 Numbers of Animals per Dosage Group in Systemic Toxicity Studies (OECD Guidances)

| Study Duration (per Sex) | Rodents (per Sex) | Nonrodents |
|--------------------------|-------------------|--------------------------------|
| 2–4 weeks | 5 | 3 |
| 13 weeks | 20 ^a | 6 |
| 26 weeks | 30 | 8 |
| Chronic | 50 | 10 |
| Carcinogenicity | 60 ^b | Applies only to contraceptives |
| Bioassays | | Applies only to contraceptives |

^a Starting with 13-week studies, one should consider adding animals (particularly to the high dose) to allow evaluation of reversal of effects.

^b In recent years there have been decreasing levels of survival in rats on 2-year studies. What is required is that at least 20–25 animals/sex/group survive at the end of the study. Accordingly, practice is beginning to use 70 or 75 animals per sex, per group.

- *To provide mechanistic insight* For example, if an agent is negative in a wide range of genetic toxicity screens but still produces tumors in animals, then one could hypothesize that an epigenetic mechanism was involved.

While not officially required, the FDA does have the authority to request, on a case-by-case basis, specific tests it feels may be necessary to address a point of concern. A genetic toxicity test could be part of such a request. In general, therefore, companies deal with genetic toxicity (after “screening”) on a case-by-case basis, dictated by good science. If more than a single administration is intended, common practice is to perform the tests prior to submitting an IND.

2.7.4 Safety Pharmacology

Midway through 2001 ICH and the related regional regulatory authorities (such as FDA, EMA, and MHW)

implemented a new set of preclinical (to be completed before initiation of human clinical trials) safety assessment requirements focused on reversible organ function alterations that could have rapid fatal effects before reversal. The general case core set of these is the freestanding GLP evaluations of cardiovascular, respiratory, pulmonary, and central nervous system (CNS) functions. There are exceptions for the “requirements” in some structural class cases. This is discussed in detail in Chapter 18 and in Gad (2012).

2.7.5 Local Tissue Tolerance

Not called out in ICH guidances but rather in the US and other pharmacopoeia are the requirements to assess local tissue effects of drugs as they potentially can occur at or around the site of drug application or administration. These effects include irritation, pyrogenicity, hemolysis, and others. There are specific requirements (as presented in Chapter 16) for all routes except oral.

2.7.6 Toxicity Testing: Biotechnology Products

As mentioned, the regulation of traditional pharmaceuticals (small molecules such as aspirin or digitalis) and biologicals (proteins such as vaccines and antitoxins derived from animal sources) has very different histories. See the discussion on biologics earlier in this chapter. Until 1972, the NIH (or its forerunning agency, the Hygienic Laboratory of the Department of the Treasury) was charged with the responsibility of administering the Virus Act of 1902. With the passage of the food and drug laws of 1906, 1938, and 1962, there was a recurring debate regarding whether these laws applied or should apply to biologicals (Pendergast, 1984). This debate was resolved when the authority for the regulation of biologics was transferred to the FDA's new Bureau of Biologics (now the CBER) in 1972. Since then, there appears to have been little difference in the matter of regulation for biologics and pharmaceuticals. The FDA essentially regulates biologics as described under the 1902 Act but then uses the rule-making authority granted under the Food and Drug Act to "fill in the gaps."

The Bureau of Biologics was once a relatively "sleepy" agency, primarily concerned with the regulation of human blood products and vaccines used for mass immunization programs. The authors of the 1902 law could hardly have foreseen the explosion in biotechnology that occurred in the 1980s. New technology created a welter of new biological products, such as recombinant DNA (rDNA)-produced proteins (e.g., tissue plasminogen activator), biological response modifiers (cytokinins and colony-stimulating factors), monoclonal antibodies, antisense oligonucleotides, and self-directed vaccines (raising an immune response to self-proteins such as gastrin for therapeutic reasons). The new products raised a variety of new questions on the appropriateness of traditional methods for evaluating drug toxicity that generated several PTC documents. For the sake of brevity, this discussion will focus on the rDNA proteins. Some of the safety issues that have been raised over the years:

- The appropriateness of testing a human-specific peptide hormone in nonhuman species
- The potential that the peptide could break down due to nonspecific metabolism, resulting in products that had no therapeutic value or even a toxic fragment
- The potential sequelae to an immune response (formation of neutralizing antibodies, provoking an autoimmune or a hypersensitivity response), pathology due to immune precipitation, etc.
- The presence of contamination with oncogenic virus DNA (depending on whether a bacterial or mammalian system was used on the synthesizing agent) or endotoxins
- The difficulty interpreting the scientific relevance of response to supraphysiological systemic doses of potent biological response modifiers

The last few intervening years have shown some of these concerns to have been more relevant than others. The "toxic peptide fragment" concern, for example, has been shown to be without merit. The presence of potentially oncogenic virus DNA and endotoxins is a quality assurance concern and is not truly a toxicological problem. Regardless of the type of synthetic pathway, all proteins must be synthesized in compliance with Good Manufacturing Practices (GMPs). Products must be as pure as possible, not only free of rDNA but also free of other types of cell debris (endotoxin). Batch-to-batch consistency with regard to molecular structure must also be demonstrated using appropriate methods (e.g., amino acid). The regulatory thinking and experience over the last 15 years has come together in the document "S6 Preclinical Safety Evaluation of Biotechnology-Derived Pharmaceuticals" prepared by the ICH. The FDA (both CDER and CBER jointly) has published the document as a Guidance for Industry (FDA, 1997; CDER, 1998). The document intended to provide basic guidance for the preclinical evaluation of biotechnology-derived products, including proteins and peptides, either produced by cell culture using rDNA technology, but did not cover antibiotics, allergenic extracts, heparin, vitamins, cellular drug products vaccines, or other products regulated as biologics. Items covered are summarized as follows:

- *Test-article specifications* In general, the product that is used in the definitive pharmacology and toxicology studies should be comparable to the product proposed for the initial clinical studies.
- *Animal species/model selection* Safety evaluation should include the use of relevant species, in which the test article is pharmacologically active due, for example, to the expression of the appropriate receptor molecule. These can be screened with *in vitro* receptor binding assays. Safety evaluation should normally include two appropriate species, if possible and/or feasible. The potential utility of gene knockout and/or transgenic animals in safety assessment is discussed.
- *Group size* No specific numbers are given, but it does state that a small sample size may lead to failure to observe toxic events.
- *Administration* The route and frequency should be as close as possible to that proposed for clinical use. Other routes can be used when scientifically warranted.
- *Immunogenicity* It has also been clearly demonstrated in the testing of rDNA protein products that animals will develop antibodies to foreign proteins. This response has been shown to neutralize (rapidly remove from circulation) the protein, but no pathological conditions have been shown to occur as a sequelae to the immune response. Bear in mind, however, that interleukins have powerful effects on immune response, but

these are due to their physiological activity and not due to an antigen–antibody response. The first has to do with “neutralizing antibodies;” that is, is the immune response so great that the test article is being removed from circulation as fast as it is being added? If this is the case, does long-term testing of such a chemical make sense? In many cases, it does not. The safety testing of any large molecule should include the appropriate assays for determining whether the test system has developed a neutralizing antibody response. Depending on the species, route of administration, intended therapeutic use, and development of neutralizing antibodies (which generally takes about 2 weeks), it is rare for a toxicity test on an rDNA protein to be of a duration longer than 4 weeks. However, if the course of therapy in humans is to be longer than 2 weeks, formation of neutralizing antibodies must be demonstrated or longer-term testing performed. The second antigen–antibody formation concern is that a hypersensitivity response will be elicited. Traditional preclinical safety assays are generally adequate to guard against this if they are 2 weeks or longer in duration and the relevant end points are evaluated.

- *Safety pharmacology* It is important to investigate the potential for unwanted pharmacological activity in appropriate animal models and to incorporate monitoring for these activities in toxicity studies.
- *Exposure assessment* Single- and multiple-dose pharmacokinetics, toxicokinetics, and tissue distribution studies in relevant species are useful. Proteins are not given orally, demonstrating absorption and mass balance are not typically primary considerations. Rather, this segment of the test should be designed to determine half-life (and other appropriate pharmacokinetic (PK) descriptor parameters), the plasma concentration associated with biological effects, and potential changes due to the development of neutralizing antibodies.
- *Reproductive performance and developmental toxicity studies* These will be dictated by the product, clinical indication, and intended patient population.
- *Genotoxicity studies* The S6 document states that the battery of genotoxicity studies routinely conducted for traditional pharmaceuticals are not appropriate for biotechnology-derived pharmaceuticals. In contrast to small molecules, genotoxicity testing with a battery of *in vitro* and *in vivo* techniques of protein molecules has not become common US industry practice. Such tests are not formally required by the FDA but, if performed, must be reported. They are, however, required by European and Japanese regulatory authorities. This has sparked a debate as to whether or not genotoxicity testing is necessary or appropriate for rDNA protein

molecules. It is the authors’ opinion that such testing is, scientifically, of little value. Firstly, large protein molecules will not easily penetrate the cell wall of bacteria or yeast, and (depending on size, charge, lipophilicity, etc.) penetration across the plasma lemma of mammalian cells will be highly variable. Secondly, if one considers the well-established mechanism(s) of genotoxicity of small molecules, it is difficult to conceive of how a protein might act in the same fashion. For example, proteins will not be metabolized to be electrophilic active intermediates that will cross-link guanine residues. In general, therefore, genotoxicity testing with rDNA proteins is wasteful of resources. It is conceivable, however, that some proteins, because of their biological mechanism of action, may stimulate the proliferation of transformed cells. For example, it is a feasible hypothesis that a colony-stimulating factor could stimulate the proliferation of leukemic cells (it should be emphasized that this is a hypothetical situation, presented here for illustrative purposes). Again, this is a question of a specific pharmacological property, and such considerations should be tested on a case-by-case basis.

- *Carcinogenicity studies* These are generally inappropriate for biotechnology-derived pharmaceuticals; however, some products may have the potential to support or induce proliferation of transformed cells—possibly leading to neoplasia. When this concern is present, further studies in relevant animal models may be needed.

These items are covered in greater detail in the S6 guidance document and in a review by Hayes and Ryffel (1997).

So, given the previous discussion, what should the toxicology testing package of a typical rDNA protein resemble? Based on the products that have successfully wended their way through the regulatory process, the following generalizations can be drawn:

- The safety tests look remarkably similar to those for traditional tests. Most have been done on three species: the rat, the dog, or the monkey. The great difference has to do with test length. It is rare for a safety test on a protein to be more than 13 weeks long.
- The dosing regimens can be quite variable and at times very technique intensive. These chemicals are almost always administered by a parenteral route of administration, normally intravenously or subcutaneously. Dosing regimens have run the range from once every 2 weeks for an antihormone “vaccine” to continuous infusion for a short-lived protein.
- As reviewed by Ryffel (1996), most side effects in man of a therapy with rDNA therapy may be predicted

by data from experimental toxicology studies, but there are exceptions. IL-6, for example, induced a sustained increase in blood platelets and acute-phase proteins, with no increase in body temperature. In human trials, however, there were increases in temperature.

- The S6 document also mentions monoclonal antibody products. Indeed, many of the considerations for rDNA products are also applicable to monoclonal antibodies (including hybridized antibodies). With monoclonal antibodies, there is the additional concern of cross-reactivity with nontarget molecules.

As mentioned, the rapid development in the biotechnology industry has created some confusion as to what arm of the FDA is responsible for such products. In October 1992, the two major reviewing groups, CBER and CDER, reached a series of agreements to explain and organize the FDA's position on products that did not easily fall into its traditional classification schemes. CDER would continue to have responsibility for traditional chemically synthesized molecules as well as those purified from mineral or plant sources (except allergens), antibiotics, hormones (including insulin, growth hormone, etc.), most fungal or bacterial products (disaccharidase inhibitors), and most products from animal or solid human tissue sources. CBER would have responsibility for products subject to licensure (BLA), including all vaccines, human blood or blood-derived products (as well as drugs used for blood banking and transfusion), immunoglobulin products, products containing intact cells, fungi, viruses, proteins produced by cell culture or transgenic animals, and synthetic allergenic products. This situation was further simplified by the introduction of the concept of "well-characterized biologics." When introduced during the debate on FDA reform in 1996, the proposed section of S.1447 stated that "Biological products that the secretary determines to be well-characterized shall be regulated solely under the Federal Food, Drug and Cosmetic Act." Under this concept, highly purified, well-characterized therapeutic rDNA proteins would be regulated by CDER, regardless of therapeutic target (Anonymous, 1996).

2.8 TOXICITY/SAFETY TESTING: CELLULAR AND GENE THERAPY PRODUCTS

Human clinical trials of cellular and gene therapies involve administration to patients of materials considered investigational biological, drug, or device products. Somatic cell therapy refers to the administration to humans of autologous, allogenic, or xenogenic cells which have been manipulated or processed *ex vivo*. Gene therapy refers to the introduction into the human body of genes or cells containing genes

foreign to the body for the purposes of prevention, treatment, diagnosing, or curing disease.

Sponsors of cellular or gene therapy clinical trials must file an IND or in certain cases an IDE with the FDA before initiation of studies in humans. It is the responsibility of the CBER to review the application and determine if the submitted data and the investigational product meet applicable standards. The critical parameters of identity, purity, potency, stability, consistency, safety, and efficacy relevant to biological products are also relevant to cellular and gene therapy products.

In 1991, FDA first published "Points to Consider in Human Somatic Cell Therapy and Gene Therapy" (Anonymous, 1991). At the time virtually all gene therapies were retroviral and were prepared as *ex vivo* somatic cell therapies. This was subsequently reviewed by Kessler et al. (1993). While the data for certain categories of information such as that regarding molecular biology were defined in previous guidance documents relating to rDNA products, the standards for preclinical and clinical development were less well defined. The field has advanced to include not only new vectors but also novel routes of administration. "Points to Consider in Human Somatic Cell Therapy and Gene Therapy" was thus amended in 1996 (Leibert, 1996) to reflect both advancements in product development and more importantly the accumulation of safety information.

FDA regulations state that the sponsor must submit, in the IND, adequate information about pharmacological and toxicological studies of the drug including laboratory animals or *in vitro* studies on the basis of which the sponsor has considered that it is reasonably safe to conduct the proposed clinical investigation. For cellular and gene therapies, designing and conducting relevant preclinical safety testing has been a challenge to both the FDA and to the sponsors. For genes delivered using viral vectors, the safety of the vector system *per se* must be considered and evaluated.

The preclinical knowledge base is initially developed by designing studies to answer fundamental questions. The development of this knowledge base is generally applicable to most pharmaceuticals as well as biopharmaceuticals and includes data to support (i) the relationship of the dose to biological activity, (ii) the relationship of the dose to toxicity, (iii) the effect of route and/or schedule on activity or toxicity, and (iv) identification of the potential risks for subsequent clinical studies. These questions are considered in the context of indication and/or disease state. In addition there are often unique concerns in relation to the specific category or product class.

For cellular therapies safety concerns may include development of a database from studies specifically designed to answer questions relating to growth factor dependence, tumorigenicity, local and systemic toxicity, and effects on

host immune responses including immune activation and altered susceptibility to disease. For viral-mediated gene therapies, specific questions may relate to the potential for overexpression of the transduced gene, transduction of normal cells/tissues, genetic transfer to germ cells and subsequent alterations to the genome, recombination/rescue with endogenous virus, reconstitutions of replication competence, potential for insertional mutagenesis/malignant transformation, altered susceptibility to disease, and/or potential risk(s) to the environment.

To date cellular and gene therapy products submitted to FDA have included clinical studies indicated for bone marrow marking, cancer, cystic fibrosis, AIDS, and inborn errors of metabolism and infectious diseases. Of the current active INDs approximately 78% have been sponsored by individual investigators or academic institutions, and 22% have also been industry sponsored. In addition to the variety of clinical indications, the cell types have also been varied. Examples include tumor-infiltrating lymphocytes (TIL) and lymphocyte-activated killer (LAK) cells, selected cells from bone marrow and peripheral blood lymphocytes (e.g., stem cells), myoblasts, tumor cells, and encapsulated cells (e.g., islet cells and adrenal chromaffin cells).

2.8.1 Cellular Therapies

Since 1984 CBER has reviewed close to 300 somatic cell therapy protocols. Examples of the specific categories include manipulation, selection, mobilization, tumor vaccines, and other.

Manipulation Autologous, allogenic, or xenogenic cells which have been expanded, propagated, or manipulated or had their biological characteristics altered *ex vivo* (e.g., TIL or LAK cells; islet cells housed in a membrane).

Selection Products designed for positive or negative selection if autologous or allogenic cells intended for therapy (e.g., purging of tumor from bone marrow, selection of CD34+ cells).

Mobilization *In vivo* mobilization of autologous stem cells intended for transplantation.

Tumor vaccines Autologous or allogenic tumor cells which are administered as vaccines (e.g., tumor cell lines, tumor cell lysates, primary explant. See FDA (1993)). This group also includes autologous antigen-presenting cells pulsed with tumor-specific peptides or tumor cell lysates.

Other Autologous, allogenic, and xenogenic cells which do not specifically fit above. This group includes cellular therapies such as extracorporeal liver assist devices.

2.8.2 Gene Therapies

The types of vectors that have been used, or proposed, for gene transduction include retrovirus, adenovirus, adeno-associated viruses, other viruses (e.g., herpes, vaccinia, etc.), and plasmid DNA. Methods for gene introduction include *ex vivo* replacement, drug delivery, marker studies, and others and *in vivo* viral vectors, plasmid vectors, and vector producer cells.

2.8.3 Ex Vivo

Replacement Cells transduced with a vector expressing a normal gene in order to correct or replace the function of a defective gene

Drug delivery Cells transduced with a vector expressing a gene encoding a therapeutic molecule which can be novel or native to the host

Marker studies Cells (e.g., bone marrow, stem cells) transduced with a vector expressing a marker or reporter gene used to distinguish it from other similar host tissues

Other Products which do not specifically fit under above (e.g., tumor vaccines in which cells are cultured or transduced *ex vivo* with a vector)

2.8.4 In Vivo

Viral vectors The direct administration of a viral vector (e.g., retrovirus, adenovirus, adeno-associated virus, herpes, vaccinia) to patients

Plasmid vectors The direct administration of plasmid vectors with or without other vehicles (e.g., lipids) to patients

Vector producer cells The direct administration of retroviral vector producer cells (e.g., murine cells producing HTK vector) to patients

2.8.5 Preclinical Safety Evaluation

The goal of preclinical safety evaluation includes recommendation of an initial safe starting dose and safe dose-escalation scheme in humans, identification of potential target organ(s) of toxicity, identification of appropriate parameters for clinical monitoring, and identification of "at-risk" patient population(s). Therefore, when feasible, toxicity studies should be performed in relevant species to assess a dose-limiting toxicity. General considerations in study design include selection of the model (e.g., species, alternative model, animal model of disease), dose (e.g., route, frequency, and duration) and study end point (e.g., activity and/or toxicity).

The approach to preclinical safety evaluation of biotechnology-derived products, including novel cellular and gene therapies, has been referred to as the "case-by-case" approach. This approach is science based, data driven, and

flexible. The major distinction from past practices from traditional pharmaceuticals is that the focus is directed at asking specific questions across various product categories. Additionally, there is a consistent reevaluation of the knowledge base to reassess real or theoretical safety concerns and hence reevaluation of the need to answer the same questions across all product categories. In some cases there may even be conditions which may not need specific toxicity studies, for example, when there is a strong efficacy model which is rationally designed to answer specific questions and/or there is previous human experience with a similar product with respect to dose and regimen.

2.8.6 Basic Principles for Preclinical Safety Evaluation of Cellular and Gene Therapies

Biotechnology-derived products in general

- Use of product in animal studies that is comparable or the same as the product proposed for clinical trial(s)
- Adherence to basic principles of GLP to ensure quality of the study including a detailed protocol prepared prospectively
- Use of the same or similar route and method of administration as proposed for clinical trials (whenever possible)
- Determination of appropriate doses delivered based upon preliminary activity obtained from both *in vitro* and *in vivo* studies (i.e., finding a dose likely to be effective and not dangerous, a no-observed-adverse-effect level, and a dose causing dose-limiting toxicity)
- Selection of one or more species sensitive to the end point being measured, for example, infections or pathologic sequelae and/or biological activity or receptor binding
- Consideration of animal model(s) of disease that may be better to assess the contribution of changes in physiologic or underlying physiology to safety and efficacy
- Determination of affect on host immune response
- Localization/distribution studies—evaluation of target tissue, normal surrounding tissue, and distal tissue sites and any alteration in normal or expected distribution
- Local reactogenicity

2.8.7 Additional Considerations for Cellular Therapies

- Evaluation of cytopathogenicity
- Evaluation of signs of cell transformation/growth factor dependence effect on animal cells, normal human cells, and cells prone to transform easily
- Determination of alteration in cell phenotype, altered cell products, and/or function
- Tumorigenicity

2.8.8 Additional Considerations for Gene Therapies

- Determination of phenotype/activation state of effector cells
- Determination of vector/transgene toxicity
- Determination of potential transfer to germline
- *In vitro* challenge studies—evaluation of recombination or complementation, potential for “rescue” for subsequent infection with wild-type virus
- Determination of persistence of cells/vector
- Determination of potential for insertional mutagenesis (malignant transformation)
- Determination of environmental spread (e.g., viral shedding)

2.9 TOXICITY TESTING: SPECIAL CASES

On paper, the general case guidelines for the evaluation of the safety of drugs are relatively straightforward and well understood. However, there are also a number of special case situations under which either special rules apply or some additional requirements are relevant. The more common of these are summarized as follows.

2.9.1 Oral Contraceptives

Oral contraceptives are subject to special testing requirements. These have recently been modified so that in addition to those preclinical safety tests generally required, the following are also required (Berliner, 1974):

- A 3-year carcinogenicity study in beagles (this is a 1987 modification in practice from earlier FDA requirements and the 1974 publication)
- A rat reproductive (segment I) study including a demonstration of return to fertility

2.9.2 Life-Threatening Diseases (Compassionate Use)

Drugs to treat life-threatening diseases are not strictly held to the sequence of testing requirements as put forth in Table 2.3 because the potential benefit on any effective therapy in these situations is so high. In the early 1990s, this situation applied to AIDS-associated diseases and cancer. The development of more effective HIV therapies (protease inhibitors) has now made cancer therapy more the focus of these considerations. Though the requirements for safety testing prior to initial human trials are unchanged, subsequent requirements are flexible and subject to negotiation and close consultation with FDA's Division of Oncology (within CDER) (FDA, 1988). The more recent thinking on anticancer agents has been reviewed by DeGeorge et al. (1998). The preclinical studies that will be required to support clinical trials and

marketing of new anticancer agents will depend on the mechanism of action and the target clinical population. Toxicity studies in animals will be required to support initial clinical trials. These studies have multiple goals:

- Determine a starting dose for clinical trials
- Identify target organ toxicity and assess recovery
- Assist in the design of clinical dosing regimens

The studies should generally conform to the protocols recommended by the National Cancer Institute, as discussed by Greishaber (1991). In general, it can be assumed that most antineoplastic cytotoxic agents will be highly toxic. Two studies are essential to support initial clinical trials (IND phase) in patients with advanced disease. These are studies of 5–14 days in length, but with longer recovery periods. A study in rodents is required that identifies those doses that produce either life-threatening or nonlife-threatening toxicity. Using the information from this first study, a second study in nonrodents (generally the dog) is conducted to determine if the tolerable dose in rodents produces life-threatening toxicity. Doses are compared on a milligram-per-square-meter basis. The starting dose in initial clinical trials is generally one-tenth of that required to produce severe toxicity in rodents (STD10) or one-tenth the highest dose in nonrodents that does not cause severe irreversible toxicity. While not required, information on PK parameters, especially data comparing the plasma concentration associated with toxicity in both species, is very highly regarded. Special attention is paid to organs with high cell division rates, bone marrow, testes, lymphoid tissue testing, and gastrointestinal (GI) tract. As these agents are almost always given intravenously, special attention needs to be given relatively early in development to intravenous irritation and blood compatibility study. Subsequent studies to support the NDA will be highly tailored, depending on the following:

- Therapeutic indication and mechanism of action
- The results of the initial clinical trials
- The nature of the toxicity
- Proposed clinical regimen

Even at the NDA stage, toxicity studies with more than 28 days of dosing are rarely required. While not required for the IND, assessment of genotoxicity and developmental toxicity will need to be addressed. For genotoxicity, it will be important to establish the ratio between cytotoxicity and mutagenicity. *In vivo* models, for example, the mouse micronucleus test, can be particularly important in demonstrating the lack of genotoxicity at otherwise subtoxic doses. For developmental toxicity, ICH stage C–D studies (traditionally known as segment II studies for teratogenicity in rat and rabbits) will also be necessary.

The emphasis of this discussion has been on purely cytotoxic neoplastic agents. Additional considerations must be given to cytotoxic agents that are administered under special circumstances: those that are photoactivated, delivered as liposomal emulsions, or delivered as antibody conjugates. These types of agents will require additional studies. For example, a liposomal agent will need to be compared to the free agent and a blank liposomal preparation. There are also studies that may be required for a particular class of agents. For example, anthracyclines are known to be cardiotoxic, so comparison of a new anthracycline agent to previously marketed anthracyclines will be expected.

In addition to antineoplastic, cytotoxic agents, there are cancer therapeutic or preventative drugs that are intended to be given on a chronic basis. This includes chemopreventatives, hormonal agents, immunomodulators, etc. The toxicity assessment studies on these will more closely resemble those of more traditional pharmaceutical agents. Chronic toxicity, carcinogenicity, and full developmental toxicity (ICH A–B, C–D, E–F) assessments will be required. For a more complete review, the reader is referred to DeGeorge et al. (1998).

2.9.3 Optical Isomers

The FDA (and like regulatory agencies, as reviewed by Daniels et al. (1997)) has become increasingly concerned with the safety of stereoisomeric or chiral drugs. Stereoisomers are molecules that are identical to one another in terms of atomic formula and covalent bonding but differ in the three-dimensional projections of the atoms. Within this class are those molecules that are nonsuperimposable mirror images of one another. These are called enantiomers (normally designated as R- or S-). Enantiomeric pairs of a molecule have identical physical and chemical characteristics except for the rotation of polarized light. Drugs have generally been mixtures of optical isomers (enantiomers), because of the difficulties in separating the isomers. It has become apparent in recent years, however, that these different isomers may have different degrees of both desirable therapeutic and undesirable toxicologic effects. Technology has also improved to the extent that it is now possible to perform chiral specific syntheses, separations, and/or analyses. It is now highly desirable from a regulatory (FDA, 1988; De Camp, 1989; Anonymous, 1992/2015; FDA, 2015) basis to develop a single isomer unless all isomers have equivalent pharmacological and toxicologic activity. The FDA has divided enantiomeric mixtures in the following categories:

- Both isomers have similar pharmacologic activity, which could be identical, or they could differ in the degrees of efficacy.

- One isomer is pharmacologically active, while the other is inactive.
- Each isomer has completely different activity.

During preclinical assessment of an enantiomeric mixture, it may be important to determine to which of these three classes it belongs. The pharmacological and toxicological properties of the individual isomers should be characterized. The PK profile of each isomer should be characterized in animal models with regard to disposition and interconversion. It is not at all unusual for each enantiomer to have a completely different PK behavior.

If the test article is an enantiomer isolated from a mixture that is already well characterized (e.g., already on the market), then appropriate bridging guides need to be performed which compare the toxicity of the isomer to that of the racemic mixture. The most common approach would be to conduct a subchronic (3 months) and a segment II type teratology study with an appropriate “positive” control group which received the racemate. In most instances no additional studies would be required if the enantiomer and the racemate did not differ in toxicity profile. If, on the other hand, differences are identified, the reasons for this difference need to be investigated and the potential implications for human subjects need to be considered.

2.9.4 Special Populations: Pediatric and Geriatric Claims

Relatively few drugs marketed in the United States (~20%) have pediatric dosing information available. Clinical trials had rarely been done specifically on pediatric patients. Traditionally, dosing regimens for children have been derived empirically by extrapolating on the basis of body weight or surface area. This approach assumes that the pediatric patient is a young adult, which simply may not be the case. There are many examples of how adults and children differ qualitatively in metabolic and/or pharmacodynamic responses to pharmaceutical agents. In their review, Shacter and DeSantis (1998) state, “The benefit of having appropriate usage information in the product label is that health care practitioners are given the information necessary to administer drugs and biologics in a manner that maximizes safety, minimizes unexpected adverse events, and optimizes treatment efficacy. Without specific knowledge of potential drug effects, children may be placed at risk. In addition, the absence of appropriate proscribing information, drugs and biologics that represent new therapeutic advances may not be administered to the pediatric population in a timely manner.” In response to the need for pediatric information, the FDA had developed a pediatric plan. This two-phase plan called first for the development of pediatric information on marketed drugs. The second phase focused on new drugs. The implementation of the plan was to be coordinated by the

Pediatric Subcommittee of the Medical Policy Coordinating Committee of CDER. The Pediatric Use Labeling Rule was a direct result of phase I in 1994 (PhRMA, 1998). Phase II resulted in 1997 from a proposed rule entitled “Pediatric Patients: Regulations Requiring Manufacturers to Assess the Safety and Effectiveness of New Drugs and Biologics.” Soon after this rule was proposed, the FDAMA of 1997 was passed. FDAMA contained provisions that specifically addressed the needs and requirements for the development of drugs for the pediatric population.

The FDAMA bill essentially codified and expanded several regulatory actions initiated by the FDA during the 1990s. Among the incentives offered by the bill, companies will be offered an additional 6 months of patent protection for performing pediatric studies (clinical trials) on already approved products. In fact, the FDA was mandated by the FDAMA to develop a list of over 500 drugs for which additional information would produce benefits for pediatric patients. The FDA is supposed to provide a written request for pediatric studies to the manufacturers (Hart, 1999).

In response to the pediatric initiatives, the FDA has published policies and guidelines and conducted a variety of meetings. CDER has established a website (<http://www.fda.gov/cder/pediatric>) which lists three pages of such information. Interestingly, the focus has been on clinical trials, and almost no attention has been given to the preclinical toxicology studies that may be necessary to support such trials. There are three pages of documents on the pediatric website. None appear to address the issue of appropriate testing. This is a situation that is just now being addressed and is in a great deal of flux.

In the absence of any guidelines from the agency for testing drugs in young or “pediatric” animals, one must fall back on the maxim of designing a program that makes the most scientific sense. As a guide, the FDA designated levels of postnatal human development and the approximate equivalent ages (in the author’s considered opinion) in various animal models are given in Table 2.9. The table is somewhat inaccurate, however, because of difference in the stages of development at birth. A rat is born quite underdeveloped when compared to a human being. A 1-day-old rat is not equivalent to a 1-day-old full-term human infant. A 4-day-old rat would be more appropriate. In terms of development, the pig may be the best model of those listed; however, one should bear in mind that different organs have different developmental schedules in different species.

Table 2.9 can be used as a rough guide in designing toxicity assessment experiments in developing animals. In designing of the treatment period, one needs to consider not only the dose and the proposed course of clinical treatment but also the proposed age of the patient and whether or not an equivalent dosing period in the selected animal model covers more than one developmental stage. For example, if the proposed patient population is human infants, initiating a

TABLE 2.9 Comparison of Postnatal Development Stages

| Stage | Human | Rat | Dog | Pig |
|------------|------------------|-----------------|------------------|------------------|
| Neonate | Birth to 1 month | Birth to 1 week | Birth to 3 weeks | Birth to 2 weeks |
| Infant | 1 month–2 years | 1–3 weeks | 3–6 weeks | 2–4 weeks |
| Child | 2–12 years | 3–9 weeks | 6 weeks–5 months | 4 weeks–4 months |
| Adolescent | 12–16 years | 9–13 weeks | 5–9 months | 4–7 months |
| Adult | Over 16 years | Over 13 weeks | Over 9 months | Over 7 months |

toxicity study of the new pharmaceutical agent in 3-day-old rats is not appropriate. Furthermore, if the proposed course of treatment in adult children is 2 weeks, it is unlikely that this would cross over into a different developmental stage. A 2-week treatment initiated in puppies, however, might easily span two developmental stages. Thus, in designing an experiment in young animals, one must carefully consider the length of the treatment period balancing the developmental age of the animal model and the proposed length of clinical treatment. Where appropriate (infant animals), one needs to also assess changes in standard developmental landmarks (e.g., eye opening, pinnae eruption, external genitalia development, etc.) as well as the more standard indicators of target organ toxicity. The need for maintaining the experimental animals past the dosing period, perhaps into sexual maturity, to assess recovery or delayed effects needs also to be carefully considered.

To summarize, the current status of assessment of toxicity in postnatal mammals, in response to the pediatric initiatives covered in FDAMA, is an extremely fluid situation. One needs to carefully consider a variety of factors in designing the study and should discuss proposed testing programs with the appropriate office at CDER.

Drugs intended for use in the elderly, like those intended for the very young, may also have special requirements for safety evaluation, but geriatric issues were not addressed in the FDAMA of 1997. The FDA has published a separate guidance document for geriatric labeling (CDER and CBER, 2001). As was the case with pediatric guidance, this document does not address preclinical testing. With the elderly, the toxicological concerns are quite different than the developmental concerns associated with pediatric patients. With the elderly, one must be concerned with the possible interactions between the test article and compromised organ function. The FDA had previously issued a guidance for clinically examining clinical safety of new pharmaceutical agents in patients with compromised renal and/or hepatic function (CDER, 1989). The equivalent ICH guideline (S5A) was issued in 1994. Whether this type of emphasis will require toxicity testing in animal models with specifically induced organ insufficiency remains to be seen. In the interim, we must realize that there is tacit evaluation of test-article-related toxicity in geriatric rodents for those agents that undergo 2-year carcinogenicity testing. As the graying of America continues, labeling for geriatric use may become more of an issue in the future.

As presented in Table 2.10 there are four special case INDs that lead to earlier approval of drugs for special cases. The prototype for these would be the orphan drug route.

2.9.5 Orphan Drugs

The development of sophisticated technologies, coupled with the rigors and time required for clinical and preclinical testing, has made pharmaceutical development very expensive. In order to recoup such expenses, pharmaceutical companies have tended to focus on therapeutic agents with large potential markets. Treatments for rare but life-threatening diseases have been “orphaned” as a result. An orphan product is defined as one targeted at a disease which affects 200,000 or fewer individuals in the United States. Alternatively, the therapy may be targeted for more than 200,000, but the developer would have no hope of recovering the initial investment without exclusivity. The Orphan Drug Act (ODA) of 1983 was passed in an attempt to address this state of affairs. Currently applicable regulations were put in place in 1992 and amended in 2013 (Anonymous, 2013). In 1994, there was an attempt in Congress to amend the Act, but it failed to be passed into law. The current regulations are administered by the Office of Orphan Products Development (OOPD). The Act offers the following incentives to encourage the development of products to treat rare diseases:

- Seven-year exclusive market following the approval of a product for an orphan disease
- Written protocol assistance from the FDA
- Tax credits for up to 50% of qualified clinical research expenses
- Available grant to support pivotal clinical trials

As reviewed by Haffner (1998), other developed countries have similar regulations.

There are significant misconceptions about the orphan drug process (Tambuyzer, 2010). The ODA did not change the requirements of testing drug products. The nonclinical testing programs are similar to those used for more conventional products. They will undergo the same FDA review process. A major difference, however, is the involvement of the OPD. A sponsor must request OPD review. Once OPD determines that a drug meets the criteria for orphan drug

TABLE 2.10 Comparison of FDA's Expedited Programs for Serious Conditions (CDER and CBER, 2014)

| Nature of program | Fast-Track | | Breakthrough Therapy | | Accelerated Approval | | Priority Review | |
|----------------------------|---|--|--|--|---|--|---|--|
| | Designation | | Designation | | Approval Pathway | | Designation | |
| Reference | <ul style="list-style-type: none"> Section 506(b) of the FD&C Act, as added by Section 112 of the Food and Drug Administration Modernization Act of 1997 (FDAMA) and amended by Section 901 of the Food and Drug Administration Safety and Innovation Act of 2012 (FDASIA) | | <ul style="list-style-type: none"> Section 506(a) of the FD&C Act, as added by Section 902 of FDASIA | | <ul style="list-style-type: none"> 21 CFR Part 314, Subpart H 21 CFR Part 601, Subpart E Section 506(c) of the FD&C Act, as amended by Section 901 of FDASIA | | <ul style="list-style-type: none"> Prescription Drug User Fee Act of 1992 | |
| Qualifying criteria | <ul style="list-style-type: none"> A drug that is intended to treat a serious condition AND nonclinical or clinical data demonstrate the potential to address unmet medical need OR A drug that has been designated as a qualified infectious disease product^e | | <ul style="list-style-type: none"> A drug that is intended to treat a serious condition AND preliminary clinical evidence indicates that the drug may demonstrate substantial improvement on a clinically significant end point(s) over available therapies | | <ul style="list-style-type: none"> A drug that treats a serious condition AND generally provides a meaningful advantage over available therapies AND demonstrates an effect on a surrogate end point that is reasonably likely to predict clinical benefit or on a clinical end point that can be measured earlier than irreversible morbidity or mortality (IMM) that is reasonably likely to predict an effect on IMM or other clinical benefit (i.e., an intermediate clinical end point) | | <ul style="list-style-type: none"> An application (original or efficacy supplement) for a drug that treats a serious condition AND, if approved, would provide a significant improvement in safety or effectiveness OR Any supplement that proposes a labeling change pursuant to a report on a pediatric study under 505A^b OR An application for a drug that has been designated as a qualified infectious disease product^c OR Any application or supplement for a drug submitted with a priority review voucher^d | |
| When to submit request | <ul style="list-style-type: none"> With IND or after Ideally, no later than the pre-BLA or pre-NDA meeting | | <ul style="list-style-type: none"> With IND or after Ideally, no later than the end-of-phase II meeting | | <ul style="list-style-type: none"> The sponsor should ordinarily discuss the possibility of accelerated approval with the review division during development, supporting, for example, the use of the planned end point as a basis for approval and discussing the confirmatory trials, which should usually be already under way at the time of approval | | <ul style="list-style-type: none"> With original BLA, NDA, or efficacy supplement | |
| Timelines for FDA response | <ul style="list-style-type: none"> Within 60 calendar days of receipt of the request | | <ul style="list-style-type: none"> Within 60 calendar days of receipt of the request | | <ul style="list-style-type: none"> Not specified | | <ul style="list-style-type: none"> Within 60 calendar days of receipt of original BLA, NDA, or efficacy supplement | |

(Continued)

TABLE 2.10 (Continued)

| Nature of program | Fast-Track | Breakthrough Therapy | | Accelerated Approval | Priority Review |
|---------------------------|---|--|--|---|-----------------|
| | Designation | Designation | | Approval Pathway | Designation |
| Features | <ul style="list-style-type: none">• Actions to expedite development and review• Rolling review | <ul style="list-style-type: none">• Intensive guidance on efficient drug development• Organizational commitment• Rolling review• Other actions to expedite review | <ul style="list-style-type: none">• Approval based on an effect on a surrogate end point or an intermediate clinical end point that is reasonably likely to predict a drug's clinical benefit | <ul style="list-style-type: none">• Shorter clock for review of marketing application (6 months compared with the 10-month standard review)^e | |
| Additional considerations | <ul style="list-style-type: none">• Designation may be rescinded if it no longer meets the qualifying criteria for fast track^f | <ul style="list-style-type: none">• Designation may be rescinded if it no longer meets the qualifying criteria for breakthrough therapy^g | <ul style="list-style-type: none">• Promotional materials• Confirmatory trials to verify and describe the anticipated effect on IMM or other clinical benefit• Subject to expedited withdrawal | <ul style="list-style-type: none">• Designation will be assigned at the time of original BLA, NDA, or efficacy supplement filing | |

Source: Extracted from CDER and CBER (2014).

^a Title VIII of FDASIA, "Generating Antibiotic Incentives Now (GAIN)," provides incentives for the development of antibacterial and antifungal drugs for human use intended to treat serious and life-threatening infections. Under GAIN, a drug may be designated as a qualified infectious disease product (QIDP) if it meets the criteria outlined in the statute. A drug that receives QIDP designation is eligible under the statute for fast-track designation and priority review. However, QIDP designation is beyond the scope of this guidance.

^b Any supplement to an application under Section 505 of the FD&C Act that proposes a labeling change pursuant to a report on a pediatric study under this section shall be considered a priority review supplement per Section 505A of the FD&C Act as amended by Section 5(b) of the Best Pharmaceuticals for Children Act.

^c See footnote a above.

^d Any application or supplement that is submitted with a priority review voucher will be assigned a priority review. Priority review vouchers will be granted to applicants of applications for drugs for the treatment or prevention of certain tropical diseases, as defined in Section 524(a)(3) and (a)(4) of the FD&C Act and for treatment of rare pediatric diseases as defined in Section 529(a)(3) of the FD&C Act.

^e As part of its commitments in PDUFA V, the FDA has established a review model, the Program. The Program applies to all new molecular entity NDAs and original BLAs, including applications that are resubmitted following a refuse-to-file action, received from October 1, 2012, through September 30, 2017. For applications filed by FDA under the Program, the PDUFA review clock will begin at the conclusion of the 60 calendar day filing review period that begins on the date of FDA receipt of the original submission.

^f A sponsor may also withdraw fast-track designation if the designation is no longer supported by emerging data or the drug development program is no longer being pursued (see Section A.5).

^g A sponsor may also withdraw breakthrough therapy designation if the designation is no longer supported by emerging data or the drug development program is no longer being pursued (see Section B.5).

status, it will work with the sponsor to provide the assistance required under the Act. The ODA does not review a product for approval. The IND/NDA process is still handled by the appropriate reviewing division (e.g., Cardiovascular) for formal review. The Act does not waive the necessity for submission of an IND, not for the responsibility of toxicological assessment. As always, in cases where there is ambiguity, a sponsor may be well served to request a pre-IND meeting at the appropriate division to discuss the acceptability of a toxicology assessment plan.

2.9.6 Botanical Drug Products

There is an old saying, “What goes around comes around,” and so it is with botanicals. At the beginning of the twentieth century, most marketed pharmaceutical agents were botanical in origin. For example, aspirin was first isolated from willow bark. These led the way in the middle part of the century, for reasons having to do with patentability, manufacturing costs, standardization, selectivity, and potency. The dawning of the twenty-first century has seen a grassroots return to botanical preparations (also sold as herbals or dietary supplements). These preparations are being marketed to the lay public as “natural” supplements to the nasty synthetic chemicals now proscribed as pharmaceutical products. In 1994, the Dietary Supplement Health and Education Act was passed which permitted the marketing of dietary supplements (including botanicals) with limited submissions to the FDA (Wu et al., 2000). If a producer makes a claim that an herbal preparation is beneficial to a specific part of the body (e.g., enhanced memory), then it may be marketed after a 75-day period of FDA review but without formal approval. On the other hand, if any curative properties are claimed, then the botanical will be regulated as a drug and producers will be required to follow the IND/NDA process. In 1997 and 1998 combined, some 26 INDs were filed for botanical products (Wu et al., 2000).

The weakness in the current regulation has to do with its ambiguity. The line between a beneficial claim and a curative claim is sometimes difficult to draw. What is the difference, for example, between an agent that enhances memory and one that prevents memory loss? Given the number of products and claims hitting the shelves every day,

this situation will probably demand increased regulatory scrutiny in the future.

2.9.7 Types of New Drug Applications (NDAs)

Actual product approvals for drugs are one form or another of NDA. While in this volume we focus on the traditional (505(b)(1)), there are two others for small molecules—505(b)(2) Applications and Abbreviated New Drug Application (ANDA) (for generic drug applications). These have minimal if any nonclinical safety requirements. While these are US FDA terms for the non-NME drug approvals, equivalents exist in other major regulatory paradigms (see, e.g., EOC Directive 2001/83/EC, amended in July of 2008).

2.10 INTERNATIONAL PHARMACEUTICAL REGULATION AND REGISTRATION

2.10.1 International Conference on Harmonization

The ICH was established to make the drug regulatory process more efficient in the United States, Europe, and Japan. The US involvement grew out of the fact that the United States is party to the General Agreement on Tariffs and Trade, which included the Agreement on Technical Barriers to Trade, negotiated in the 1970s, to encourage reduction of nontariff barriers to trade (Barton, 1998). The main purpose of ICH is, through harmonization, to make new medicines available to patients with a minimum of delay. More recently, the need to harmonize regulation has been driven, according to ICH, by the escalation of the cost of R&D. The regulatory systems in all countries have the same fundamental concerns about safety, efficacy, and quality, yet sponsors had to repeat many time-consuming and expensive technical tests to meet country-specific requirements. Secondarily, there was a legitimate concern over the unnecessary use of animals. Conference participants include representatives from the drug regulatory bodies and research-based pharmaceutical industrial organizations of three regions; the European Union (EU), the United States, and Japan were over 90% of world's pharmaceutical industry. Representation is summarized in Table 2.11. The biennial conference met regularly

TABLE 2.11 ICH Representation

| Country/Region | Regulatory | Industry |
|-------------------------|---|--|
| European Union | European Commission (2) | European Federation of Pharmaceutical Industries and Associations (2) |
| Japan | Ministry of Health and Welfare (2) | Japan Pharmaceutical Manufacturers Association (2) |
| United States | Food and Drug Administration (2) | Pharmaceutical Research and Manufacturers of America (2) |
| Observing organizations | World Health Organization, European Free Trade Association, Canadian Health Protection Branch | International Federation of Pharmaceutical Manufacturers & Associations (2); also provides the secretariat |

Numbers in parentheses are number of representatives on the ICH Steering Committee.

TABLE 2.12 Steps in ICH Guideline Development and Implementation

| | |
|---|--|
| 1 | Building scientific consensus in joint regulatory/industry expert working groups |
| 2 | Agreement by the steering committee to release the draft consensus text for wider consultation |
| 3 | Regulatory consultation in the three regions. Consolidation of the comments |
| 4 | Agreement on a harmonized ICH guideline; adopted by the regulators ^a |
| 5 | Implementation in the three ICH regions ^a |

^a ICH (1997).

beginning in 1991, rotating between sites in the United States, Europe, and Japan.

The ICH meets its objectives by issuing guidelines for the manufacturing, development, and testing of new pharmaceutical agents that are acceptable to all three major parties. For each new guideline, the ICH Steering Committee establishes an expert working group with representation from each of the six major participatory ICH bodies. Each new draft guideline goes through the five various steps of review and revision summarized in Table 2.12. So far, ICH has proposed or adopted over 40 safety, efficacy, and quality guidelines (listed in Table 2.13) for use by the drug regulatory agencies in the United States, Europe, and Japan. Those guidelines specifically applying to nonclinical drug safety evaluation, in their most current state, are listed in Table 2.14.

The guidelines are organized under broad categories: the “E” series having to do with clinical trials, the “Q” series having to do with quality (including chemical manufacturing and control as well as traditional GLP issues), and the “S” series having to do with safety. Guidelines may be obtained from the ICH secretariat, c/o of IFPMA, 30 rue de St.-Jean, PO Box 9, 1211 Geneva 18, Switzerland, or may be downloaded directly from the ICH website (<http://www.ich.org/products/guidelines.html>). They are also published in the Federal Register. It is the guidelines of the “S” series that will have the most impact on toxicologists. The biggest changes having to do with toxicological assessment are summarized as follows.

2.10.1.1 Carcinogenicity Studies Carcinogenicity studies are covered in Guidelines S1A, S1B, and S1C. The guidelines are almost more philosophical than they are technical. In comparison to the EPA guidelines, for example, the ICH guidelines contain little in the way on concrete study criteria (e.g., the number of animals, the necessity for clinical chemistry, etc.). There is discussion on when carcinogenicity studies should be done, whether two species are more appropriate than one, and how to set dosages on the basis of human clinical PK data. The major changes being wrought by these guidelines are the following:

- Only one 2-year carcinogenicity study should be generally required. Ideally, the species chosen should be the one most like man in terms of metabolic transformations of the test article.
- The traditional second long-term carcinogenicity study can be replaced by a shorter-term alternative model. In practical terms, this guideline is beginning to result in sponsors conducting a 2-year study in the rat and a 6-month study in an alternative mouse model, such as the P53 or the TG.AC genetically manipulated mouse strains.
- In the absence of target organ toxicity with which to set the high dose at the maximally tolerated dose, the high dose can be set at the dose that produces an area under the curve (AUC). This is 25-fold higher than that obtained in human subjects.

2.10.1.2 Chronic Toxicity Traditionally, chronic toxicity of new pharmaceuticals in the United States was assessed in studies of 1-year duration in both the rodent and the nonrodent species of choice. The European view was that studies of 6 months are generally sufficient. The resulting guideline (S4A) was a compromise. Studies of 6-month duration were recommended for the rodent, as rodents would also be examined in 2-year studies. For the nonrodent (dog, non-human primate, and pig), studies of 9-month duration were recommended.

2.10.1.3 Developmental and Reproductive Toxicity This was an area in which there was considerable international disagreement and the area in which ICH has promulgated the most technically detailed guidelines (S5A and S5B). Some of the major changes include the following:

- The traditional segment I, II, and III nomenclature has been replaced with different nomenclature, as summarized in Table 2.15.
- The dosing period of the pregnant animals during studies on embryonic development (traditional segment II) studies has been standardized.
- New guidelines for fertility assessment (traditional segment I) studies that have shortened the premating dosing schedule (e.g., in male rats from 10 to 4 weeks). There has been an increased interest in assessment of spermatogenesis and sperm function.
- The new guidelines allow for a combination of studies in which the end point typically assessed in the traditional segment II and segment III studies is now examined under a single protocol.

For a more complete review of the various study designs, the reader is referred to the review by Manson (1994).

TABLE 2.13 International Conference on Harmonization Guidelines

| References | Guideline | Date |
|------------|---|---|
| E1 | The Extent of Population Exposure to Assess Clinical Safety | October 1994 |
| E2A | Clinical Safety Data Management: Definitions and Standards for Expedited Reporting | October 1994 |
| E2B(R3) | Clinical Safety Data Management: Data Elements for Transmission of Individual Case Safety Reports | November 2014 |
| E2C(R2) | Clinical Safety Data Management: Periodic Safety Update Reports for Marketed Drugs | November 2012 |
| E2D | Definitions and Standards for Expedited Reporting | November 2003 |
| E2E | Pharmacovigilance Planning | November 2004 |
| E2F | Development Safety Update Report | August 2010 |
| E3 | Structure and Content of Clinical Study Reports | November 1995 |
| E4 | Dose-Response Information to Support Drug Registration | March 1994 |
| E5(R1) | Ethnic Factors in the Acceptability of Foreign Clinical Data | February 1998 |
| E6(R1) | Good Clinical Practice | May 1996 (R2 draft integrated addendum: June 2015) |
| E7 | Studies in Support of Special Populations: Geriatrics | June 1993 |
| E8 | Guidance on General Considerations for Clinical Trials; Notice | July 1997 |
| E9 | Guideline on Statistical Principles for Clinical Trials; Notice of Availability | February 1998 |
| E10 | Choice of Control Group and Related Issues in Clinical Trials | July 2000 |
| E11 | Clinical Investigation of Medicinal Products in the Pediatric Population | July 2000 (R1 concept paper: August 2014) |
| E12 | Principles for Clinical Evaluation of New Antihypertensive Drugs | March 2000 |
| E14 | The Clinical Evaluation of QT/QTc Interval Prolongation and Proarrhythmic Potential for Non-Antiarrhythmic Drugs | May 2005 |
| E15 | Definitions for Genomic Biomarkers, Pharmacogenomics, Pharmacogenetics, Genomic Data and Sample Coding Categories | November 2007 |
| E16 | Biomarkers Related to Drug or Biotechnology Product Development: Context, Structure and Format of Qualification Submissions | August 2010 |
| E17 | General Principle on Planning/Designing Multi-Regional Clinical Trials | May 2016 (Current Step 2 Version) |
| E18 | Draft Guideline: Genomic Sampling and Management of Genomic Data | December 2015 |
| M3(R2) | Guidance on Nonclinical Safety Studies for the Conduct of Human Clinical Trials for Pharmaceuticals | June 2009 |
| Q1A(R2) | Stability Testing of New Drug Substances and Products | February 2003 |
| Q1B | Stability Testing of New Drug Substances and Products | November 1996 |
| Q1C | Stability Testing for New Dosage Forms | November 1996 |
| Q1D | Bracketing and Matrixing Designs for Stability Testing of Drug Substances and Drug Products | February 2002 |
| Q1E | Evaluation of Stability Data | February 2003 |
| Q2(R1) | Validation of Analytical Procedures: Text and Methodology | October 1994 |
| Q3A(R2) | Impurities in New Drug Substances | October 2006 |
| Q3B(R2) | Impurities in New Drug Products | June 2006 |

(Continued)

TABLE 2.13 (Continued)

| References | Guideline | Date |
|---------------------|---|--|
| Q3C(R5) | Impurities: Guideline for Residual Solvents | February 2011 (R6 draft revision: June 2015) |
| Q3D | Guidelines for Elemental Impurities | December 2014 |
| Q4 | Pharmacopoeias | November 2007 |
| Q4A | Pharmacopoeial Harmonisation | November 2007 |
| Q4B | Evaluation and Recommendation of Pharmacopoeial Texts | November 2007 |
| Q4B Annex 1(R1) | Evaluation and Recommendation of Pharmacopoeial Texts: Residue on Ignition/Sulphated Ash General Chapter | September 2010 |
| Q4B Annex 2(R1) | Evaluation and Recommendation of Pharmacopoeial Texts: Test for Extractable Volume of Parenteral Preparations General Chapter | September 2010 |
| Q4B Annex 3(R1) | Evaluation and Recommendation of Pharmacopoeial Texts: Test for Particulate Contamination: Sub-Visible Particles General Chapter | September 2010 |
| Q4B Annex 4A(R1) | Microbiological Examination of Non-Sterile Products: Microbial Enumeration Tests General Chapter | September 2010 |
| Q4B Annex 4B(R1) | Microbiological Examination of Non-Sterile Products: Tests for Specified Micro-Organisms General Chapter | September 2010 |
| Q4B Annex 4C(R1) | Microbiological Examination of Non-Sterile Products: Acceptance Criteria for Pharmaceutical Preparation Substances for Pharmaceutical Use General Chapter | September 2010 |
| Q4B Annex 5(R1) | Disintegration Test General Chapter | September 2010 |
| Q4B Annex 6 | Uniformity of Dosage Units General Chapter | November 2013 |
| Q4B Annex 7(R2) | Dissolution Test General Chapter | November 2010 |
| Q4B Annex 8(R1) | Sterility Test General Chapter | September 2010 |
| Q4B Annex 9(R1) | Tablet Friability General Chapter | September 2010 |
| Q4B Annex 10(R1) | Polyacrylamide Gel Electrophoresis General Chapter | September 2010 |
| Q4B Annex 11 | Capillary Electrophoresis General Chapter | June 2010 |
| Q4B Annex 12 | Analytical Sieving General Chapter | June 2010 |
| Q4B Annex 13 | Bulk Density and Tapped Density of Powders General Chapter | June 2012 |
| Q4B Annex 14 | Bacterial Endotoxins Test General Chapter | October 2012 |
| Q5A(R1) | Quality of Biotechnological Products: Viral Safety Evaluation of Biotechnology Products Derived from Cell Lines of Human or Animal Origin | September 1999 |

(Continued)

TABLE 2.13 (Continued)

| References | Guideline | Date |
|---|--|---------------|
| Q5B | Quality of Biotechnology Products: Analysis of the Expression Construct in Cells Used for Production of r-DNA Derived Protein Product | November 1995 |
| Q5C | Quality of Biotechnological Products: Stability Testing of Biotechnological/Biology Products | November 1995 |
| Q5D | Availability of Draft Guideline on Quality of Biotechnological/Biological Products: Derivation and Characterization of Cell Substrates Used for Production of Biotechnological/Biological Products | July 1997 |
| Q5E | Comparability of Biotechnological/Biological Products Subject to Changes in their Manufacturing Process | November 2004 |
| Q6A | Specifications: Test Procedures and Acceptance Criteria for New Drug Substances and New Drug Products: Chemical Substances (including Decision Trees) | October 1999 |
| Q6B | Specifications: Test Procedures and Acceptance Criteria for Biotechnological/Biological Products | March 1999 |
| Q7 | Good Manufacturing Practice Guide for Active Pharmaceutical Ingredients | November 2000 |
| Q8(R2) | Pharmaceutical Development | August 2009 |
| Q9 | <i>Quality Risk Management</i> | November 2005 |
| Q10 | <i>Pharmaceutical Quality System</i> | June 2008 |
| Q11 | Development and Manufacture of Drug Substances (Chemical Entities and Biotechnological/Biological Entities) | May 2012 |
| Q12 | Final Concept Paper: Technical and Regulatory Considerations for Pharmaceutical Product Lifecycle Management | July 2014 |
| See Table 2.14 for current safety guidance list | | |

While they were not quite as sweeping in approach as the aforementioned guidelines, a toxicologist working in pharmaceutical safety assessment should become familiar with the all the other ICH guidelines in the S series.

In an interesting recent article, Ohno (1999) discussed not the harmonization of nonclinical guidelines but also the need to harmonize the timing of nonclinical tests in relation to the conduct of clinical trials. For example, there are regional differences in the inclusion of women of childbearing potential in clinical trials. In the United States, including woman in such trials is becoming more important, and therefore evaluation of embryo-fetal development will occur earlier in the drug development process than in Japan. Whether or not such timing or staging of nonclinical tests becomes part of an ICH guideline in the near future remains to be established.

2.10.2 Other International Considerations

The United States is the single largest pharmaceutical market in the world. But the rest of the world (particularly, but not limited to the second and third largest markets, Japan and the EU) represents in aggregate a much larger market, so no one develops a new pharmaceutical for marketing in just the United States. The effort at harmonization (exemplified by

the ICH) has significantly reduced differences in requirements for these other countries, but certainly not obliterated them. Though a detailed understanding of their regulatory schemes is beyond this volume, the bare bones and differences in toxicology requirements are not.

2.10.2.1 European Union The standard EU toxicology and pharmacologic data requirements for a pharmaceutical include:

- Single-dose toxicity
- Repeat-dose toxicity (subacute and chronic trials)
- Reproduction studies (fertility and general reproductive performance, embryotoxicity, and peri-/postnatal toxicity)
- Mutagenic potential (*in vitro* and *in vivo*)
- Carcinogenicity
- Pharmacodynamics
 - Effects related to proposed drug indication
 - General pharmacodynamics
 - Drug interactions
- Pharmacokinetics
 - Single dose
 - Repeat dose

TABLE 2.14 ICH Current Guidelines Governing Nonclinical Safety Evaluation

| Reference Number and Classification | Title | Adopted Originally | Revisions | Link to Document |
|--|---|---|---|---|
| S1 Carcinogenicity studies | Rodent Carcinogenicity Studies for Human Pharmaceuticals | November 14, 2012 Endorsed as Final Concept Paper | N/A | http://www.ich.org/fileadmin/Public_Web_Site/ICH_Products/Guidelines/Safety/S1/S1_Concept_Paper_14_November_2012.pdf |
| S1A Carcinogenicity studies | Guideline on the Need for Carcinogenicity Study of Pharmaceuticals | November 29, 1995 | N/A | http://www.ich.org/fileadmin/Public_Web_Site/ICH_Products/Guidelines/Safety/S1A/Step4/S1A_Guideline.pdf |
| S1B Carcinogenicity studies | Testing for Carcinogenicity of Pharmaceuticals | July 16, 1997 | N/A | http://www.ich.org/fileadmin/Public_Web_Site/ICH_Products/Guidelines/Safety/S1B/Step4/S1B_Guideline.pdf |
| S1C(R2) Carcinogenicity studies | Dose Selection for Carcinogenicity Studies of Pharmaceuticals | October 27, 1994 | R1: November, 2005 R2: March 11, 2008 | http://www.ich.org/fileadmin/Public_Web_Site/ICH_Products/Guidelines/Safety/S1C_R2/Step4/S1C_R2_Guideline.pdf |
| S2(R1) Genotoxicity studies | Guidance on Genotoxicity Testing and Data Interpretation for Pharmaceuticals Intended for Human Use | November 9, 2011 | The tripartite harmonized ICH guideline was finalized under <i>Step 4</i> in November 2011. It replaces and combines the ICH S2A and S2B guidelines | http://www.ich.org/fileadmin/Public_Web_Site/ICH_Products/Guidelines/Safety/S2_R1/Step4/S2R1_Step4.pdf |
| S3A Toxicokinetics and pharmacokinetics | Note for Guidance on Toxicokinetics: The Assessment of Systemic Exposure in Toxicity Studies | October 27, 1994 | N/A | http://www.ich.org/fileadmin/Public_Web_Site/ICH_Products/Guidelines/Safety/S3A/Step4/S3A_Guideline.pdf |
| S3B Toxicokinetics and pharmacokinetics | Pharmacokinetics: Guidance for Repeated Dose Tissue Distribution Studies | October 27, 1994 | N/A | http://www.ich.org/fileadmin/Public_Web_Site/ICH_Products/Guidelines/Safety/S3B/Step4/S3B_Guideline.pdf |
| S4 Toxicity testing | Duration of Chronic Toxicity Testing in Animals (Rodent and Nonrodent Toxicity Testing) | September 2, 1998 | N/A | http://www.ich.org/fileadmin/Public_Web_Site/ICH_Products/Guidelines/Safety/S4/Step4/S4_Guideline.pdf |
| S5(R2) Reproductive toxicology | Detection of Toxicity to Reproduction for Medicinal Products and Toxicity to Male Fertility | June 24, 1993 | Addendum dated November 9, 2000 incorporated in November 2005 R3: Concept paper endorsed by steering | http://www.ich.org/fileadmin/Public_Web_Site/ICH_Products/Guidelines/Safety/S5/Step4/S5_R2_Guideline.pdf |
| S6(R1) Biotechnological products | Preclinical Safety Evaluation of Biotechnology-Derived Pharmaceuticals | July 16, 1997 | Addendum dated June 12, 2011 incorporated at the end of June 2011 | http://www.ich.org/fileadmin/Public_Web_Site/ICH_Products/Guidelines/Safety/S6_R1/Step4/S6_R1_Guideline.pdf |

| | | | | |
|--|---|---------------------------------|---|--|
| S7A | Safety Pharmacology Studies for Human Pharmaceuticals | November 8, 2000 | N/A | http://www.ich.org/fileadmin/Public_Web_Site/ICH_Products/Guidelines/Safety/S7A/Step4/S7A_Guideline.pdf |
| S7B | The Nonclinical Evaluation of the Potential for Delayed Ventricular Repolarization (QT Interval Prolongation) by Human Pharmaceuticals | May 12, 2005 | N/A | http://www.ich.org/fileadmin/Public_Web_Site/ICH_Products/Guidelines/Safety/S7B/Step4/S7B_Guideline.pdf |
| S8 | Immunotoxicity Studies for Human Pharmaceuticals | September 15, 2005 | N/A | http://www.ich.org/fileadmin/Public_Web_Site/ICH_Products/Guidelines/Safety/S8/Step4/S8_Guideline.pdf |
| S9 | Nonclinical Evaluation for Anticancer Pharmaceuticals | October 29, 2009 | N/A | http://www.ich.org/fileadmin/Public_Web_Site/ICH_Products/Guidelines/Safety/S9/Step4/S9_Step4_Guideline.pdf |
| S10 | Photosafety Evaluation of Pharmaceuticals | November 13, 2013 | N/A | http://www.ich.org/fileadmin/Public_Web_Site/ICH_Products/Guidelines/Safety/S10/S10_Step_4.pdf |
| S11 | Nonclinical Safety Testing in Support of Development of Pediatric Medicines | November 10, 2014 | N/A | http://www.ich.org/fileadmin/Public_Web_Site/ICH_Products/Guidelines/Safety/S11/S11_Final_Concept_Paper_10_November_2014.pdf |
| Juvenile toxicity | | Endorsed as final concept paper | | |
| M3R2 | Guidance on Nonclinical Safety Studies for the Conduct of Human Clinical Trials and Marketing Authorization for Pharmaceuticals | July 16, 1997 | R1: November 9, 2000 R2: June 11, 2009 | http://www.ich.org/fileadmin/Public_Web_Site/ICH_Products/Guidelines/Multidisciplinary/M3_R2/Step4/M3_R2_Guideline.pdf |
| M7R1 | Assessment and Control of DNA Reactive (Mutagenic) Impurities in Pharmaceuticals to Limit Potential Carcinogenic Risk <i>And</i> Addendum | June 23, 2014 | R1: June 9, 2015 | http://www.ich.org/fileadmin/Public_Web_Site/ICH_Products/Guidelines/Multidisciplinary/M7/M7_Step_4.pdf <i>and</i> http://www.ich.org/fileadmin/Public_Web_Site/ICH_Products/Guidelines/Multidisciplinary/M7/M7_Addendum_Step_2.pdf |
| Updated April 26, 2016 | | | | |
| All guidelines may also be accessed through the ICH website at http://www.ich.org/products/guidelines.html | | | | |

TABLE 2.15 Comparison of Traditional and ICH Guidelines for Reproductive and Developmental Toxicology

| Traditional Protocol | Stages Covered | ICH Protocol | Dosing Regimen |
|---|--|---|---|
| Segment I (rats) | A. Premating to conception B. Conception to implantation | Fertility and early embryonic development, including implantation | Males: 4 weeks premating, mating (1–3 weeks) plus 3 weeks postmating Females: 2 weeks premating, mating through day 7 of gestation |
| Segment II (rabbits) | C. Implantation to closure of hard palate D. Closure of hard palate to the end of pregnancy | Embryo-fetal development | Females: day 6–20 of pregnancy |
| Study Title | Termination | Endpoints: In-Life | Endpoints: Postmortem |
| Fertility and early embryonic development, including implantation | Females: Day 13–15 of pregnancy Males: Day after completion of dosing | Clinical signs and mortality Body weights and feed intake Vaginal cytology | Macroscopic examination plus histology on gross lesions Collection of reproductive organs for possible histology Quantitation of corpora lutea and implantation sites Seminology (count, motility and morphology) |
| Embryo-fetal development | | Clinical signs and mortality Body weights and changes Feed intake | Macroscopic examination plus histology on gross lesions Quantitation of corpora lutea and implantation sites Fetal body weights Fetal abnormalities |
| Pre- and postnatal development, including maternal function | | Clinical signs and mortality Body weights and changes Feed intake Duration of pregnancy Parturition | Macroscopic examination plus histology on gross lesions Implantation Abnormalities (including terata) Live/dead offspring at birth Pre- and postweaning survival and growth (F ₁) Physical development (F ₁) Sensory functions and reflexes (F ₁) Behavior (F ₁) |

- Distribution in normal and pregnant animals
- Biotransformation
- Local tissue tolerance
- Environmental toxicity

In general, the registration process in the EU allows one to either apply to an overall medicines authority or to an individual national authority. Either of these steps is supposed to lead to mutual recognition by all the individual members.

2.10.2.2 Japan In Japan, the Koseisho is the national regulatory body for new drugs. The standard LD₅₀ test is no longer a regulatory requirement for new medicines in the United States, the EU, or Japan. The Japanese guidelines were the first to be amended in accordance with this agreement, with the revised guidelines becoming effective in August 1993. The Japanese may still anticipate that

single-dose (acute) toxicity studies should be conducted in at least two species, one rodent and one nonrodent (the rabbit is not accepted as a nonrodent). Both males and females should be included from at least one of the species selected: if the rodent, then a minimum of five per sex and if the nonrodent, at least two per sex. In nonrodents, both the oral and parenteral routes should be used, and normally the clinical route of administration should be employed. In nonrodents, only the intended route of administration needs to be employed; if the intended route of administration in humans is intravenous, then use of this route in both species is acceptable. An appropriate number of doses should be employed to obtain a complete toxicity profile and to establish any dose–response relationship. The severity, onset, progression, and reversibility of toxicity should be studied during a 14-day follow-up period, with all animals being necropsied. When macroscopic changes are noted, the tissue must be subjected to histological examination.

Chronic and subchronic toxicity studies are conducted to define the dose level, when given repeatedly, that cause toxicity and the dose level that does not lead to toxic findings. In Japan, such studies are referred to as repeated-dose toxicity studies. As with single-dose studies, at least two animal species should be used, one rodent and one nonrodent (rabbit not acceptable). In rodent studies, each group should consist of at least 10 males and 10 females; in nonrodent species, three of each sex are deemed adequate. Where interim examinations are planned, however, the numbers of animals employed should be increased accordingly. The planned route of administration in human subjects is normally explored. The duration of the study will be dictated by the planned duration of clinical use (Table 2.16).

At least three different dose groups should be included, with the goals of demonstrating an overtly toxic dose and a no-effect dose and establishing any dose–response relationship. The establishment of a nontoxic dose within the framework of these studies is more rigorously adhered to in Japan than elsewhere in the world. All surviving animals should also be necropsied, either at the completion of the study or during its extension recovery period, to assess reversal of toxicity and the possible appearance of delayed toxicity. Full histological examination is mandated on all nonrodent animals used in a chronic toxicity study; at a minimum, the highest-dose and control groups of rodents must be submitted to a full histological examination.

While the value of repeated-dose testing beyond 6 months has been questioned (Lumley et al., 1992), such testing is a regulatory requirement for a number of agencies, including the US FDA and the Koseisho. In Japan, repeated-dose testing for 12 months is required only for new medicines expected to be administered to humans for periods in excess of 6 months (Yakuji Nippo, 1994). At the first ICH held in Brussels, the consensus was that 12-month toxicity studies in rodents could be reduced to 6 months where carcinogenicity

studies are required. While not yet adopted in the Japanese guidelines, 6-month repeated-dose toxicity studies have been accepted by the agencies of all three regions. Japan—like the EU—accepts a 6-month duration if accompanied by a carcinogenicity study. The United States still requires a 9-month nonrodent study.

With regard to reproductive toxicology, as a consequence of the first ICH, the United States, the EU, and Japan agreed to recommend mutual recognition of their respective current guidelines. A tripartite harmonized guideline on reproductive toxicology has achieved ICH step 4 status and should be incorporated into the local regulations of all three regions soon. This agreement represents a very significant achievement that should eliminate many obstacles to drug registration.

Preclinical Male Fertility Studies Before conducting a single-dose male volunteer study in Japan, it is usually necessary to have completed a preclinical male fertility study (segment 1) that has an in-life phase of 10 or more weeks (i.e., 10 weeks of dosing, plus follow-up). Although government guidelines do not require this study to be completed before phase I trials begin, the responsible institutional review board, or the investigator usually imposes this condition. Japanese regulatory authorities are aware that the segment 1 male fertility study is of poor predictive value. The rat, which is used in this study, produces a marked excess of sperm. Many scientists therefore believe that the test is less sensitive than the evaluation of testicular weight and histology that constitute part of the routine toxicology assessment

Female Reproductive Studies Before entering a female into a clinical study, it is necessary to have completed the entire reproductive toxicology program, which consists of the following studies:

- Segment 1: Fertility studies in the rat or mouse species used in the segment 2 program
- Segment 2: Teratology studies in the rat or mouse and the rabbit
- Segment 3: Late gestation and lactation studies in a species used in the segment 2 studies

Such studies usually take approximately 2 years. Although the US regulations state the need for completion of segments 1 and 2 and the demonstration of efficacy in male patients, where appropriate, before entering females into a clinical program, the current trend in the United States is toward relaxation of the requirements to encourage investigation of the drug both earlier and in a larger number of females during product development. Growing pressure for the earlier inclusion of women in drug testing may encourage selection of this issue as a future ICH topic. The trend in the United

TABLE 2.16 Required Duration of Dosing in Nonclinical Study to Support Clinical Dosing

| Duration of Dosing in Toxicity Study (months) | Duration of Human Exposure |
|---|--|
| 1 | Single dose or repeated dosage not exceeding 1 week |
| 3 | Repeated dosing exceeding 1 week and to a maximum of 4 weeks |
| 6 | Repeated dosing exceeding 4 weeks and to a maximum of 6 months |
| 12 ^a | Repeated dosing exceeding 6 months or where this is deemed to be appropriate |

Source: New Drugs Division Notification No. 43, June 1992. CDER and CBER (2014), ICH (1997).

^aWhere carcinogenicity studies are to be conducted, the Koseisho had agreed to forego chronic dosage beyond 6 months.

States and the EU toward including women earlier in the critical program has not yet been embraced in Japan, however.

The three tests required in Japan for genotoxicity evaluation are a bacterial gene mutation test, *in vitro* cytogenetics, and *in vivo* tests for genetic damage. The Japanese regulations state these tests to be the minimum requirement and encourage additional tests. Currently, Japanese guidelines do not require a mammalian cell gene mutation assay. Harmonization will likely be achieved by the Koseisho recommending all four tests, which will match requirements in the United States and the EU; at present, this topic is at step 1 in the ICH harmonization process. The mutagenicity studies should be completed before the commencement of phase II clinical studies.

Guidelines presented at the second ICH are likely to alter the preclinical requirements for registration in Japan; they cover toxicokinetics and when to conduct repeated-dose tissue distribution studies. The former document may improve the ability of animal toxicology studies to predict possible adverse events in humans; currently, there are not toxicokinetic requirements in Japan, and their relevance is questioned by many there. Although there is general agreement on the registration requirement for single-dose tissue distribution studies, implementation of the repeated-dose study requirement has been inconsistent across the three ICH parties.

2.10.3 Safety Pharmacology

Japan was the first major country to required extensive pharmacological profiling on all new pharmaceutical agents as part of the safety assessment profile. Prior to commencement of initial clinical studies, the drug's pharmacology must be characterized in animal models. In the United States and Europe, these studies have been collectively called safety pharmacology studies. For a good general review of the issues surrounding safety pharmacology, the reader is referred to Hite (1997). The Japanese guidelines for such characterizations were published in 1991. They include:

- Effects on general activity and behavior
- Effects on the CNS
- Effects on the autonomic nervous system and smooth muscle
- Effects on the respiratory and cardiovascular systems
- Effects on the digestive system
- Effects on water and electrolyte metabolism
- Other important pharmacological effects

Source: New Drugs Division Notification No. 4, January 1991.

In the United States, pharmacological studies in demonstration of efficacy have always been required, but specific safety pharmacological studies have never been required.

TABLE 2.17 Composition of the Common Technical Document (ICH Format)

| Module | |
|--------|-------------------------------------|
| 1 | Regional administrative information |
| 2 | Quality overall summary |
| | Nonclinical overview |
| | Nonclinical summary |
| | Clinical overview |
| | Clinical summary |
| 3 | Quality data |
| 4 | Nonclinical study reports |
| 5 | Clinical study reports |

Special situational or mechanistic data would be requested on a case-by-case basis. This is a situation that is changing. In the United States the activities of the Safety Pharmacology Discussion Group, for example, have helped bring attention to the utility and issues surrounding safety pharmacology data. In 1999 and 2000, the major toxicological and pharmacological societal meetings had symposia on safety pharmacological testing. Many major US pharmaceutical companies are in the process of implementing programs in safety pharmacology. The issue has been taken up by ICH and the draft guideline is currently at the initial stages of review. This initial draft (Guideline S7) includes core tests in the assessment of CNS, cardiovascular, and respiratory function. Studies will be expected to be performed under GLP guidelines.

Even with harmonization as per ICH, there remain significant variations over the length of the entire process that takes a drug through to market (Hirako et al., 2007; Gad, 2012; Brock et al., 2013). These require guidance from a knowledgeable team of experts over the course of the process. This is especially true for emerging markets such as China (Deng and Kaitin, 2004). But the promulgation and near complete acceptance of a single format (the Common Technical Document—CTD) for worldwide regulatory submissions (see Table 2.17 for an outline of components) has been a huge step for global harmonization.

2.11 COMBINATION PRODUCTS

Recent years have seen a vast increase in the number of new therapeutic products which are not purely drug, device, or biologic, but rather a combination of two or more of these. This leads to a problem of deciding which of the three centers shall have ultimate jurisdiction.

The Center for Devices and Radiological Health (CDRH) is designated the center for major policy development and for the promulgation and interpretation of procedural regulations for medical devices under the Act. The CDRH regulates all medical devices inclusive of radiation-related device that are not assigned categorically or specifically to

CDER. In addition, the CDRH will independently administer the following activities (references to “Sections” are the provisions of the Act):

- A. Small business assistance programs under Section 10 of the amendments (See PL 94-295). Both CDER and CDRH will identify any unique problems relating to medical device regulation for small business.
- B. Registration and listing under Section 510 including some CDER-administered device applications. CDER will receive printouts and other assistance, as requested.
- C. Color additives under Section 706, with review by CDER, as appropriate.
- D. GMPs Advisory Committee. Under Section 520(f) (3), CDER will regularly receive notices of all meetings, with participation by CDER, as appropriate.
- E. Medical Device Reporting. The manufacturers, distributors, importers, and users of all devices, including those regulated by CDER, shall report to CDRH under Section 519 of the Act as required. The CDRH will provide monthly reports and special reports as needed to CDER for investigation and follow-up of those medical devices regulated by CDER.

2.11.1 Device Programs That CDER and CBRH Each Will Administer

Both CDER and CDRH will administer and, as appropriate, enforce the following activities for medical devices assigned to their respective centers (references to “Sections” are the provisions of the Act):

- A. Surveillance and compliance actions involving general controls violations, such as misbranded or adulterated devices under Section 301, 501, and 502
- B. Warning letters, seizures, injunctions, and prosecutions under Section 302, 303, and 304
- C. Civil penalties under Section 303(f) and administrative restraint under Section 304(g)
- D. Nonregulatory activities, such as educational programs directed at users, participation in voluntary standards organizations, etc.
- E. Promulgation of performance standards and applications of special controls under Section 514
- F. Premarket notification; investigational device exemptions including humanitarian exemptions; premarket approval; product development protocols; classification; device tracking; petitions for reclassification; postmarket surveillance under Sections 510(k), 513, 515, 519, 520(g) & (m), and 522; and the advisory committees necessary to support these activities
- G. Banned devices under Section 516

- H. FDA-requested and firm-initiated recalls whether under Section 518 or another authority and other Section 518 remedies such as recall orders
- I. Exemptions, variances, and applications of CGMP regulations under Section 520(f)
- J. Government-wide quality assurance program
- K. Requests for export approval under Sections 801(e) and 802

2.11.2 Coordination

The centers will coordinate their activities in order to assure that manufacturers do not have to independently secure authorization to market their product from both centers unless this requirement is specified in Section VII.

2.11.3 Submissions

Submissions should be made to the appropriate center, as specified herein, at the addresses provided as follows:

Address update:

Food and Drug Administration
 Center for Drug Evaluation and Research
 Central Document Room (Room #2–14)
 12420 Parklawn Drive
 Rockville, MD 20852
 or
 Food and Drug Administration
 Center for Devices and Radiological Health
 Document Mail Center (HFZ-401)
 1390 Piccard Drive
 Rockville, MD 20850

For submissions involving medical devices and/or drugs that are not clearly addressed in this agreement, sponsors are referred to the product jurisdiction regulations (21 CFR Part 3). These regulations have been promulgated to facilitate the determination of regulatory jurisdiction but do not exclude the possibility for a collaborative review between the centers.

2.11.3.1 Center Jurisdiction The following subsections provide details concerning status, market approval authority, special label/regulatory considerations, investigational options, and intercenter consultations for the categories of products specified. Section VII provides the general criteria that CDRH and CDER will apply in reaching decisions as to which center will regulate a product.

- A. 1. a. Device with primary purpose of delivering or aiding in the delivery of a drug that is distributed without a drug (i.e., unfilled)

Examples

Devices that calculate drug dosages
 Drug delivery pump and/or catheter infusion pump for implantation
 Iontophoresis device
 Medical or surgical kit (e.g., tray) with reference in instructions for use with specific drug (e.g., local anesthetic)
 Nebulizer
 Small particle aerosol generator (SPAG) for administering drug to ventilated patient
 Splitter block for mixing nitrous oxide and oxygen
 Syringe, jet injector, and storage and dispensing equipment

Status Device and drug, as separate entities.

Market approval authority CDRH and CDER, respectively, unless the intended use of the two products, through labeling, creates a combination product.

Special label/regulatory considerations The following specific procedures will apply depending on the status of the drug delivery device and drugs that will be delivered with the device:

- (i) It may be determined during the design or conduct of clinical trials for a new drug that it is not possible to develop adequate performance specifications data on those characteristics of the device that are required for the safe and effective use of the drug. If this is the case, then drug labeling cannot be written to contain information that makes it possible for the user to substitute a generic, marketed device for the device used during developments to use with the marketed drug. In these situations, CDER will be the lead center for regulation of the device under the device authorities.
- (ii) For a device intended for use with a category of drugs that are on the market, CDRH will be the lead center for regulation for the device under the device authorities. The effects of the device use on drug stability must be addressed in the device submission, when relevant. An additional showing of clinical effectiveness of the drug when delivered by the specific device will generally not be required. The device and drug labeling must be mutually conforming with respect to indication, general mode of delivery (e.g., topical, IV), and drug dosage/schedule equivalents.
- (iii) For a drug delivery device and drug that are developed for marketing to be used together as a system, a lead center will be designated to be the contact point with the manufacturer(s). If a

drug has been developed and marketed and the development and studying of device technology predominate, the principal mode of action will be deemed to be that of the device, and CDRH would have the lead. If a device has been developed and marketed and the development and studying of drug predominate, then, correspondingly, CDER would have the lead. If neither the drug nor the device is on the market, the lead center will be determined on a case-by-case basis.

Investigation options IDE or IND, as appropriate.

Intercenter consultation CDER, when lead center, will consult with CDRH if CDER determines that a specific device is required as part of the NDA process. CDRH as lead center will consult with CDER if the device is intended for use with a marketed drug and the device creates a significant change in the intended use, mode of delivery (e.g., topical, IV), or dose/schedule of the drug.

- H. Device with primary purpose of delivering or aiding in the delivery of a drug and distributed containing a drug (i.e., “prefilled delivery system”)

Examples

Nebulizer
 Oxygen tank for therapy and OTC emergency use
 Prefilled syringe
 Transdermal patch

Status Combination product.

Market approval authority CDER using drug authorities and device authorities, as necessary.

Special label/regulatory considerations None.

Investigation options IND.

Intercenter consultations Optional.

2. Device incorporating a drug component with the combination product having the primary intended purpose of fulfilling a device function

Examples

Bone cement containing antimicrobial agent
 Cardiac pacemaker lead with steroid-coated tip
 Condom, diaphragm, or cervical cap with contraceptive or antimicrobial agent (including virucidal) agent
 Dental device with fluoride
 Dental wood wedge with hemostatic agent
 Percutaneous cuff (e.g., for a catheter or orthopedic pin) coated/impregnated with antimicrobial agent
 Skin closure or bandage with antimicrobial agent
 Surgical or barrier drape with antimicrobial agent
 Tissue graft with antimicrobial or other drug agent

Urinary and vascular catheter coated/impregnated with antimicrobial agent

Wound dressing with antimicrobial agent

Status Combination product.

Market approval authority CDRH using device authorities.

Special label/regulatory considerations These products have a drug component that is present to augment the safety and/or efficacy of the device.

Investigation options IDE.

Intercenter consultation Required if a drug or the chemical form of the drug has not been legally marketed in the United States as a human drug for the intended effect.

3. Drug incorporating a device component with the combination product having the primary intended purpose of fulfilling a drug function

Examples

Skin-prep pads with antimicrobial agent

Surgical scrub brush with antimicrobial agent

Status Combination product.

Market approval authority CDER using drug authorities and, as necessary, device authorities.

Special label/regulatory considerations Marketing of such a device requires a submission of an NDA with safety and efficacy data on the drug component or it meets monograph specifications as generally recognized as safe (GRAS) and generally recognized as effective (GRAE). Drug requirements, for example, CGMPs, registration and listing, and experience reporting, apply to products.

Investigation options IND.

Intercenter consultation Optional.

4. a. Device used in the production of a drug either to deliver directly to a patient or for the use in the producing medical facility (excluding use in a registered drug manufacturing facility)

Examples

Oxygen concentrators (home or hospital)

Oxygen generator (chemical)

Ozone generator

Status Device

Market approval authority CDER, applying both drug and device authorities.

Special label/regulatory consideration May also require an NDA if the drug produced is a new drug. Device requirements (e.g., CGMPs, registration and listing, experience reporting) will apply to products.

Investigation options IDA, or NDA, as appropriate,

Intercenter consultation Optional.

- b. Drug/device combination product intended to process a drug into a finished package form

Examples

Device that uses drug concentrates to prepare large-volume parenterals

Oxygen concentrator (hospital) output used to fill oxygen tanks for use within that medical facility

Status Combination product.

Market approval authority CDER, applying both drug and device authorities.

Special label/regulatory considerations Respective drug and device requirements (e.g., CGMPs, registration and listing, experience reporting) will apply.

Investigation options IDE or NDA, as appropriate.

Intercenter consultation Optional, but will be routinely obtained.

- B. 1. Device used concomitantly with a drug to directly activate or to augment drug effectiveness

Examples

Biliary lithotripter used in conjunction with dissolution agent

Cancer hyperthermia used in conjunction with chemotherapy

Current generator used in conjunction with an implanted silver electrode (drug) that produces silver ions for an antimicrobial purpose

Materials for blocking blood flow temporarily to restrict chemotherapy drug to the intended site of action

UV and/or laser activation of oxsoralen for psoriasis or cutaneous T-cell lymphoma

Status Device and drug, as separate entities.

Market approval authority CDRH and CDER, respectively.

Special label/regulatory considerations The device and drug labeling must be mutually conforming with respect to indications, general mode of delivery (e.g., topical, IV), and drug dosage/schedule equivalence. A lead center will be designated to be the contact point with the manufacturer. If a drug has been developed and approved for another use and the development and studying of device technology predominate, then CDRH would have lead. If a device has been developed and marketed for another use and the development

and studying of drug action predominate, then CDER would have lead. If neither the drug nor the device is on the market, the lead center will be determined on a case-by-case basis. If the labeling of the drug and device creates a combination product, as defined in the combination product regulations, then the designation of the lead center for both applications will be based upon a determination of the product's primary mode of action.

Investigation options IDE or IND, as appropriate.

Intercenter consultations Required.

2. Device kits labeled for use with drugs that include both device(s) and drug(s) as separate entities in one package with the overall primary intended purpose of the kit fulfilling a device function

Examples

Medical or surgical kit (e.g., tray) with drug component
Status Combination product.

Market approval authority CDRH, using device authorities, is responsible for the kit if the manufacturer is repackaging a market drug. Responsibility for overall packaging resides with CDRH. CDER will be consulted as necessary on the use of drug authorities for the repackaged drug component.

Special label/regulatory consideration Device requirements (e.g., CGMPs, registration and listing, experience reporting) apply to kits. Device manufacturers must assure that manufacturing steps do not adversely affect drug components of the kit. If the manufacturing steps do affect the marketed drug (e.g., the kit is sterilized by irradiation), ANDA or NDA would also be required with CDRH as lead center.

Investigation options IDA or IND, as appropriate.

Intercenter consultation Optional if ANDA or NDA not required.

- C. Liquids, gases, or solids intended for use as devices (e.g., implanted, or components, parts, or accessories to devices)

Examples

Dye for tissues used in conjunction with laser surgery to enhance absorption of laser light in target tissue

Gas mixtures for pulmonary function testing devices

Gases used to provide "physical effects"

Hemodialysis fluids

Hemostatic devices and dressings

Injectable silicon, collagen, and Teflon

Liquids functioning through physical action applied to the body to cool or freeze tissues for therapeutic purposes

Liquids intended to inflate, flush, or moisten (lubricate) indwelling device (in or on the body)

Lubricants and lubricating jellies

Ophthalmic solutions for contact lenses

Organ/tissue transport and/or perfusion fluid with antimicrobial or other drug agent, that is, preservation solutions

Powders for lubricating surgical gloves

Sodium hyaluronate or hyaluronic acid for use as a surgical aid

Solution for use with dental "chemical drill"

Spray on dressings not containing a drug component

Status Device

Market approval authority CDRH

Special label/regulatory considerations None

Investigation options IDE

Intercenter consultation Required if the device has direct contact with the body and the drug or the chemical form of the drug has not been legally marketed as a human drug

D. Products regulated as drugs

Examples

Irrigation solutions

Purified water or saline in prefilled nebulizers for use in inhalation therapy

Skin protectants (intended for use on intact skin)

Sun screens

Topical/internal analgesic-antipyretic

Status Drug

Market approval authority CDER

Special label/regulatory considerations None

Investigation options IND

Intercenter consultations Optional

E. Ad Hoc Jurisdictional Decisions.

| Examples | Status | Center |
|--|--------|--------|
| Motility marker constructed of radiopaque plastic | Device | CDRH |
| Brachytherapy capsules, needles, etc., that are radioactive and may be removed from the body after radiation therapy has been administered | Device | CDRH |
| Skin markers | Device | CDRH |

Status Device or drug

Market approval authority CDRH or CDER as indicated

Special label/regulatory considerations None

Investigation options IDE or IND, as appropriate

Intercenter consultation Required to assure agreement on drug/device status

2.11.3.2 General Criteria Affecting Drug/Device Determination The following represent the general criteria that will apply in making device/drug determinations:

A. Device criteria

1. A liquid, powder, or other similar formulation intended only to serve as a component, part, or accessory to a device with a primary mode of action that is physical in nature will be regulated as a device by CDRH.
2. A product that has the physical attributes described in 201(h) (e.g., instrument, apparatus) of the Act and does not achieve its primary intended purpose through chemical action within or on the body, or by being metabolized, will be regulated as a device by CDRH.
3. The phrase “within or on the body” as used in 201(h) of the Act does not include extracorporeal systems or the solutions used in conjunction with such equipment. Such equipment and solutions will be regulated as devices by CDRH.
4. An implant, including an injectable material, placed in the body for primarily a structural purpose even though such an implant may be absorbed or metabolized by the body after it has achieved its primary purpose will be regulated as a device by CDRH.
5. A device containing a drug substance as a component with the primary purpose of the combination being to fulfill a device function is a combination product and will be regulated as a device by CDRH.
6. A device (e.g., machine or equipment) marketed to the user, pharmacy, or licensed practitioner that produces a drug will be regulated as a device or combination product by CDER. This does not include equipment marketed to a registered drug manufacturer.
7. A device whose labeling or promotional materials make reference to a specific drug or generic class of drugs unless it is prefilled with a drug ordinarily remains a device regulated by CDRH. It may, however, also be subject to the combination products regulation.

B. Drug criteria

1. A liquid, powder, tablet, or other similar formulation that achieves its primary intended purpose through chemical action within or on the body, or by being metabolized, unless it meets one of the specified device criteria, will be regulated as a drug by CDER.
2. A device that serves as a container for a drug or a device that is a drug delivery system attached to the drug container where the drug is present in the container is a combination product that will be regulated as a drug by CDER.
3. A device containing a drug substance as a component with the primary purpose of the combination product being to fulfill a drug purpose is a combination product and will be regulated as a drug by CDER.
4. A drug whose labeling or promotional materials makes reference to a specific device or generic class of devices ordinarily remains a drug regulated by CDER. It may, however, also be subject to the combination products regulation.

2.12 CONCLUSIONS

In summary, we have touched upon the regulations that currently control the types of preclinical toxicity testing done on potential human pharmaceuticals and medical device products. We have reviewed the history, the law, the regulations themselves, the guidelines, and common practices employed to meet regulatory standards. Types of toxicity testing were discussed, as were the special cases pertaining to, for example, biotechnology products.

REFERENCES

- Alder S, Zbinden G. (Eds.) (1988) *National and International Drug Safety Guidelines*. M.T.C. Verlag, Zollikon.
- Anderson OE. (1958) *The Health of a Nation: Harvey W. Wiley and the Fight for Pure Food*. University of Chicago Press, Chicago, IL.
- Anonymous. (1991) Points to consider in human somatic cell therapy and gene therapy. *Hum. Gene Ther.* 2(3):251–265.
- Anonymous. (1992/2015) FDA’s policy statement for the development of new stereoisomeric drugs. *Chirality*. 4(5):338–340.
- Anonymous. (1996) “Well-Characterized” biologics would be regulated under FD&C act by Kassebaum FDA reform bill. *F-D-C Rep. (Pink Sheets)*. 58:11–12.
- Anonymous (2013) *Orphan Drug Regulations*. 21 CFR Part 316. 57 FR 62085, December 29, 1992, as amended at 78 FR 35133, June 12, 2013.

- Barton BL. (1998) International conference on harmonization—good clinical practices update. *Drug Inf. J.* 32(4):1143–1147.
- Berliner VR. (1974) U.S. Food and Drug Administration requirements for toxicity testing of contraceptive products. *Acta Endocrinol. Suppl. (Copenhagen)*. 185:240–265.
- Blakeslee D. (1998) *Thalidomide: A Background Briefing*. JAMA HIV/AIDS Information Center. Centers for Disease Control (CDC), Atlanta, GA.
- Brock WJ, Hastings KL, McGown KM. (2013) *Nonclinical Safety Assessment: A Guide to International Pharmaceutical Regulations*. John Wiley & Sons, Inc., Hoboken, NJ.
- Brusick D. (1987) *Principles of Genetic Toxicology*, 2nd Ed. Plenum Press, New York.
- Burns J. (1983) Overview of safety regulations governing food, drug and cosmetics. In: *The United States in Safety and Evaluation and Regulation of Chemicals 3: Interface between Law and Science* (Homberger F, Ed.). Karger, New York.
- CDER (1989) *Guideline for the Study of Drugs Likely to be Used in the Elderly*. CDER, Rockville, MD.
- CDER (1998) *Repeal of Section 507 of the Federal Food, Drug, and Cosmetic Act*. CDER, Rockville, MD. Available at: <http://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/ucm078754.pdf> (accessed November 11, 2015).
- CDER and CBER (1995) *Content and Format of Investigational New Drug Applications (INDs) for Phase 1 Studies of Drugs, Including Well-Characterized, Therapeutic, Biotechnology-derived Products*. CDER and CBER, Rockville, MD. Available at: <http://www.fda.gov/downloads/drugs/guidancecomplianceRegulatoryInformation/guidances/ucm071597.pdf> (accessed November 16, 2015).
- CDER and CBER (2001) *Content and Format for Geriatric Labeling*. CDER and CBER, Rockville, MD. Available at: <http://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/ucm075062.pdf> (accessed November 11, 2015).
- CDER and CBER (2014) *Expedited Programs for Serious Conditions—Drugs and Biologics*. CDER and CBER, Silver Spring, MD. Available at: <http://www.fda.gov/downloads/drugs/guidancecomplianceRegulatoryInformation/guidances/ucm358301.pdf> (accessed November 16, 2015).
- Daniels JM, Nestmann ER, Kerr AI. (1997) Development of stereoisomeric (chiral) drugs: a brief review of scientific and regulatory considerations. *Drug Inf. J.* 31(3):639–646.
- De Camp WH. (1989) The FDA perspective on the development of stereoisomers. *Chirality*. 1(1):2–6.
- DeGeorge JJ, Ahn C-H, Andrews PA, Brower ME, Giorgio DW, Goheer MA, Lee-Ham DY, McGuinn WD, Schmidt W, Sun CJ, Tripathi SC. (1998) Regulatory considerations for preclinical development of anticancer drugs. *Cancer Chemother. Pharmacol.* 41(3):173–185.
- Deng R, Kaitin KI. (2004) The regulation and approval of new drugs in China. *Drug Inf. J.* 38(1):29–39.
- DiMasi JA, Hansen RW, Grabowski HG, Lasagna L. (1991) Cost of innovation in the pharmaceutical industry. *J. Health Econ.* 10(2):107–142.
- DiMasi JA, Seibring MA, Lasagna L. (1994) New drug development in the United States from 1963–1992. *Clin. Pharmacol. Ther.* 55(6):609–622.
- Ember LR. (2001) Raising the bar for quality drugs. *Chem. Eng. News*. 79(12):26.
- FDA. (1987) *Environmental Assessment Technical Assistance Handbook*. National Technical Information Service, Washington, DC.
- FDA (1988) *Investigational New Drug, Antibiotic, and Biological Drug Product Regulations; Procedures for Drugs Intended to Treat Life-Threatening and Severely Debilitating Illnesses*. CFR 21 Parts 312 and 314, Rockville, MD.
- FDA. (1993) *Points to Consider in the Characterization of Cell Lines Used to Produce Biologicals*. FDA, Washington, DC.
- FDA (1997) *Points to Consider in the Manufacture and Testing of Monoclonal Antibody Products for Human Use*. Center for Biologics Evaluation and Research, Rockville, MD. Available at: <http://www.fda.gov/downloads/BiologicsBloodVaccines/GuidanceComplianceRegulatoryInformation/OtherRecommendationsforManufacturers/UCM153182.pdf> (accessed December 4, 2015).
- FDA (2015) *Good Laboratory Practice for Nonclinical Laboratory Studies*, Silver Spring, MD. Available at: <http://www.accessdata.fda.gov/scripts/cdrh/cfdocs/cfcfr/cfrsearch.cfm?cfrpart=58> (accessed November 16, 2015).
- Fung M, Thornton A, Mybeck K, Wu JH-H, Hornbuckle K, Muniz E. (2001) Evaluation of the characteristics of safety withdrawal of prescription drugs from worldwide pharmaceutical markets—1960–1999. *Drug Inf. J.* 35(1):293–317.
- Gad SC. (2012) *Safety Pharmacology in Pharmaceutical Development: Approval and Post Marketing Surveillance*, 2nd Ed. CRC Press, Boca Raton, FL.
- Goldenthal EI. (1968) Current views on safety evaluation of drugs. *FDA Pap.* 2(4):13–18.
- Grabowski HG, Vernon JM. (1983) *The Regulation of Pharmaceuticals: Balancing the Benefits and Risks*. American Enterprise Institute for Public Policy Research, Washington, DC.
- Greishaber C. (1991) Prediction of human toxicity of new antineoplastic drugs from studies in animals. In: *The Toxicity of Anticancer Drugs* (Powis G and Hacker M, Eds.). Pergamon Press, New York, pp. 10–24.
- Haffner ME. (1998) Orphan drug development—international program and study design issues. *Ther. Innov. Regul. Sci.* 32(1):93–99.
- Hart C. (1999) Getting past the politics and paperwork of pediatric drug studies. *Mod. Drug Discovery*. 2:15–16.
- Hayes TJ, Ryffel B. (1997) Symposium-in-writing: safety considerations of recombinant protein therapy; Introductory comments. *Clin. Immunol. Immunother.* 83:1–4.
- Hirako M, McAuslane N, Salek S, Anderson C, Walker S. (2007) A comparison of the drug review process at five international regulatory agencies. *Drug Inf. J.* 41(3):291–308.
- Hite M. (1997) Safety pharmacology approaches. *Int. J. Toxicol.* 16(1):23–31.
- Holtzman DC. (1993) European biotech firms invest in US facilities. *Bioworld Today*.

- Hutt PB. (1983a) Investigations and reports respecting FDA regulations of new drugs (Part I). *Clin. Pharmacol. Ther.* 33:537–548.
- Hutt PB. (1983b) Investigations and reports respecting FDA regulations of new drugs (Part II). *Clin. Pharmacol. Ther.* 33:174–187.
- Hutt PB. (1987) Progress in new drug regulation. *Clin. Res. Pract. Drug Regul. Aff.* 5(5–6):307–318.
- ICH. (1997) *International Conference on Harmonization Safety Steps 4/5 Documents*. Interpharm Press, Buffalo Grove, IL.
- ICH (2009) *Guidance on Nonclinical Safety Studies for the Conduct of Human Clinical Trials and Marketing Authorization for Pharmaceuticals M3(R2)*. International Council for Harmonisation, Geneva, Switzerland. Available at: http://www.ich.org/fileadmin/Public_Web_Site/ICH_Products/Guidelines/Multidisciplinary/M3_R2/Step4/M3_R2__Guideline.pdf (accessed April 28, 2016).
- Jackson CO. (1970) *Food and Drug Legislation and the New Deal*. Princeton University Press, Princeton, NJ.
- Kessler DA, Siegel JP, Noguchi PD, Zoon KC, Feiden KL, Woodcock J. (1993) Regulation of somatic-cell therapy and gene therapy by the Food and Drug Administration. *N. Engl. J. Med.* 329(16):1169–1173.
- Leibert MA. (1996) Addendum to points to consider in human somatic cell and gene therapy (1991). *Hum. Gene Ther.* 7:1181–1190.
- Lumley CE, Parkinson C, Walker S. (1992) An international appraisal of the minimum duration of chronic animal toxicity studies. *Hum. Exp. Toxicol.* 11(3):155–162.
- Manson JM. (1994) Testing of pharmaceutical agents for reproductive toxicity. In: *Developmental Toxicology* (Kimmel CA and Beulke-Sam J, Eds.). Raven Press, New York, pp. 379–402.
- Marwick C. (1998) Implementing the FDA modernization act. *JAMA.* 279(11):815–816.
- Merrill RA. (1994) Regulation of drugs and devices: an evolution. *Health Aff.* 13(3):47–69.
- Ohno Y. (1999) Harmonization of the timing of non-clinical tests in relation to the conduct of clinical trials. *J. Control. Release.* 62(1–2):57–63.
- Pendergast W. (1984) Biological Drug Regulation. In: *75th Anniversary Commemorative Volume of Food and Drug Law* (Institute FaDL, Ed.). Food and Drug Law Institute, Washington, DC, pp. 293–305.
- PhRMA. (1998) *Annual Report of the Pharmaceutical Research and Manufacturer Association (Priority #2: Improved FDA Regulation of Drug Development)*. PhRMA, Washington, DC.
- Pilot LR, Waldmann DR. (1998) Food and Drug Administration Modernization Act of 1997: medical device provisions. *Food Drug Law J.* 53:267–295.
- Ryffel B. (1996) Unanticipated human toxicology of recombinant proteins. *Arch. Toxicol. Suppl.* 18:333–341.
- Shacter E, DeSantis PL. (1998) Labeling of drug and biologic products for pediatric use. *Drug Inf. J.* 32:299–303.
- Tambuyzer E. (2010) Rare diseases, orphan drugs and their regulation: questions and misconceptions. *Nat. Rev. Drug Discov.* 9(12):921–929.
- Temin P. (1980) *Taking Your Medicine: Drug Regulation in the United States*. Harvard University Press, Cambridge, MA.
- Tufts. (2014) *Cost to Develop and Win Marketing Approval of a New Drug is \$2.6 Billion Press Release*. Tufts University Center for the Study of Drug Development, Boston, MA.
- USP (2015) *United States Pharmacopeia XXXIX, National Formulary 34*. Available at: <http://www.usp.org/usp-nf> (accessed November 16, 2015).
- Wu K-M, DeGeorge JG, Atrakchi A, Barry E, Bigger A, Chen C, Du T, Freed L, Geyer H, Goheer A, Jacobs A, Jean D, Rhee H, Osterberg R, Schmidt W, Farrelly JG. (2000) Regulatory science: a special update from the United States Food and Drug Administration; preclinical issues and status of investigation of botanical drug products in the United States. *Toxicol. Lett.* 111(3):199–202.
- Yakuji Nippo. (1994) *Guidelines for Toxicity Study of Drugs Manual*. Yakuji Nippo, Tokyo.

FURTHER READING

- Beusing MA, McSweeney E. (1999) Submitting biologics applications to the Center for Biologics Evaluation and Research electronically. *Drug Inf. J.* 33(1):1–15.
- Cadden S. (1999) *New Drug Approval in Canada*. Parexel, Waltham, MA.
- CDRH (1990) *Premarket Notification, 510(k): Regulatory Requirements for Medical Devices*. CDRH, Rockville, MD. Available at: <http://babel.hathitrust.org/cgi/pt?id=mdp.39015041531420;view=1up;seq=1;size=150> (accessed March 18, 2015).
- Currie WJ. (1995) *New Drug Approval in Japan*. Parexel, Waltham, MA.
- DiMasi JA, Reichert JM, Feldman L, Malins A. (2013) Clinical approval success rates for investigational cancer drugs. *Clin. Pharmacol. Ther.* 94(3):107–142.
- Evers PT. (1995) *New Drug Approval in the European Union*. Parexel, Waltham, MA.
- FDA. (1973) *FDA Introduction to Total Drug Quality*. U.S. Government Printing Office, Washington, DC.
- FDA. (1985) *Points to Consider in the Production and Testing of New Drugs and Biologicals Produced by Recombinant DNA Technology*. Office of Biologics Research and Review, Bethesda, MD.
- FDA. (1998) *Guidance for Human Somatic Cell Therapy and Gene Therapy*. CBER, Rockville, MD.
- Gad SC. (2009) *Drug Safety Evaluation*, 2nd Ed. John Wiley & Sons, Inc., Hoboken, NJ.
- Gad SC. (2011) *Safety Evaluation of Pharmaceuticals and Medical Devices: International Regulatory Guidelines*. Springer US, New York.
- Guarino RA. (2009) *New Drug Approval Process*. CRC Press, Boca Raton, FL.

- Hess G. (1999) FDA accelerates it's approval of new drugs. *Chem. Mark. Newsl.* 255:21.
- Mantus D, Pisano DJ. (2014) *FDA Regulatory Affairs*. CRC Press, Boca Raton, FL.
- PhRMA. (2015) *2015 Biopharmaceutical Research Industry Profile*. Pharmaceutical Research and Manufacturers of America, Washington, DC.
- Ryffel B. (1997) Safety of human recombinant proteins. *Biomed. Environ. Sci.* 10(1):65–72.
- Wessinger J. (1989) Nonclinical pharmacologic and toxicologic considerations for evaluating biologic products. *Regul. Toxicol. Pharmacol.* 10(3):255–263.
- Yakuji Nippo. (1993) *Drug Registration Requirements in Japan*. Yakuji Nippo, Tokyo.
- Young F, Nightingale S, Mitchell W, Beaver L. (1989) The United States Drug Export Amendment Act of 1986: perspectives from the Food and Drug Administration. *Int. Dig. Health Legis.* 40:246–252.

DATA MINING: SOURCES OF INFORMATION FOR CONSIDERATION IN STUDY AND PROGRAM DESIGN AND IN SAFETY EVALUATION

3.1 INTRODUCTION

The appropriate starting place for the safety assessment of any new chemical entity (NCE), particularly a potential new drug, is to first determine what is already known about the molecule, its structural and therapeutics or functional class analogs (pharmacological analogs being agents with assumed similar pharmacological mechanisms), and the disease one seeks to treat. Such a determination requires the fullest possible access and review of the available literature. Here we try to provide a fundamental overview of the current range of approaches to gathering such data (Table 3.1). In using this information, one must keep in mind that there is both an initial requirement to build a data file or database and a continuing need to update such a database or files on a regular basis, serving as part of the project record. Updating a database requires not merely adding to what is already there but also discarding out-of-date (i.e., now known to be incorrect) information and reviewing the entire structure for connections and organization.

Such data is first used in selecting which possible compounds should be carried forward in development as a possible new drug (as illustrated in Figure 3.1 and explored in detail in Gad (2005)). A multitude of reasons for collecting and for uses of data should be recognized and considered.

3.1.1 Claims

Claims are what is said in labeling and advertising and may be either of a positive (therapeutic or beneficial) or negative (lack of an adverse effect) nature. The positive or efficacy claims are not usually the direct concern of the toxicologist though it must be kept in mind that such claims both must

be proven and can easily exceed the limits of the statutory definition of a device, turning the product into a drug or combination product.

Negative claims such as “nonirritating” or “hypoallergenic” also must be proven and are generally the responsibility of the product safety professional to substantiate. There are special tests for such claims.

3.1.2 Time and Economies

The final factors of influence or arbitrator of test conduct and timing are the requirements of the marketplace, the resources of the organization, and the economic worth of the product.

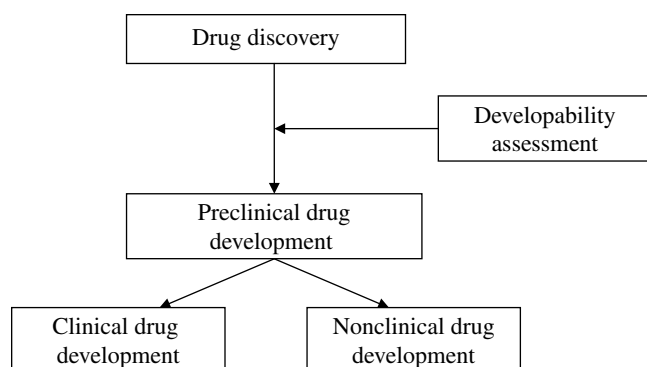
Plans for filings with regulatory agencies and for market launches are typically set before actual testing (or final stage development) is undertaken, as the need to be in the marketplace within a certain time frame is critical. Such timing and economic issues are beyond the scope of this volume but must be considered.

3.1.3 Prior Knowledge

The appropriate starting place for the safety assessment of any NCE, particularly a potential new material for a medical device, is to first determine what is already known about the material and whether there are any close structural or pharmacological analogs (pharmacological analogs being agents with assumed similar pharmacological mechanisms). Such a determination requires complete access to the available literature. In using this information, one must keep in mind that there is both an initial requirement to build a data file or database and a need to update

TABLE 3.1 Sources of Prior Art

| |
|--|
| Internet |
| FDA: Inactive ingredients for currently marketed drug products, http://www.accessdata.fda.gov/scripts/cder/iig/index.cfm |
| Proprietary databases |
| MEDLINE/TOXLINE/journals |
| Book (monographs and edited) |
| Personal network/meetings |
| Obscure databases |

**FIGURE 3.1** Prior art in assessing pharmaceutical developability.

such a store on a regular basis. Updating a database requires not merely adding to what is already there but also discarding out-of-date (i.e., now known to be incorrect) information and reviewing the entire structure for connections and organization.

The first step in any new literature review is to obtain as much of the following information as possible:

- Correct chemical identity including molecular formula, Chemical Abstracts Service (CAS) Registry Number, common synonyms, trade names, and a structural diagram. Gosselin et al. (1984), Ash and Ash (2007, 2008), and the USP (2015) (and ongoing) are excellent sources of information on existing commercial products and their components and uses. Limited by being print sources but still relevant.
- Chemical composition (if a mixture) and major impurities.
- Production and use information.
- Chemical and physical properties (physical state, vapor pressure, pH, solubility, chemical reactivity, etc.).
- Any structurally related chemical substances that are already on the market or in production.
- Known or presumed pharmacological properties.

Collection of the previous information is not only important for hazard assessment (high vapor pressure would indicate high inhalation potential, just as high and low pH would

indicate high irritation potential), but the prior identification of all intended use and exposure patterns may provide leads to alternative information sources; for example, drugs to be used as antineoplastics or antibiotics may already have extensive toxicology data obtainable from government or private sources. A great deal of the existing toxicity information (particularly information on acute toxicity) is not available in the published or electronic literature because of concerns about the proprietary nature of this information and the widespread opinion that it does not have enough intrinsic scholarly value to merit publication. This unavailability is unfortunate, as it leads to a lot of replication of effort and expenditure of resources that could be better used elsewhere. It also means that an experienced toxicologist must use an informal search of the unpublished literature and the knowledge of their colleagues as a supplement to searches of the published and electronic literature.

There are now numerous published texts that should be considered for use in literature-reviewing activities. An alphabetic listing of 36 of the more commonly used hard copy sources for safety assessment data is presented in Table 3.2 and included in the reference section of this chapter. Obviously, this is not a complete listing and consists of only the general multipurpose texts that have a wider range of applicability for toxicology. Texts dealing with specialized classes of agents (e.g., disinfectants, excipients, and pharmaceutical salts) or with specific target organ toxicity (neurotoxins and teratogens) are generally beyond the scope of this text. Parker (1988) should be consulted for details on the use of these texts. Wexler (2009), Parker (1988), and Sidhu et al. (1989) should be consulted for more extensive listings of the literature and computerized databases. Such sources can be of direct (free) Internet sources (where one must beware of garbage in, garbage out (GIGO)), commercial databases, and package products, to mention just the major categories. Appendix C provides addresses for major free Internet sources.

3.1.4 Miscellaneous Reference Sources

There are some excellent published information sources covering some specific classes of chemicals, for example, heavy metals, plastics, resins, or petroleum hydrocarbons. The National Academy of Science series *Medical and Biologic Effects of Environment Pollutants* covers 10–15 substances considered to be environmental pollutants. *Scientific American Medicine* presents a current (and consistently updated) summary of knowledge of diseases and treatments. *CRC Critical Reviews in Toxicology* is a well-known scientific journal that over the years has compiled over 26 volumes of extensive literature reviews of a wide variety of chemical substances. A photocopy of this journal's topical index will prevent one from overlooking information that may be contained in this important source. Trade organizations

TABLE 3.2 Key Safety Assessment Reference Texts

-
- Abraham DJ (Ed.) (2010) *Burger's Medicinal Chemistry and Drug Discovery*, 7th Ed. John Wiley & Sons, New York.
- American Conference of Governmental Industrial Hygienists (ACGIH) (2012) *Documentation of the Threshold Limit Values and Biological Exposure Indices*, 7th Ed. ACGIH, Cincinnati, OH.
- Ash M and Ash I. (2007) Pharmaceutical additives. In: *Electronic Handbook* 3rd Ed. Gower, Brookfield, VT.
- Ash M and Ash I. (2008) Food additives. In: *Electronic Handbook* 3rd Ed. Gower, Brookfield, VT.
- Bingham E and Cohrssen B. (Eds.) (2012) *Patty's Toxicology*, 6th Ed. John Wiley & Sons, New York.
- Brayfield A. (Ed.) (2014) *Martindale: The Complete Drug Reference*. Pharmaceutical Press, London.
- Center for Disease Control (2016) *Registry of the Toxic Effects of Chemical Substances (RTECS)*. National Institute for Occupational Safety and Health.
- Cronin E. (1980) *Contact Dermatitis*. Churchill Livingstone, Edinburgh.
- Dart RC. (2004) *Medical Toxicology*, 3rd Ed. Lippincott Williams, Wilkins, PA.
- Deichmann W and Gerard H. (1996) *Toxicology of Drugs and Chemicals*. Academic Press, New York.
- Dipiro JT, Talbert RL, Yee GC, Matzke GR, Wells BG, and Posey LM. (Eds.) (2014) *Pharmacotherapy: A Pathophysiologic Approach*, 9th Ed. McGraw-Hill, New York.
- FDA (2015) *Inactive Ingredient Database*. Available at <http://www.fda.gov/downloads/drugs/informationondrugs/UCM080154.zip> (accessed January 12, 2016).
- Ford MD. (Ed.) (2001) *Clinical Toxicology*. W.B. Saunders Company, Philadelphia.
- Gad SC and Chengelis CP. (1998) *Acute Toxicology Testing*, 2nd Ed. Academic Press, San Diego, CA.
- Gosselin RE, Smith RP, and Hodge HC. (Eds.) (1984) *Clinical Toxicology of Commercial Products*, 5th Ed. Williams & Wilkins, Baltimore, MD.
- Grant WM. (1993) *Toxicology of the Eye*, 4th Ed. Charles C. Thomas, Springfield, IL.
- Harbison RD, Bourgeois MM, and Johnson GT. (Eds.) (2015) *Hamilton and Hardy's Industrial Toxicology*, 6th Ed. John Wiley & Sons, Inc., Hoboken, NJ.
- Klaassen CD. (Ed.) (2013) *Casarett and Doull's Toxicology: The Basic Science of Poisons*, 7th Ed. McGraw-Hill, New York.
- Lewis RJ. (Ed.) (1991) *Carcinogenically Active Chemicals: A Reference Guide*. Van Nostrand Reinhold, New York.
- Lewis RJ. (2012) *Sax's Dangerous Properties of Industrial Materials*, 12th Ed. John Wiley & Sons, Inc., Hoboken, NJ.
- National Toxicology Program (NTP). (2000) *Nineteenth Annual Report on Carcinogens*. PB 85-134633. Department of Health and Human Services, Washington, DC.
- O'Neil MJ. (2013) *The Merck Index: An Encyclopedia of Chemicals, Drugs, and Biologicals*, 15th Ed. Royal Society of Chemistry, Whitehouse Station, NJ.
- PDR. (2015) *Physician's Desk Reference*, 69th Ed. PDR Network, Montvale, NJ.
- Pohanish RP. (Ed.) (2011) *Sittig's Handbook of Toxic and Hazardous Chemicals and Carcinogens*, 6th Ed. Elsevier, Oxford.
- Proctor NH and Hughes JP. (1978) *Chemical Hazards of the Workplace*. J.B. Lippincott, Philadelphia.
- Schardein JL. (2000) *Chemically Induced Birth Defects*, 3rd Ed. Marcel Dekker, New York.
- Scientific American Medicine (SciMed). (2015) Decker, New York. Updated monthly. Available at <https://www.deckerip.com/products/scientific-american-medicine/> (accessed March 16, 2016).
- Shannon MW, Borron SW, and Burns MJ. (2007) *Haddad and Winchester's Clinical Management of Poisoning and Drug Overdose*, 4th Ed. Saunders, Philadelphia.
- Shepard TH and Lemire RJ. (2010) *Catalog of Teratogenic Agents*, 13th Ed. Johns Hopkins University Press, Baltimore, MD.
- Turco S and Davis NM. (1973) Clinical significance of particular matter: a review of the literature. *Hosp. Pharm.*, 8:137–140.
- Wexler P. (Ed.) (2014) *Encyclopedia of Toxicology*, 3rd Ed. Academic Press, San Diego, CA.
- Wiley-Interscience. (Ed.) (2007) *Wiley Handbook of Current and Emerging Drug Therapies*. John Wiley & Sons, Inc., Hoboken, NJ.
-

such as the Fragrance Industry Manufacturers Association and the Chemical Manufacturers Association have extensive toxicology databases from their research programs that are readily available to toxicologists of member companies. Texts that deal with specific target organ toxicity—neurotoxicity, hepatotoxicity, or hematotoxicity—often contain detailed information on a wide range of chemical structures. Published information sources like the *Target of Organ Toxicity* series (Taylor & Francis, now partway through a third set of revisions) are examples of the types of publications that often contain important information on many industrial chemicals that may be useful either directly or by

analogy. Upon discovery that the material one is evaluating may possess target organ toxicity, a cursory review of these types of texts is warranted.

In the last decade, for most toxicologists the online literature search has changed from an occasional, sporadic activity to a semicontinuous need. Usually, in many companies, nontoxicology-related search capabilities are already in place. Therefore, all that is needed is to expand the information source to include some of the databases that cover the types of toxicology information one desires. However, if no capabilities exist within an organization (increasingly the case), one can approach a university,

consultant, or a private contract laboratory and utilize their online system at a reasonable rate. It is of course possible to access almost all of these sources from home (or home office) using a personal computer. The major available online databases are as follows:

A. *National Library of Medicine*. The National Library of Medicine (NLM) information retrieval service contains the well-known and frequently used Medical Information Online (MEDLINE), Toxicology Information Online (TOXLINE), and CANCERLIT databases. Databases commonly used by toxicologists for acute data in the NLM service are the following:

1. TOXLINE is a bibliographic database covering the pharmacological, biochemical, physiological, environmental, and toxicological effects of drugs and other chemicals. It contains approximately 1.7 million citations, most of which are complete with abstract, index terms, and CAS Registry Numbers. TOXLINE citations have publication dates of 1981 to the present. Older information is on TOXLINE 65 (pre-1965 through 1980).
2. MEDLINE is a database containing approximately 7 million references to biomedical journal articles published since 1966. These articles, usually with an English abstract, are from over 3000 journals. Coverage of previous years (back to 1966) is provided by back files, searchable online, that total some 3.5 million references.
3. Toxicology Data Network (TOXNET) is a computerized network of toxicologically oriented data banks. TOXNET offers a sophisticated search and retrieval package that accesses the following three subfiles:
 - a. Hazardous Substances Data Bank (HSDB) is a scientifically reviewed and edited data bank containing toxicological information enhanced with additional data related to the environment, emergency situations, and regulatory issues. Data are derived from a variety of sources including government documents and special reports. This database contains records for over 4100 chemical substances.
 - b. Toxicology Data Bank (TDB) is a peer-reviewed data bank focusing on toxicological and pharmacological data, environmental and occupational information, manufacturing and use data, and chemical and physical properties. References have been extracted from a selected list of standard source documents.
 - c. Chemical Carcinogenesis Research Information System (CCRIS) is a National Cancer Institute-sponsored database derived from both short- and

long-term bioassays on 2379 chemical substances. Studies cover carcinogenicity, mutagenicity, promotion, and cocarcinogenicity.

4. Registry of Toxic Effects of Chemical Substances (RTECS) is the NLM's online version of the National Institute for Occupational Safety and Health's (NIOSH) annual compilation of substances with toxic activity. The original collection of data was derived from the 1971 Toxic Substances Lists. RTECS data contains threshold limit values, aquatic toxicity ratings, air standards, National Toxicology Program carcinogenesis bioassay information, and toxicological/carcinogenic review information. The NIOSH is responsible for the file content in RTECS and for providing quarterly updates to NLM: RTECS currently covers toxicity data on more than 106 000 substances.

E. *The Merck Index*. *The Merck Index* is now available online for up-to-the-minute access to new chemical entities.

3.1.5 Search Procedure

As mentioned in Section 3.1, chemical composition and identification information should already have been obtained before the chemical is to be searched. With most information retrieval systems, this is a relatively straightforward procedure. Citations on a given subject may be retrieved by entering the desired free-text terms as they appear in titles, keywords, and abstracts of articles. The search is then initiated by entering the chemical CAS number and/or synonyms. If you are only interested in a specific target organ effect—for instance, carcinogenicity—or specific publication years, searches can be limited to a finite number of abstracts before requesting the printout.

Often it is unnecessary to request a full printout (author, title, abstract). You may choose to review just the author and title listing before selecting out the abstracts of interest. In the long run, this approach may save you computer time, especially if the number of citations being searched is large.

Once you have reviewed the abstracts, the last step is to request photocopies of the articles of interest. Extreme caution should be used in making any final health hazard determination based solely on an abstract or nonprimary literature source.

3.1.6 Monitoring Published Literature and Other Research in Progress

Although there are a few other publications offering similar services, the *Life Sciences* edition of *Current Contents* is the publication most widely used by toxicologists for monitoring the published literature. *Current Contents* monitors

over 1180 major journals and provides a weekly listing by title and author. Selecting out those journals you wish to monitor is one means of selectively monitoring the major toxicology journals.

Aids available to the toxicologist for monitoring research in progress are quite variable. Ongoing research include monitoring research in progress by reviewing abstracts presented at the annual meetings of professional societies such as the Society of Toxicology, Teratology Society, Environmental Mutagen Society, and American College of Toxicology. These societies usually have their abstracts prepared in printed form; for example, the most recent (2015) *Toxicologist* contains over 2600 abstracts presented at the annual meeting. Copies of the titles and authors of these abstracts are usually listed in the societies' respective journals, which, in many cases, would be reproduced and could be reviewed through *Current Contents*.

3.1.7 Kinds of Information

The kinds of information described here are found on two types of physical PC media—CD-ROM and laser disks. The gradual emergence of this technology during the past decade blossomed with the introduction of several CD-ROM products that deal with safety issues surrounding the toxicology and safety of chemicals. CD-ROM media with such information can generally be characterized by two major advantages: they are relatively easy to use and are amazingly quick in retrieving data of interest. The products run the gamut of allowing one to assess current developments on a weekly basis, as well as carry out more traditional reviews of historical information. The general types of information one can cover include basic pharmacology, preclinical toxicology, competitive products, and clinical safety.

The specific products discussed are as follows: CD-ROM products called Current Contents, Toxic Release Inventory (TRI), Material Safety Data Sheets (MSDS), CCINFODisc, Pollution/Toxicology, and MEDLINE Ondisc and a laser disk product entitled the Veterinary Pathology Slide Bank. We provide a brief synopsis of the major features of each as well as a description of their integration into a functional, PC-based Toxicology Information Center (TIC). It must be noted, however, that all of these are being replaced to a large extent by online resources.

When such a TIC is established, one will find that some unusual benefits accrue. One now has immediate and uninterrupted access to libraries of valuable and comprehensive scientific data. This access is free of "online" constraints and designed to be user friendly, with readily retrievable information available 24 h a day, 7 days a week. The retrieved information can also usually be manipulated in electronic form, so one can use it in reports and/or store it in machine-readable form as ASCII files.

The minimal hardware requirements, which are certainly adequate for all items discussed here, are a current Windows or Mac OS operating a single CD-ROM drive, at least a 1 GB system hard disk drive, a CD-ROM drive, a VGA color monitor, and a printer. The basic point here is that hardware requirements are minimal and readily available. In the case of the laser disk products, a laser disk drive and high-resolution (VGA) monitor are also required.

3.1.8 Toxic Release Inventory (TRI)

Before embarking on a discussion of products describing health, toxicology, and safety issues, it is well to be aware of a new, pilot CD-ROM version of the Environmental Protection Agency's (EPA) 1987 Toxic Chemical Release Inventory and Hazardous Substances Fact Sheets. This TRI resource, which contains information regarding the annual inventory of hundreds of named toxic chemicals from certain facilities (since 1987), as well as the toxicological and ecological effects of chemicals, is available from the National Technical Information Service (NTIS), US Department of Commerce, Springfield, Virginia, 22161.

The list of toxic chemicals subject to reporting was originally derived from those designed for similar purposes by the states of Maryland and New Jersey. As such, over 300 chemicals and categories are noted. (After appropriate rule making, modifications to the list can be made by the EPA.) The inventory is designed to inform the public and government officials about routine and accidental releases of toxic chemicals to the environment.

The CD-ROM version of the database can be efficiently searched with a menu-driven type of software called SearchExpress. It allows one to search with Boolean expressions as well as individual words and/or frequency of "hits" as a function of the number of documents retrieved on a given topic. Numerous searchable fields have been included, allowing one to retrieve information by a variety of means—for example, the compound name; the chemical registry number; the amount of material released into the air, water, or land; the location of the site of release; and the SIC code of the releasing party. One can also employ ranging methods with available numeric fields and sorting of output.

It is hoped that this shared information will help to increase the awareness, concern, and action by individuals to ensure a clean and safe environment. The TRI database is a significant contribution to that effort, and the CD-ROM version is a superb medium with which to widely publicize and make accessible the findings.

3.1.9 Material Safety Data Sheets (MSDS)

The MSDS CD-ROM is a useful resource that contains over 33000 MSDS on chemicals submitted to the Occupational Safety and Health Administration (OSHA) by chemical

manufacturers. This resource contains complete MSDS information as well as other important information such as the chemical formula, structure, physical properties, synonyms, registry number, and safety information.

Users can easily search the CD-ROM by employing the Aldrich catalog number, CAS number, chemical name, or molecular formula. One can also export the chemical structures to some supported software for subsequent inclusion into work processing programs. The product is available from Aldrich Chemical Company, Inc., 940 West Street, Paul Avenue, Milwaukee, Wisconsin, 54233.

3.1.10 Canadian Centre for Occupational Health and Safety (CCINFO)

This set of four CD-ROM disks contains several valuable databases of information that are updated on a quarterly basis: MSDS, CHEM Data, Occupational Health and Safety (OHS) Source, and OHS Data. The MSDS component currently contains over 60000 MSDS supplied by chemical manufacturers and distributors. It also contains several other databases (RIPP, RIPA, Pest Management Research Information System (PRIS)), one of which (PRIS) even includes information on pest management products, including their presence and allowable limits in food.

A second disk in the series (CHEM Data) contains comprehensive information from the CHEMINFO, RTECS, and Chemical Evaluation Search and Retrieval System (CESARS) databases, as well as recommendations on Transport of Dangerous Goods (TDG)/Hazardous Materials (49CFR).

The third and fourth disks include OHS information. These disks contain databases on resource organizations, resource people, case law, jurisprudence, fatalities, mining incidents, and ADISCAN. Furthermore, information on noise levels, NIOSH nonionizing radiation levels, and a document information directory system is readily retrievable. These CD-ROM materials are available from the Canadian Centre for Occupational Health and Safety, 250 Main Street East, Hamilton, Ontario, L8N 1H6.

3.1.11 Pollution and Toxicology (POLTOX)

This CD-ROM library also focuses our attention on environmental health and safety concerns. Scientists working in any industry or capacity that deals with toxic or potentially toxic chemicals will find it very useful. It allows one access to seven major databases in this field in a single search through its use of "linking" features in its software. The distributors of this product have provided us with a spectrum of information dealing with toxic substances and environmental health.

The collection of these databases include five that are available exclusively from Cambridge Scientific Abstracts

(CSA)—Pollution Abstracts, Toxicology Abstracts, Ecology Abstracts, Health and Safety Science Abstracts, and Aquatic Pollution and Environmental Quality. The abstracts come from journals or digests published by CSA on important issues including environmental pollution, toxicological studies of industrial chemicals, ecological impacts of biologically active chemicals, as well as health, safety, and risk management in occupational situations. The POLTOX CD-ROM contains over 200000 records from these sources since 1981.

POLTOX also contains two other useful databases—TOXLINE (described earlier) and the Food Science and Technology Abstracts (FSTA) libraries. The FSTA component is a reasonably comprehensive collection of information regarding toxicological aspects of compounds found in food, including contamination, poison, and carcinogenic properties. The CD-ROM product is available from Compact Cambridge, 7200 Wisconsin Avenue, Bethesda, Maryland, 20814.

3.1.12 MEDLINE

The MEDLINE database, which comes from the NLM, is a superb, indispensable reference library that is particularly strong in its wide coverage of research activities in the biomedical literature. It also encompasses the areas of clinical medicine, health policy, and healthcare services. Each year, over 300000 articles are reviewed and indexed into the database. The full bibliographic citations of these articles, usually including the abstract of the published work, are available from numerous vendors in CE-ROM form and are usually updated on a monthly basis.

Information can be accessed from MEDLINE in a variety of ways: by author, title, subject, CAS Registration Number, keyword, publication year, and journal title. MEDLINE Ondisc is the CD-ROM product we employ (from Dialog Information Services, Inc., 3460 Hillview Avenue, Palo Alto, California, 94304). It allows one access to the full MEDLINE files back to 1984. Each year from that time until 1988 is covered on a single CD-ROM disk; starting in 1989, each disk covers only a 6-month time period. The information is accessed through either an easily employed "menu-driven" system or a more standard online type of "command language."

Gower Publishing (Brookfield, VT) has published a series of "electronic handbooks" providing approved ingredient information on materials used in cosmetics, personal care additives, food additives, and pharmaceuticals. Academic Press, through its Sci-Vision branch, has just (2000) launched an ambitious service of CD-ROM-based toxicity database products which are structure and substructure searchable.

It is worth noting that the CD-ROM-based system has been seamlessly integrated with (proprietary) both record-keeping and communications software so that one can

optionally monitor the use of the online services and easily continue searching in the dialog “online” environment after using the CD-ROM-based MEDLINE library. Another very useful feature includes the storage of one’s search logic so that repetitive types of searches, over time, for example, can be done very easily.

3.2 PC-BASED INFORMATION PRODUCTS: LASER DISC

3.2.1 International Veterinary Pathology Slide Bank (IVPSB)

This application represents an important complementary approach toward training and awareness using laser disk technology. The International Veterinary Pathology Slide Bank (IVPSB) provides a quality collection of transparencies, laser videodisks, and interactive computer/videodisk training programs. In particular, the videodisk contains over 21 000 slides from over 60 contributors representing 37 institutions from 6 countries. These slides are accessible almost instantaneously because of the tremendous storage capacity and rapid random search capabilities of the videodisk through the interactive flexibility of the computer. The information available is of particular interest to toxicologists and pathologists because the visuals illustrate examples of gross lesions of infectious diseases, regional diseases, clinical signs or external microscopy, histopathology, normal histology, cytology and hematology, and parasitology.

The laser disk, a catalog of the entries, a computer database, and selected interactive programs can be obtained from Dr. Wayne Crowell, Department of Veterinary Pathology, University of Georgia, Athens, Georgia, 30602.

3.3 CONCLUSIONS

This brief overview of some of the readily available PC-based information resources will, hopefully, encourage more widespread use of this type of technology. Toxicologists and pathologists, in particular, can avail themselves of these useful resources in a way that was simply not possible just a few years ago. The information one needs to make decisions is now far more accessible to many more of us for relatively reasonable expenditures of money for software and hardware.

An effective approach to provide maximal access to these resources is to set up a “TIC,” which consists of the earlier noted PC hardware and single, centrally available copies of the noted CD-ROM-based and laser disk products. By employing a menu-based system (available commercially or by shareware) to access the respective products, one can usually provide entry into each of the products discussed here with a single keystroke.

As time goes on, one can grow with the system by considering networking the CD-ROM-based resources and/or setting up other strategically located TICs on one’s campus. The important concept here is that we wish to make the superb “new” PC-based information products as available as we can to interested scientists.

A critical part of the strategy for delivery of information to the end user is that one can anticipate marked increased usage of the more traditional hard copy-based resources of the centralized library. The tools described here are frequently complementary to the pivotal library-based information center. What one can anticipate, however, is a much more focused use of hard copy-based information.

REFERENCES

- Abraham DJ. (Ed.) (2010) *Burger’s Medicinal Chemistry and Drug Discovery*, 7th Ed. John Wiley & Sons, New York.
- American Conference of Governmental Industrial Hygienists (ACGIH). (Ed.) (2012) *Documentation of the Threshold Limit Values for Substances in Workroom Air*. ACGIH, Cincinnati, OH.
- Ash M, Ash I. (2007) Pharmaceutical additives. In: *Electronic Handbook*, 3rd Ed. Gower, Brookfield, VT.
- Ash M, Ash I. (2008) Food additives. In: *Electronic Handbook*, 3rd Ed. Gower, Brookfield, VT.
- Bingham E, Cohn B. (Eds.) (2012) *Patty’s Toxicology*, 6th Ed. John Wiley & Sons, New York.
- Brayfield A. (Ed.) (2014) *Martindale: The Complete Drug Reference*. Pharmaceutical Press, London.
- Center for Disease Control. (2016) *Registry of the Toxic Effects of Chemical Substances (RTECS)*. National Institute for Occupational Safety and Health, Washington, DC.
- Cronin E. (1980) *Contact Dermatitis*. Churchill Livingstone, Edinburgh.
- Dart RC. (2004) *Medical Toxicology*, 3rd Ed. Lippincott Williams, Wilkins, PA.
- Deichmann W, Gerard H. (1996) *Toxicology of Drugs and Chemicals*. Academic Press, New York.
- Dipiro JT, Talbert RL, Yee GC, Matzke GR, Wells BG, Posey LM. (Eds.) (2014) *Pharmacotherapy: A Pathophysiologic Approach*, 9th Ed. McGraw-Hill, New York.
- FDA (2015) *Inactive Ingredient Database*. Available at <http://www.fda.gov/downloads/drugs/informationondrugs/UCM080154.zip> (accessed January 12, 2016).
- Ford MD. (Ed.) (2001) *Clinical Toxicology*. W.B. Saunders Company, Philadelphia.
- Gad SC. (Ed.) (2005) *Drug Discovery Handbook*. John Wiley & Sons, Inc., Hoboken, NJ.
- Gad SC, Chengelis CP. (1998) *Acute Toxicology Testing*, 2nd Ed. Academic Press, San Diego, CA.
- Gosselin RE, Smith RP, Hodge HC. (Eds.) (1984) *Clinical Toxicology of Commercial Products*, 5th Ed. Williams & Wilkins, Baltimore, MD.

- Grant WM. (1993) *Toxicology of the Eye*, 4th Ed. Charles C. Thomas, Springfield, IL.
- Harbison RD, Bourgeois MM, Johnson GT. (Eds.) (2015) *Hamilton and Hardy's Industrial Toxicology*, 6th Ed. John Wiley & Sons, Inc., Hoboken, NJ.
- Klaassen CD. (Ed.) (2013) *Casarett and Doull's Toxicology: The Basic Science of Poisons*, 7th Ed. McGraw-Hill, New York.
- Lewis RJ. (Ed.) (1991) *Carcinogenically Active Chemicals: A Reference Guide*. Van Nostrand Reinhold, New York.
- Lewis RJ. (2012) *Sax's Dangerous Properties of Industrial Materials*, 12th Ed. John Wiley & Sons, Inc., Hoboken, NJ.
- National Toxicology Program (NTP) (2000) *Nineteenth Annual Report on Carcinogens*. PB 85-134633. Department of Health and Human Services, Washington, DC.
- O'Neil MJ. (2013) *The Merck Index: An Encyclopedia of Chemicals, Drugs, and Biologicals*, 15th Ed. Royal Society of Chemistry, Whitehouse Station, NJ.
- Parker CM. (1988) Available toxicology information sources and their use. In: *Product Safety Evaluation Handbook* (Gad SC, Ed.). Marcel Dekker, New York, pp. 23–41.
- PDR. (2015) *Physician's Desk Reference*, 69th Ed. PDR Network, Montvale, NJ.
- Pohanish RP. (Ed.) (2011) *Sittig's Handbook of Toxic and Hazardous Chemicals and Carcinogens*, 6th Ed. Elsevier, Oxford.
- Proctor NH, Hughes JP. (1978) *Chemical Hazards of the Workplace*. J.B. Lippincott, Philadelphia.
- Schardein JL. (2000) *Chemically Induced Birth Defects*, 3rd Ed. Marcel Dekker, New York.
- Scientific American Medicine (SciMed) (2015) Decker, New York. Updated monthly. Available at <https://www.deckerip.com/products/scientific-american-medicine/> (accessed March 16, 2016).
- Shannon MW, Borron SW, Burns MJ. (2007) *Haddad and Winchester's Clinical Management of Poisoning and Drug Overdose*, 4th Ed. Saunders, Philadelphia.
- Shepard TH, Lemire RJ. (2010) *Catalog of Teratogenic Agents*, 13th Ed. Johns Hopkins University Press, Baltimore, MD.
- Sidhu KS, Stewart TM, Netton EW. (1989) Information sources and support networks in toxicology. *J. Am. Coll. Toxicol.* 8:1011–1026.
- Turco S, Davis NM. (1973) Clinical significance of particular matter: a review of the literature. *Hosp. Pharm.* 8:137–140.
- United States Pharmacopeia XXXIX, National Formulary 34. (2015) Available at <http://www.usp.org/usp-nf> (accessed November 16, 2015).
- Wexler P. (Ed.) (2009) *Information Resources in Toxicology*, 4th Ed. Elsevier, New York.
- Wexler P. (Ed.) (2014) *Encyclopedia of Toxicology*, 3rd Ed. Academic Press, San Diego, CA.
- Wiley-Interscience. (2007) *Wiley Handbook of Current and Emerging Drug Therapies*. John Wiley & Sons, Inc., Hoboken, NJ.

SCREENS IN SAFETY AND HAZARD ASSESSMENT

4.1 INTRODUCTION

In biological research, screens are tests designed and performed to identify agents or organisms having a certain set of characteristics that will either exclude them from further consideration or cause them to be selected for closer attention. In pharmaceutical safety assessment, our use of screens is usually negative (i.e., no activity is found)—agents or objects possessing certain biochemical activities are considered to present enough of a hazard that they are not studied further or developed as potential therapeutic agents without compelling reasons (in cases of extreme benefit such as life-saving qualities).

In the broadest terms, what is done in preclinical (and, indeed, in phase I clinical) studies can be considered a form of screening (Zbinden et al., 1984). Starting with selection of candidates (Gad, 2005) and the development of formulations (Boersen et al., 2013), screens are employed throughout the evaluation and development process. What varies is the degree of effectiveness of (or our confidence in) each of the tests used. As a general rule, though we think of the expensive and labor intensive “pivotal” studies required to support regulatory requirements (4-week–1-year toxicity studies, carcinogenicity, and segment I–III studies, etc.) as definitive, in fact, they are highly effective (or so, at least, we generally believe) but not necessarily efficient screens. An initial screen, now frequently required by FDA, is of binding of candidate therapeutic molecules to receptors in a defined panel (such as presented in Table 4.1), as is discussed in Bowes et al. (2012). The purpose is to determine the specificity of the candidate for its target, the potential risk arising from off-target hits.

Though toxicologists in the pharmaceutical industry are familiar with the broad concepts of screening, they generally do not recognize the applicability of screens. The principles underlying screening are also not generally well recognized or understood. The objective behind the entire safety

assessment process in the pharmaceutical industry is to identify those compounds for which the risk of harming humans does not exceed the potential benefit to them.

In most cases this means that if a test or screen identifies a level of risk that we have confidence in (our “activity criterion”), then the compound that was tested is no longer considered a viable candidate for development. In this approach, what may change from test to test is the activity criterion (i.e., our basis for and degree of confidence in the outcome). We are interested in minimizing the number of false negatives in safety assessment. Anderson and Hauck (1983) should be consulted for statistical methods to minimize false-negative results.

Figure 4.1 illustrates how currently decisions are more likely to be made on a multidimensional basis, which creates a need for balance between (1) degree of benefit, (2) confidence that there *is* a benefit (efficacy is being evaluated in “models” or screens at the same time safety is), (3) type of risk (with, e.g., muscle irritation, mutagenicity, acute lethality, and carcinogenicity having various degrees of concern attached to them), and (4) confidence in, and degree of, risk. This necessity for balance is commonly missed by many who voice opposition to screens because “they may cause us to throw out a promising compound based on a finding in which we have only (for example) 80% confidence.” Screens, particularly those performed early in the research and development process, should be viewed as the biological equivalent of exploratory data analysis (EDA). They should be very sensitive, which by definition means that they will have a lot of “noise” associated with them. Screens generally do not establish that an agent is (or is not) a “bad actor” for a certain end point. Rather, they confirm that if interest in a compound is sufficient, a more definitive test (a confirmatory test) is required, which frequently will provide a basis for selecting between multiple candidate compounds.

TABLE 4.1 Pharmacologic Receptor Screen Panel*Neurotransmitter-Related Ion Channels*

Adenosine
 Adrenergic, alpha 1
 Adrenergic, alpha 2
 Calcium channel
 Adrenergic, beta
 Dopamine transporter
 Potassium channel
 Dopamine
 GABA A, agonist site
 GABA A, BDX, alpha 1 site 5.62% **sodium, site 2 96.49%**
 GABA B 1.30%
 Glutamate, AMPA site (ionotropic) 11.65%
 Growth factors/hormones
 Glutamate, kainate site (ionotropic)
 Glutamate, NMDA agonist site (ionotropic) nonselective
 Glutamate, NMDA, glycine
 (Stry-insens site) (ionotropic)
 Platelet-activating factor, PAF
 Glycine, strychnine sensitive
 Histamine, H1
 Histamine, H2
 Histamine, H3
 Melatonin, nonselective
 Muscarinic, M1 (h)
 Muscarinic, M2 (h)
 Muscarinic, nonselective, central
 Muscarinic, nonselective, peripheral
 Nicotinic, neuronal (a-BnTx insensitive)
 Norepinephrine transporter
 Opioid, nonselective
 Opioid, orphanin, ORL1 (h)
 Serotonin transporter
 Serotonin, nonselective
 Sigma, nonselective

Nonselective

Steroids, vasopressin 1
 Estrogen
 Testosterone (cytosolic) (h)
 Decarboxylase, glutamic acid
 Second messenger esterase
 Oxidase, MAO-A, peripheral
 Nitric oxide, NOS (neuronal binding)
 Transferase, choline acetyl
 Prostaglandins
 Leukotriene, LTB4 (BLT)

Normal practice is to screen these 62 receptors at 10 μ M for binding to evaluate selectivity.

- Sensitivity: the ratio of true positives to total actives
- Specificity: the ratio of true negatives to total inactives
- Positive accuracy: the ratio of true to observed positives
- Negative accuracy: the ratio of true to observed negatives
- Capacity: the number of compounds that can be evaluated
- Reproducibility: the probability that a screen will produce the same results at another time (and, perhaps, in some other lab)

These characteristics may be optimized for a particular use, if we also consider the mathematics underlying them and “errors.”

A brief review of the basic relationships between error types and power starts with considering each of five interacting factors (Gad, 1982a, 1999) that serve to determine power and define competing error rates:

α , the probability of us committing a type I error (a false positive)

β , the probability of us committing a type II error (a false negative)

Δ , the desired sensitivity in a screen (such as being able to detect an increase of 10% in mutations in a population)

σ , the variability of the biological system and the effects of chance errors

n , the necessary sample size needed to achieve the desired levels of each of these factors

We can, by our actions, generally change only this portion of the equation, since n is proportional to

$$\frac{\sigma}{\alpha, \beta, \text{ and } \Delta}$$

The implications of this are, therefore, that (1) the greater σ is, the larger n must be to achieve the desired levels of α , β , and/or Δ and (2) the smaller the desired levels of α , β , and/or Δ , if n is constant, the larger σ must be.

What are the background response level and the variability in our technique? As any good toxicologist will acknowledge, matched concurrent control (or standardization) groups are essential to minimize within-group variability as an “error” contributor. Unfortunately, in *in vivo* toxicology test systems, large sample sizes are not readily attainable, and there are other complications to this problem that we shall consider later.

In an early screen, a relatively large number of compounds will be tested. It is unlikely that one will stand out so

4.2 CHARACTERISTICS OF SCREENS

The terminology involved in screen design and evaluation and the characteristics of a screen should be clearly stated and understood. The characteristics of screen performance are defined as:

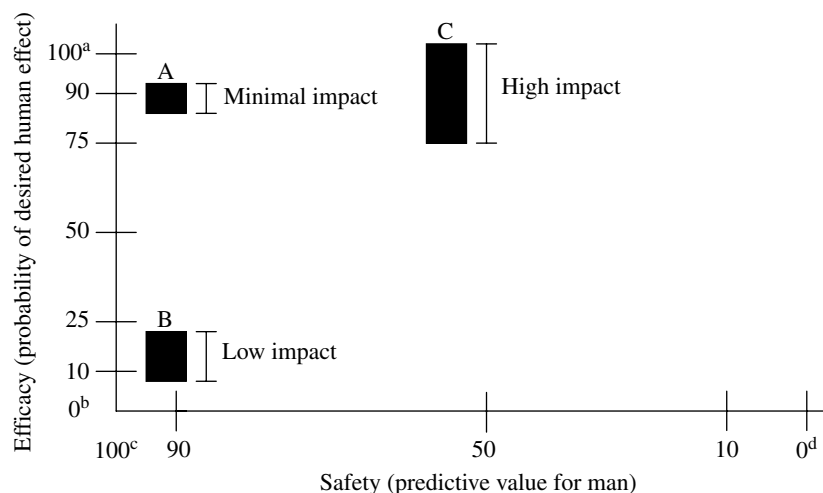


FIGURE 4.1 Decision making for pharmaceutical candidates based on outcome of screening tests. (a) A 100% probability of efficacy means that every compound that has the observed performance in the model(s) used has the desired activity in man. (b) A 0% probability of efficacy means that every compound that has the observed performance in the model(s) used does not have the desired activity in man. (c) A 100% probability of a safety finding means that such a compound would definitely cause this toxicity in man. (d) A 0% probability means this will never cause such a problem in man. Note: these four cases (a, b, c, and d) are almost never found. The height of the “impact” column refers to the relative importance (“human risk”) of a safety finding. Compound A has a high probability of efficacy but also a high probability of having some adverse effect in man. But if that adverse effect is of low impact—say, transitory muscle irritation for a life-saving antibiotic—A should go forward. Likewise, B, which has a low probability of efficacy and a high probability of having an adverse effect with moderate impact, should not be pursued. Compound C is at a place where the high end of the impact scale should be considered. Though there is only a 5% probability of this finding (say, neurotoxicity or carcinogenicity) being predictive in man, the adverse effect is not an acceptable one. Here a more definitive test is called for or the compound should be dropped.

TABLE 4.2 Discovery and Discrimination of Toxicants

| Screen Outcome | Actual Activity of Agent Tested | |
|----------------|---------------------------------|----------|
| | Positive | Negative |
| Positive | <i>a</i> | <i>b</i> |
| Negative | <i>c</i> | <i>d</i> |

Discovery (sensitivity) = $a/(a+c)$, where a = all toxicants found positive and $a+c$ = all toxicants tested.

Discrimination (specificity) = $d/(b+d)$, where d = all nontoxicants found negative and $b+d$ = all nontoxicants tested.

much as to have greater statistical significance than all the other compounds (Bergman and Gittins, 1985). A more or less continuous range of activities will be found instead. Compounds showing the highest (beneficial) or lowest (adverse) activity will proceed to the next assay or tier of tests in the series and may be used as lead compounds in a new cycle of testing and evaluation.

The balance between how well a screen discovers activities of interest versus other effects (specificity) is thus critical. Table 4.2 presents a graphic illustration of the dynamic relationship between discovery and discrimination.

Both discovery and discrimination in screens hinge on the decision criterion that is used to determine if activity has or has not been detected. How sharply such a criterion

is defined and how well it reflects the working of a screening system are two of the critical factors driving screen design.

An advantage of testing many compounds is that it gives the opportunity to average activity evidence over structural classes or to study quantitative structure–activity relationships (QSARs). QSARs can be used to predict the activity of new compounds and thus reduce the chance of *in vivo* testing on negative compounds. The use of QSARs can increase the proportion of truly active compounds passing through the system.

It should be remembered that maximization of the performance of a series of screening assays required close collaboration among the toxicologist, chemist, and statistician. Screening, however, forms only part of a much larger research and development context. Screens thus may be considered the biological equivalent of EDA. EDA methods, in fact, provide a number of useful possibilities for less rigid and yet utilitarian approaches to the statistical analysis of the data from screens and are one of the alternative approaches presented and evaluated here (Tukey, 1977; Redman, 1981; Hoaglin et al., 1983; Hoaglin et al., 1985). Over the years, the author has published and consulted on a large number of screening studies and projects. These have usually been directed at detecting or identifying potential behavioral toxicants or neurotoxicants, but some have been

directed at pharmacological, immunotoxic, and genotoxic agents (Gad, 1988, 1989a).

The general principles or considerations for screening in safety assessments are as follows:

1. Screens almost always focus on detecting a single point of effect (such as mutagenicity, lethality, neurotoxicity, or developmental toxicity) and have a particular set of operating characteristics in common.
2. A large number of compounds are evaluated, so ease and speed of performance (which may also be considered efficiency) are very desirable characteristics.
3. The screen must be very sensitive in its detection of potential effective agents. An absolute minimum of active agents should escape detection; that is, there should be very few false negatives (in other words, the type II error rate or beta level should be low). Stated yet another way, the signal gain should be way up.
4. It is desirable that the number of false positives be small (i.e., there should be a low type I error rate or alpha level).
5. Items 2–4, which are all to some degree contradictory, require the involved researchers to agree on a set of compromises, starting with the acceptance of a relatively high alpha level (0.10 or more), that is, a higher noise level.
6. In an effort to better serve Item 1, safety assessment screens frequently are performed in batteries so that multiple end points are measured in the same operation. Additionally, such measurements may be repeated over a period of time in each model as a means of supporting Item 2.
7. The screen should use small amounts of compound to make Item 1 possible and should allow evaluation of materials that have limited availability (such as novel compounds) early on in development.
8. Any screening system should be validated initially using a set of blind (positive and negative) controls. These blind controls should also be evaluated in the screening system on a regular basis to ensure continuing proper operation of the screen. As such, the analysis techniques used here can then be used to ensure the quality or modify performance of a screening system.
9. The more that is known about the activity of interest, the more specific the form of screen that can be employed. As specificity increases, so should sensitivity. However, generally the size of what constitutes a meaningful change (that is, the Δ) must be estimated and is rarely truly known.
10. Sample (group) sizes are generally small.
11. The data tend to be imprecisely gathered (often because researchers are unsure what they are looking for) and therefore possess extreme within-group variability or modify test performance.
12. Proper dose selection is essential for effective and efficient screen design and conduct. If insufficient data are available, a suitably broad range of doses must be evaluated (however, this technique is undesirable on multiple grounds, as has already been pointed out).

Much of the mathematics involved in calculating screen characteristics came from World War II military-based operations analysis and research, where it was important for design of radar, antiair, and antisubmarine warfare systems and operations (Garrett and London, 1970).

4.3 USES OF SCREENS

The use of screens that first occurs to most pharmaceutical scientists is in pharmacology (Martin et al., 1988). Early experiences with the biological effects of a new molecule are almost always in some form of efficacy or pharmacology screen. The earliest of these tend to be with narrowly focused models, not infrequently performed *in vitro*. The later pharmacology screens, performed *in vivo* to increase confidence in the therapeutic potential of a new agent or to characterize its other activities (cardiovascular, CNS, etc.), can frequently provide some information of use in safety assessment also (even if only to narrow the limits of doses to be evaluated), and the results of these screens should be considered in early planning. In the new millennium, requirements for specific safety pharmacology screens have been promulgated. Additionally, since the late 1990s, two new areas of screening have become very important in pharmaceutical safety assessment. The first is the use of screens for detecting compounds with the potential to cause fatal cardiac arrhythmias. These are almost always preceded by the early induction of a prolongation of the Q–T interval. While this should be detected in the EKGs performed in repeat-dose canine studies, several early screens (such as the HERG) are more rapid and efficient (though not conclusive) for selecting candidate compounds for further development.

The other area is the use of microassays in toxicogenomic screening—early detection of the potential for compounds to alter gene expressions with adverse consequences (Nuwaysir et al., 1999; Pennie, 2000).

Safety assessment screens are performed in three major settings—discovery support, development (what is generally considered the “real job” of safety assessment), and occupational health/environmental assessment testing.

Discovery support is the most natural area of employment of screens and is the place where effective and efficient screen design and conduct can pay the greatest long-range benefits. If compounds with unacceptable safety profiles can be identified before substantial resources are invested in them—and structures modified to maintain efficacy while avoiding early safety concerns—then long-term success of the entire research and development effort is enhanced. In the discovery support phase, one has the greatest flexibility in the design and use of screens. Here screens truly are used to select from among a number of compounds.

Examples of the use of screens in the development stage are presented in some detail in the next section.

The use of screens in environmental assessment and occupational health is fairly straightforward. On the occupational side, concerns (as addressed in Chapter 25 of this volume) address the potential hazards to those involved in making the bulk drug. The need to address potential environmental concerns covers both true environmental items (aquatic toxicity, photostability, environmental, accumulation, etc.) and potential health concerns for environmental exposures of individuals. The resulting work tends to be either regulatorily defined tests (for aquatic toxicity) or defined end points such as dermal irritation and sensitization, which have been (in a sense) screened for already in other nonspecific tests.

The most readily recognized examples of screens in toxicology are those that focus on a single end point. The traditional members of this group include genotoxicity tests, lethality tests (particularly recognizable as a screen when in the form of limit tests), and tests for corrosion, irritation (both eye and skin), and skin sensitization. Others that fit this same pattern, as will be shown, include the carcinogenicity bioassay (especially the transgenic mouse models) and developmental toxicity studies.

The “chronic” rodent carcinogenicity bioassay is thought of as the “gold standard” or definitive study for carcinogenicity, but, in fact, it was originally designed as (and functions as) a screen for strong carcinogens (Page, 1977). It uses high doses to increase its sensitivity in detecting an effect in a small sample of animals. The model system (be it rats or mice) has significant background problems of interpretation. As with most screens, the design has been optimized (by using inbred animals, high doses, etc.) to detect one type of toxicant—strong carcinogens. Indeed, a negative finding does not mean that a material is not a carcinogen but rather than it is unlikely to be a potent one.

Many of the studies done in safety assessment are multiple end-point screens. Such study types as a 90-day toxicity study or immunotox/neurotox screens are designed to measure multiple end points with the desire of increasing both sensitivity and reliability (by correspondence/correlation checks between multiple data sets).

4.4 TYPES OF SCREENS

There are three major types of screen designs: the single stage, sequential, and tiered. Both the sequential and tiered are multistage approaches, and each of these types also varies in terms of how many parameters are measured. But these three major types can be considered as having the following characteristics.

4.4.1 Single Stage

A single test will be used to determine acceptance or rejection of a test material. Once an activity criterion (such as X score in a righting reflex test) is established, compounds are evaluated based on being less than X (i.e., negative) or equal to or greater than X (i.e., positive). As more data are accumulated, the criterion should be reassessed.

4.4.2 Sequential

Two or more repetitions of the same test are performed, one after the other, with the severity of the criterion for activity being increased in each sequential stage. This procedure permits classification of compounds into a set of various ranges of potencies. As a general rule, it appears that a two-stage procedure, by optimizing decision rules and rescreening compounds before declaring compounds “interesting,” increases both sensitivity and positive accuracy; however, efficiency is decreased (or is throughput rate).

4.4.3 Tier (or Multistage)

In this procedure, materials found active in a screen are reevaluated in one or more additional screens or tests that have greater discrimination. Each subsequent screen or test is both more definitive and more expensive.

For purposes of our discussion here, we will primarily focus on the single-stage system, which is the simplest. The approaches presented here are appropriate for use in any of these screening systems, although establishment of activity criteria becomes more complicated in successive screens. Clearly, the use of multistage screens presents an opportunity to obtain increased benefits from the use of earlier (lower-order) screening data to modify subsequent screen performance and the activity criterion.

4.5 CRITERION: DEVELOPMENT AND USE

In any early screen, a relatively large number of compounds will be evaluated with the expectation that a minority will be active. It is unlikely that any one will stand out so much as to have greater statistical significance than all the other compounds based on a formal statistical test. A more or less

continuous range of activities will be found. Compounds displaying a certain degree of activity will be identified as “active” and handled as such. For safety screens, those which are “inactive” go on to the next test in a series and may be used as lead compounds in a new cycle of testing and evaluation. The single most critical part of the use of screens is how to make the decision that activity has been found.

Each test or assay has an associated activity criterion. If the result for a particular test compound meets this criterion, the compound is “active” and handled accordingly. Such a criterion could have a statistical basis (e.g., all compounds with observed activities significantly greater than the control at the 5% level could be tagged). However, for early screens, a statistical criterion may be too strict, given the power of the assay, resulting in a few compounds being identified as “active.” In fact, a criterion should be established (and perhaps modified over time) to provide a desired degree of confidence in the predictive value of the screen.

A useful indicator of the efficiency of an assay series is the frequency of discovery of truly active compounds. This is related to the probability of discovery and to the degree of risk (hazard to health) associated with an active compound passing a screen undetected. These two factors, in turn, depend on the distribution of activities in the series of compounds being tested and the chances of rejecting and accepting compounds with given activities at each stage.

Statistical modeling of the assay system may lead to the improvement of the design of the system by reducing the interval between discoveries of active compounds. The objectives behind a screen and considerations of (1) costs for producing compounds and testing and (2) the degree of uncertainty about test performance will determine the desired performance characteristics of specific cases. In the most common case of early toxicity screens performed to remove possible problem compounds, preliminary results suggest that it may be beneficial to increase the number of

compounds tested, decrease the numbers of animals (or other test models) per assay, and increase the range and number of doses. The result will be less information on more structures, but there will be an overall increase in the frequency of discovery of active compounds (assuming that truly active compounds are entering the system at a random and steady rate).

The methods described here are well suited to analyzing screening data when the interest is truly in detecting the absence of an effect with little chance of false negatives. There are many forms of graphical analysis methods available, including some newer forms that are particularly well suited to multivariate data (the type that is common in more complicated screening test designs). It is intended that these aspects of analysis will be focused on in a later publication.

The design of each assay and the choice of the activity criterion should, therefore, be adjusted, bearing in mind the relative costs of retaining false positives and rejecting false negative (Bickis, 1990). Decreasing the group sizes in the early assays reduced the chance of obtaining significance at any particular level (such as 5%) so that the activity criterion must be relaxed, in a statistical sense, to allow more compounds through. At some stage, however, it becomes too expensive to continue screening many false positives, and the criteria must be tightened accordingly. Where the criteria are set depends on what acceptable noise levels are in a screening system.

Criteria can be simple (lethality) or as complex (a number of clinical chemical and hematologic parameters) as required. The first step in establishing them should be an evaluation of the performance of test systems that have not been treated (i.e., negative controls). There will be some innate variability in the population, and understanding this variability is essential to setting some “threshold” for “activity” that has an acceptably low level of occurrence in a control population. Figure 4.2 illustrates this approach.

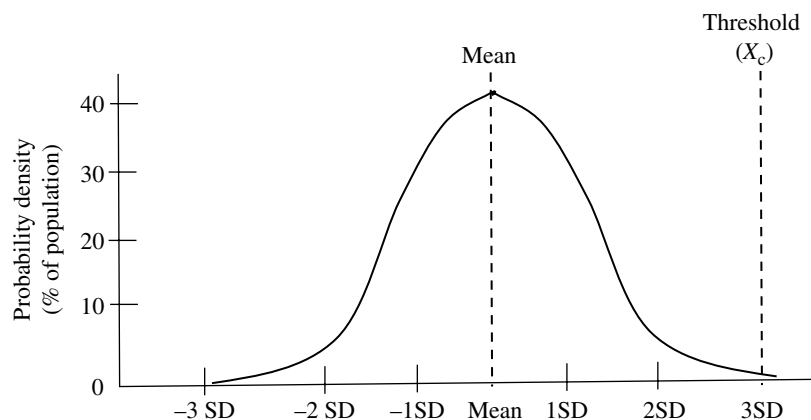


FIGURE 4.2 Setting thresholds using historical control data. The figure shows a Gaussian (“normal”) distribution of screen parameters; 99.7% of the observations in the population are within three standard deviations (SD) of the historic mean. Here the threshold (i.e., the point at which a datum is outside of normal) was set at $X_c = \text{mean} + 3 \text{ SD}$. Note that such a screen is one sided.

What end points are measured as inputs to an activity criterion are intrinsic in the screen system but may be either direct (i.e., having some established mechanistic relationship to the end point that is being predicted in man, such as gene mutations as predictive of carcinogenicity in man) or correlative. Correlated variables (such as many of those measured in *in vitro* systems) are “black box” predictors—compounds causing certain changes in these variables have a high probability of having a certain effect in man, though the mechanisms (or commonality of mechanism) are not established. There is also, as it should be noted, a group of effects seen in animals the relevance of which in man is not known. This illustrates an important point to consider in the design of a screen—one should have an understanding (in advance) of the actions to be taken given each of the possible outcomes of a screen.

4.6 ANALYSIS OF SCREENING DATA

Screening data presents a special case that, due to its inherent characteristics, is not well served by traditional approaches (Gad, 1982b, 1988, 1989a, b, c).

Why? First consider which factors influence the power of a statistical test. Gad (1988) established the basic factors that influence the statistical performance of any bioassay in terms of its sensitivity and error rates. Recently, Healy (1987) presented a review of the factors that influence the power of a study (the ability to detect a dose-related effect when it actually exists). In brief, the power of a study depends on six aspects of study design:

- Sample size
- Background variability (error variance)
- Size of true effect to be detected (i.e., objective of the study)
- Type of significance test
- Significance level
- Decision rule (the number of false positives one will accept)

There are several ways to increase power—each with a consequence:

| Action | Consequence |
|--|--|
| Increase the sample size | Greater resources required |
| Design test to detect larger differences | Less useful conclusions |
| Use a more powerful significance test | Stronger assumptions required |
| Increase the significance level | Higher statistical false-positive rate |
| Use one-tailed decision rule | Blind to effects in the opposite direction |

Timely and constant incorporation of knowledge of test system characteristics and performance will reduce background variability and allow sharper focus on the actual variable of interest. There are, however, a variety of nontraditional approaches to the analysis of screening data.

4.7 UNIVARIATE DATA

4.7.1 Control Charts

The control chart approach (Montgomery, 1985), commonly used in manufacturing quality control in another form of screening (for defective product units), offers some desirable characteristics.

By keeping records of cumulative results during the development of screen methodology, an initial estimate of the variability (such as standard deviation) of each assay will be available when full-scale use of the screen starts. The initial estimates can then be revised as more data are generated (i.e., as we become more familiar with the screen).

The following example shows the usefulness of control charts for control measurements in a screening procedure. Our example test for screening potential muscle strength suppressive agents measures reduction of grip strength by test compounds compared with a control treatment. A control chart was established to monitor the performance of the control agent (1) to establish the mean and variability of the control and (2) to ensure that the results of the control for a given experiment are within reasonable limits (a validation of the assay procedure).

As in control charts for quality control, the mean and average range of the assay were determined from previous experiments. In this example, the screen had been run 20 times previous to collecting the data shown. These initial data showed a mean grip strength \bar{X} of 400 g and a mean range R of 90 g. These values were used for the control chart (Figure 4.3). The subgroups are of size 5. The action limits for the mean and range charts were calculated as follows:

$$\bar{X} \pm 0.58R = 400 \pm 0.58 \times 90 = 348 - 452 \text{ (from the } \bar{X} \text{ chart)}$$

Then, using the upper limit (du) for an n of 5,

$$2.11R = 2.11 \times 90 = 190 \text{ (the upper limit for the range)}$$

Note that the range limit, which actually established a limit for the variability of our data, is, in fact, a “detector” for the presence of outliers (extreme values).

Such charts may also be constructed and used for proportion or count types of data. By constructing such charts for the range of control data, we may then use them as rapid and efficient tools for detecting effects in groups being assessed for that same activity end point.

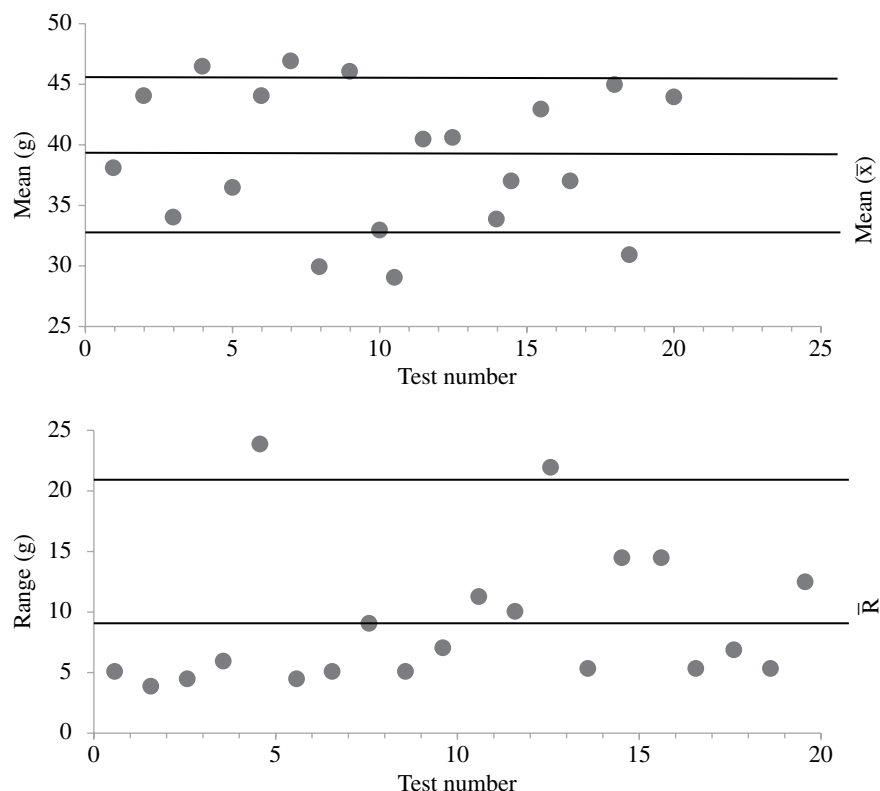


FIGURE 4.3 Example of a control chart used to “prescreen” data (actually, explore and identify influential variables) from a portion of a functional observational battery.

4.7.2 Central Tendency Plots

The objective behind our analysis of screen data is to have a means of efficiently, rapidly, and objectively identifying those agents that have a reasonable probability of being active. Any materials that we so identify may be further investigated in a more rigorous manner, which will generate data that can be analyzed by traditional means. In other words, we want a method that makes out-of-the-ordinary results stand out. To do this we must first set the limits on “ordinary” (summarize the control case data) and then overlay a scheme that causes those things that are not ordinary to become readily detected (“exposed,” in EDA terms) (Velleman and Hoaglin, 1981; Tufte, 1983). One can then perform “confirmatory” tests and statistical analysis (using traditional hypothesis testing techniques), if so desired.

If we collect a set of control data on a variable (say scores on our observations of the righting reflex) from some number of “ordinary” animals, we can plot it as a set of two histograms (one for individual animals and the second for the highest total score in each randomly assigned group of five animals), such as those shown in Figure 4.4 (the data for which came from 200 actual control animals).

Such a plot identifies the nature of our data, visually classifying them into those that will not influence our analysis (in the set shown, clearly scores of 0 fit into this category)

and those that will critically influence the outcome of an analysis. In so doing, the position of control (“normal”) observations is readily revealed as a “central tendency” in the data (hence the name for this technique).

We can (and should) develop such plots for each of our variables. Simple inspection makes clear that answers having no discriminatory power (0 values in Figure 4.4) do not interest us or influence our identifying of an outlier in a group and should simply be put aside or ignored before continuing on with analysis. This first stage, summarizing the control data, thus gives us a device for identifying data with discriminatory power (extreme values), thus allowing us to set aside the data without discriminatory power.

Focusing our efforts on the remainder, it becomes clear that although the incidence of a single, low, nonzero observation in a group means nothing, total group scores of 2 or more occurred only 5% of the time by chance. So we can simply perform an extreme value screen on our “collapsed” data sets, looking for total group values or individual values that are beyond our acceptance criteria.

The next step in this method is to develop a histogram for each ranked or quantal variable by both individual and group. “Useless” data (those that will not influence the outcome of the analysis) are then identified and dropped from analysis. Group scores may then be simply evaluated against



FIGURE 4.4 Plotting central tendency. Possible individual scores for righting reflexes may range from 0 to 8 (Gad, 1982a). Group total scores would thus range from 0 to 40. (Shown are the number of groups that contain individual scores in the individual categories.)

the baseline histograms to identify those groups with scores divergent enough from control to be either true positives or acceptably low-incidence false positives. Additional control data can continue to be incorporated in such a system over time, both increasing the power of the analysis and providing a check on screen performance.

4.7.3 Multivariate Data

The traditional acute, subchronic, and chronic toxicity studies performed in rodents and other species also can be considered to constitute multiple end-point screens. Although the numerically measured continuous variables (body weight, food consumption, hematology values) generally can be statistically evaluated individually by traditional means, the same concerns of loss of information present in the interrelationship of such variables apply. Generally,

traditional multivariate methods are not available, efficient, sensitive, or practical (Young, 1985).

4.7.4 The Analog Plot

The human eye is extremely good at comparing the size, shape, and color of pictorial symbols (Anderson, 1960; Andrews, 1972; Davison, 1983; Schmid, 1983; Cleveland and McGill, 1985). Furthermore, it can simultaneously appreciate both the minute detail and the broad pattern.

The simple way of transforming a table of numbers to a sheet of pictures is by using analog plots. Numbers are converted to symbols according to their magnitude. The greater the number, the larger the symbol. Multiple variables can be portrayed as separate columns or as differently shaped or colored symbols (Wilk and Gnanadesikan, 1986).

The conversion requires a conversion chart from the magnitude of the number to the symbol size. The conversion function should be monotonic (e.g., dose and the measured responses should each change in one direction according to a linear, logarithmic, or probit function). Log conversion will give more emphasis to differences at the lower end of the scale, whereas a probit will stabilize the central range of response (16–84%) of a percentage variable. For example, for numbers x , symbol radius r , and plotting scaling factor k , a log mapping will give:

$$\begin{array}{ll} x=1 & r=k \\ x=10 & r=2k \\ x=100 & r=3k \end{array}$$

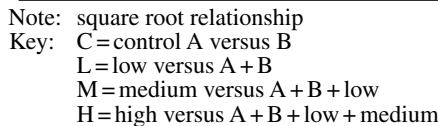
To compare different variables on the same sheet requires some form of standardization to put them on the same scale. Also, a choice must be made between displaying the magnitude of the numbers and their significance (Kruskal, 1964; Kass, 1980). Two possibilities are:

- Express each mean as a percentage change from a control level or overall mean (*a means plot*).
- Calculate effects for meaningful contrasts (*a contrasts plot*).

The analog plot chart in Figure 4.5 shows relationships for five measures on a time versus dose basis, allowing ready evaluation of interrelationships and patterns.

A study using 50 rats of each sex in each of five groups (two controls and three increasing doses) measured body weight and food and liquid consumption every week or month for 2 years. This resulted in 3 variables \times 2 sexes \times 5 groups \times 53 times \times 50 animals. Means alone constituted some 1600 four-digit numbers.

Body weight gains from the period immediately preceding each consumption measurement were used, since these were less correlated. For each variable and at each



In this case, various effects—day of dosing, dose–response, and magnitude of response—are simultaneously portrayed, with the size of each circle being proportional to the magnitude of the measured value.

into four meaningful contrasts:

Control A versus control B

Control A+B versus low

Control A+B+low versus medium

Control A+B+low+medium versus high

were standardized by the within-group standard deviations. Contrast involving doses can be compared with the contrast for the difference between the controls, which should be random. The clearest feature is the high-dose effect for food consumption. However, this seems not to be closely correlated with changes in body weight gains. Certain changes can be seen at the later measurement times, probably because of dying animals.

turing all the information in a set of multiend-point data. When the data are continuous in nature, approaches such as the analog plot can be used (Chernoff, 1973; Chambers et al., 1983). A form of control chart also can be derived for such uses when detecting effect rather than exploring relationships between variables is the goal. When the data are discontinuous,

can be adapted to analyzing attribute data, an analog plot can be adapted. Other methods are also available.

REFERENCES

- Anderson E. (1960) A semigraphical method for the analysis of complex problems. *Technometrics*. 2(3):387–391.
- Anderson S, Hauck W. (1983) A new procedure for testing equivalence in comparative bioavailability and other clinical trials. *Commun. Stat. Theory Methods*. 12:2663–2692.
- Andrews DF. (1972) Plots of high dimensional data. *Biometrics*. 28:125–136.
- Bergman SW, Gittins JC. (1985) Screening procedures for discovering active compounds. In: *Statistical Methods for Pharmaceutical Research Planning* (Peace K, Ed.). Marcel Dekker, New York.
- Bickis MG. (1990) Experimental design. In: *Handbook of In Vivo Toxicity Testing* (Arnold DL, Grice HC and Krewski DR, Eds.). Academic Press, San Diego, pp. 128–134.
- Boersen N, Lee T, Hui H-W. (2013) Development of Preclinical Formulations for Toxicology Studies. In: *A Comprehensive Guide to Toxicology in Preclinical Drug Development* (Faqi AS, Ed.). Academic Press, San Diego, CA, pp. 69–86.

- Bowes J, Brown AJ, Hamon J, Jarolimek W, Sridhar A, Waldron G, Whitebread S. (2012) Reducing safety-related drug attrition: the use of *in vitro* pharmacological profiling. *Nat. Rev. Drug Discov.* 11:909–921.
- Chambers JM, Cleveland WS, Kliner B, Tukey PA. (1983) *Graphical Methods for Data Analysis*. Duxbury Press, Boston, MA.
- Chernoff H. (1973) The use of faces to represent points in K-dimensional space graphically. *J. Am. Stat. Assoc.* 68:361–368.
- Cleveland WS, McGill R. (1985) Graphical perception and graphical methods for analyzing scientific data. *Science*. 229:828–833.
- Davison ML. (1983) *Multidimensional Scaling*. John Wiley & Sons, New York.
- Gad SC. (1982a) A neuromuscular screen for use in industrial toxicology. *J. Toxicol. Environ. Health*. 9:691–704.
- Gad SC. (1982b) Statistical analysis of behavioral toxicology data and studies. *Arch. Toxicol. Suppl.* 5:256–266.
- Gad SC. (1988) An approach to the design and analysis of screening studies in toxicology. *J. Am. Coll. Toxicol.* 8:127–138.
- Gad SC. (1989a) Principles of screening in toxicology with special emphasis on applications to neurotoxicology. *J. Am. Coll. Toxicol.* 8:21–27.
- Gad SC. (1989b) Screens in neurotoxicity: objectives, design and analysis, with the observational battery as a case example. *J. Am. Coll. Toxicol.* 8:1–18.
- Gad SC. (1989c) Statistical analysis of screening studies in toxicology with special emphasis on neurotoxicology. *J. Am. Coll. Toxicol.* 8:171–183.
- Gad SC. (1999) *Statistics and Experimental Design for Toxicologists*, 3rd Ed. CRC Press, Boca Raton, FL.
- Gad SC. (Ed.) (2005) *Drug Discovery Handbook*. John Wiley & Sons, Hoboken, NJ.
- Garrett RA, London JP. (1970) *Fundamentals of Naval Operations Analysis*. US Naval Institute, Annapolis, MD.
- Healy GF. (1987) Power calculations in toxicology. *Altern. Lab. Anim.* 15:132–139.
- Hoaglin DC, Mosteller F, Tukey JW. (1983) *Understanding Robust and Exploratory Data Analysis*. John Wiley & Sons, New York.
- Hoaglin DC, Mosteller F, Tukey JW. (1985) *Exploring Data Tables, Trends, and Shapes*. John Wiley & Sons, New York.
- Kass GV. (1980) An exploratory technique for investigating large quantities of categorical data. *J. Appl. Stat.* 29:119–127.
- Kruskal JB. (1964) Multidimensional scaling by optimizing goodness of fit to a nonmetric hypothesis. *Psychometrika*. 29:1–27.
- Martin YC, Kutter E, Austel V. (1988) *Modern Drug Research*. Marcel Dekker, New York.
- Montgomery DC. (1985) *Introduction to Statistical Quality Control*. John Wiley & Sons, New York.
- Nuwaisir EF, Bittner M, Trent J, Barrett JC, Afshari CA. (1999) Microarrays and toxicology: the advent of toxicogenomics. *Mol. Carcinog.* 24:153–159.
- Page NP. (1977) Concepts of a bioassay program in environmental carcinogenesis. In: *Environmental Cancer. Hemisphere Publishing* (Kraybill HF and Mehlman MA, Eds.), New York, pp. 87–171.
- Pennie WD. (2000) Use of cDNA microarrays to probe and understand the toxicological consequences of altered gene expression. *Toxicol. Lett.* 112–113:473–477.
- Redman C. (1981) Screening compounds for clinically active drugs. In: *Statistics in the Pharmaceutical Industry* (Buncher CR and Tsya J, Eds.). Marcel Dekker, New York, pp. 19–42.
- Schmid CF. (1983) *Statistical Graphics*. John Wiley & Sons, New York.
- Tufte ER. (1983) *The Visual Display of Quantitative Information*. Graphic Press, Cheshire, CT.
- Tukey JW. (1977) *Exploratory Data Analysis*. Addison-Wesley, Reading, PA.
- Velleman PF, Hoaglin DC. (1981) *Applications, Basics and Computing of Exploratory Data Analysis*. Duxbury Press, Boston, MA.
- Wilk MB, Gnanadesikan R. (1986) Probability plotting methods for the analysis of data. *Biometrics*. 55:1–17.
- Young FW. (1985) Multidimensional scaling. In: *Encyclopedia of Statistical Sciences* (Katz S and Johnson NL, Eds.), Vol. 5. John Wiley & Sons, New York, pp. 649–659.
- Zbinden G, Elsner J, Boelsterli UA. (1984) Toxicological screening. *Regul. Toxicol. Pharmacol.* 4:275–286.

FORMULATIONS, ROUTES, AND DOSAGE REGIMENS

The perfect drug would be along the lines of Paul Ehrlich's "magic bullet"—as illustrated in Figure 5.1, a drug molecule is readily administered, completely absorbed, moves to the desired therapeutic target site (receptor), does what is supposed to, and is completely eliminated. The most pressing (and rewarding, if successful) area for drug development currently is optimizing the drug a therapeutic target delivery part of this process. One of the key steps in the nonclinical and clinical formulation of the drug is the selection of vehicles and of the inactive ingredients (excipients). Excipients are essential components of drug products (DP) in the United States, and one must adequately address the safety of the proposed exposure to the excipients in those products. The specific safety data that may be needed will vary depending upon the clinical situation, including such factors as the duration, level, and route of exposure (i.e., actual means of clinical drug administration).

Many guidances exist to aid in the development of pharmaceutical drugs, but very few guidances exist to aid in the formulation of drugs for nonclinical safety evaluation or for the assessment of pharmaceutical excipient safety. The Food and Drug Administration (FDA)/Center for Drug Evaluation and Research (CDER) adopted, in 2005, the guidance for industry "Nonclinical Studies for Development of Pharmaceutical Excipients" which focuses on the development of safety profiles to support use of new excipients as components of drug or biological products.

A similar guidance was published by International Pharmaceutical Excipients Council (IPEC), "Excipient Safety Evaluation Guidance," in 1995 (updated in IPEC (2012)). These guidelines are presented in a tiered approach of recommended data that should be available on an excipient to provide a pharmaceutical formulator with a rational basis for including a new excipient in a drug formulation.

The objective of the current proposal is to codify existing excipients and provide a logical and rational approach to

qualifying new excipients. The final aim of these safety evaluation guidelines for excipients is to provide an important element in the acceptability of a new excipient by regulators independent of the approval of a specific drug formulation.

The three essential requirements of active pharmaceutical ingredient (API) principles are compared with those of excipients. Fundamental to both are quality and safety. The requirement of therapeutic efficacy for drugs is replaced by that of functionality for the excipient, defined as "the physical, physicochemical and biopharmaceutical properties" of the same.

Throughout the development process for pharmaceuticals, formulation development is proceeding with several objectives in mind. The importance of each of these factors changes over time (Monkhouse and Rhodes, 1998) and as illustrated in Figure 5.2. First is optimizing the bioavailability of the therapeutic agent at the target organ site by the intended clinical route. Clinical route(s) is selected on a number of grounds (nature of the drug, patient acceptance, issues of safety, marketability, competition). Second is minimizing any safety concerns. This means not just the systemic toxicity but also local tissue tolerance at the site of application. Third is optimizing the stability of the drug active ingredient. Its activity and integrity must be maintained for a period of sufficient duration to be made effectively available to patients. Early on in preclinical development, simplicity and maximized bioavailability are essential. Early single-dose studies in animals are the starting place and usually bear no relationship to what is used later.

Formulations used to administer potential drugs undergoing development occupy an unusual place in pharmaceutical safety assessment compared to the rest of industrial toxicology. Eventually, a separate function in the pharmaceutical company developing a drug will develop a specific formulation that is to be administered to people—a formulation that optimizes the conditions of absorption and stability for

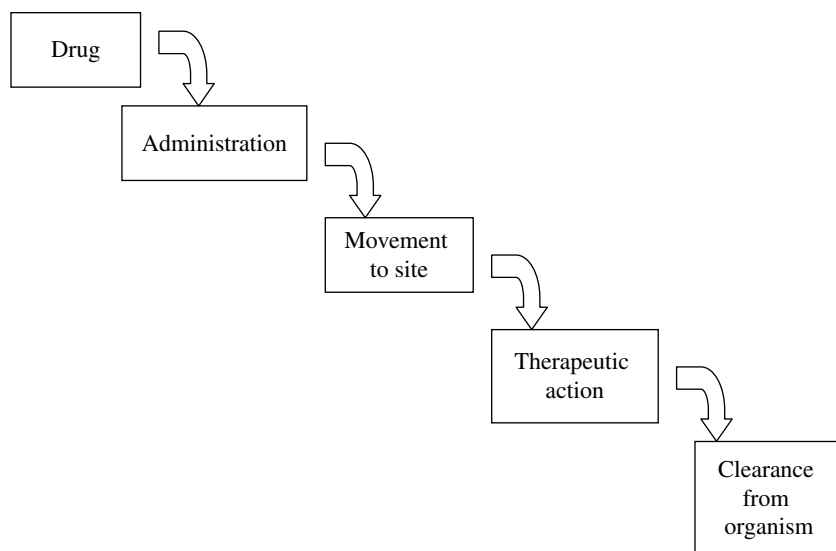


FIGURE 5.1 The magic bullet concept.

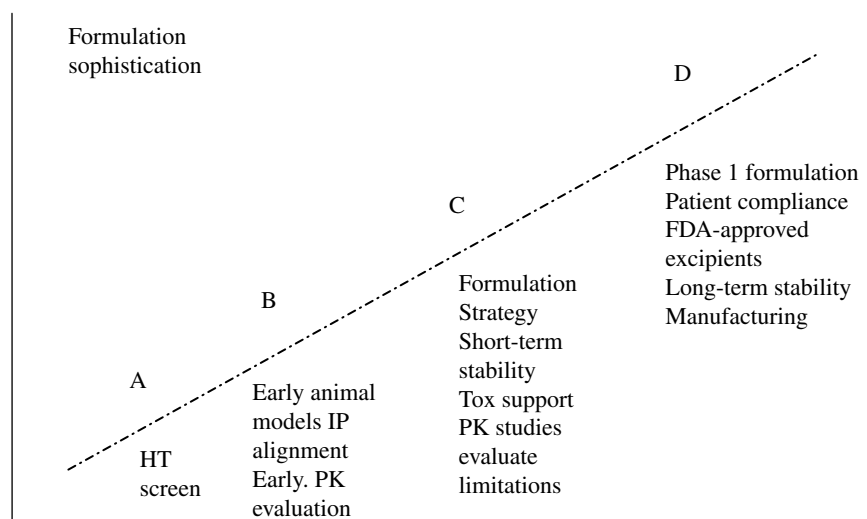


FIGURE 5.2 Evolution of formulations through phase I.

the drug entity (Racy, 1989). The final formulation will need to be assessed to see if it presents any unique local or short-term hazards, but as long as its nonactive constituents are drawn from the approved formulary lists, no significant separate evaluation of their safety is required preclinically. They can, of course, alter the toxicity of the drug under study.

Simultaneous with this development of an optimized clinical formulation, however, preclinical evaluations of the safety of the drug moiety must be performed. Separate preclinical formulations (which generally are less complex than the clinical ones) are developed, sometimes by a formulation group and other times by a toxicology group itself. These preclinical formulations will frequently include much higher concentrations of the drug moiety being tested than do any clinical formulations. The preclinical formulations are developed and

evaluated with the aim of reproducibly delivering the drug (if at all possible by the route intended in man), maintaining drug stability through an optimum period of time, and occluding the observed effects of the drug with vehicle effects to the minimum extent possible. And these preclinical formulations are not restricted to materials that will or could be used in final clinical formulations (Gad, 2008; Boersen et al., 2014).

In pivotal studies, the actual blood levels of an active moiety that are achieved will be determined so that correlations to later clinical studies can be made.

The formulations that are developed and used for preclinical studies are sometimes specific for the test species to be employed, but their development always starts with consideration of the route of exposure that is to be used clinically and, if possible, in accordance with a specified

regimen of treatment (mirroring the intended clinical protocol as much as possible). One aspect of both nonclinical and clinical formulation and testing which prevents an important but often overlooked aspect of pharmaceutical safety assessment is the special field of excipients. These will be considered at the end of this chapter.

Among the cardinal principles of both toxicology and pharmacology is that the means by which an agent comes in contact with or enters the body (i.e., the route of exposure or administration) does much to determine the nature and magnitude of an effect. However a rigorous understanding of formulations, routes, and their implications to the design and analysis of safety studies is not widespread. And in the day-to-day operations of performing studies in animals, such an understanding of routes, their manipulation, means and pitfalls of achieving them, and the art and science of vehicles and formulations is essential to the sound and efficient conduct of a study.

As presented in Table 5.1 there are at least 26 potential routes of administration, of which 10 are commonly used in safety assessment and, therefore, are specifically addressed here.

5.1 MECHANISMS

There are three primary sets of reasons why differences in formulations and the route of administration are critical in determining the effect of an agent of the biological system. These are (1) local effects, (2) absorption and distribution, and (3) metabolism.

5.1.1 Local Effects

Local effects are those that are peculiar to the first area or region of the body to which a test material gains entry or that it contacts. For the dermal route, these include irritation, corrosion, and sensitization. For the parenteral routes, these include irritation, pyrogenicity, sterility, and blood compatibility. In general, these same categories of possible adverse effects (irritation, immediate immune response, local tissue/cellular compatibility, and physicochemical interactions) are the mechanisms of, or basis for, concern.

In general, no matter what the route, certain characteristics will predispose a material to have local effects (and, by definition, if present, tend to limit the possibility of local effects). These factors include pH, redox potential, high molar concentration, and the low level of flexibility (malleability) and sharp edges of certain solids. These characteristics will increase the potential for irritation by any route and, subsequent to the initial irritation, other appropriate regional adaptive responses (for orally administered materials, e.g., emesis and diarrhea).

5.1.2 Absorption and Distribution

For a material to be toxic, it must be absorbed into the organism (local effects are largely not true toxicities by this definition).

TABLE 5.1 Potential Routes of Administration

| |
|---|
| A. Oral routes |
| 1. Oral (PO) ^a |
| 2. Inhalation ^a |
| 3. Sublingual |
| 4. Buccal |
| B. Placed into a natural orifice in the body other than the mouth |
| 1. Intranasal |
| 2. Intra-auricular |
| 3. Rectal |
| 4. Intravaginal |
| 5. Intrauterine |
| 6. Intraurethral |
| C. Parenteral (injected into the body or placed under the skin) |
| 1. Intravenous (IV) ^a |
| 2. Subcutaneous (SC) ^a |
| 3. Intramuscular (IM) ^a |
| 4. Intra-arterial |
| 5. Intradermal (ID) ^a |
| 6. Intralesional |
| 7. Epidural |
| 8. Intrathecal |
| 9. Intracisternal |
| 10. Intracardial |
| 11. Intraventricular |
| 12. Intraocular |
| 13. Intraperitoneal (IP) ^a |
| D. Topical routes |
| 1. Cutaneous ^a |
| 2. Transdermal (also called percutaneous) ^a |
| 3. Ophthalmic ^a |

^a Commonly used in safety assessment.

TABLE 5.2 Normal pH Range for Human Physiologic Fluids

| Medium | Normal pH Range |
|------------------|-----------------|
| Tears | 7.35–7.45 |
| Saliva | 6.0–8.0 |
| Gastric juice | 1.5–6.5 |
| Intestinal juice | 6.5–7.6 |
| Blood | 7.35–7.45 |
| Skin (sweat) | 4.0–6.8 |

There are characteristics that influence absorption by the different routes, and these need to be understood by any person trying to evaluate and/or predict the toxicities of different moieties. Some key characteristics and considerations are summarized in the following by route.

Table 5.2 presents the normal pH ranges for human physiologic fluids. These need to be considered in terms of the impact on solubility and stability of a formulation and active drug:

- A. Oral and rectal routes (gastrointestinal (GI) tract)
 1. Lipid-soluble compounds (nonionized) are more readily absorbed than water-soluble compounds (ionized).

- a. Weak organic bases are in the nonionized, lipid-soluble form in the intestine and tend to be absorbed there.
 - b. Weak organic acids are in the nonionized, lipid-soluble form in the stomach, and one would suspect that they would be absorbed there, but absorption in the intestine is greater because of time and area of exposure.
 2. Specialized transport systems exist for some moieties: sugars, amino acids, pyrimidines, calcium, and sodium.
 3. Almost everything is absorbed—at least to a small extent (if it has a molecular weight below 10000).
 4. Digestive fluids may modify the structure of a chemical.
 5. Dilution increases toxicity because of more rapid absorption from the intestine, unless stomach contents bind or degrade the moiety.
 6. Physical properties are important; for example, dissolution of metallic mercury is essential to allow its absorption.
 7. Age is important; for example, neonates have a poor intestinal barrier.
 8. Effect of fasting on absorption depends on the properties of the chemical of interest.
- B. Inhalation (lungs)**
1. Aerosol deposition
 - a. Nasopharyngeal—5 μm or larger in man, less in common laboratory animals
 - b. Tracheobronchial—1 to 5 μm
 - c. Alveolar—1 μm
 2. If inhalant is a solid, mucociliary transport from lungs to GI tract may clear it out of the pulmonary system.
 3. Lungs are anatomically good for absorption.
 - a. Large surface area (50–100 m^2)
 - b. Blood flow being high
 - c. Accessibility to blood (10 μm between gas media and blood)
 4. Absorption of gases is dependent on solubility of the gas in blood.
 - a. Chloroform, for example, has high solubility and is all absorbed, though respiration is limited.
 - b. Ethylene has low solubility and only a small percentage is absorbed—blood flow limits absorption.
- C. Parenteral routes**
1. Irritation at the site of injection is influenced by solubility, toxicity, temperature, osmolality, and pH of injected solution.
 2. Pyrogenicity and blood compatibility are major concerns for intravenously administered materials.
 3. Solubility of test material in an aqueous or modified aqueous solution is the chief limitation on how much material may be given intravenously (Stiegeland and Noseworthy, 1963).
- D. Dermal routes**
1. In general, any factor that increases the absorption through the stratum corneum will also increase the severity of an intrinsic response. Unless this factor mirrors potential exposure conditions, it may, in turn, adversely affect the relevance of test results.
 2. The physical nature of solids must be carefully considered both before testing and in interpreting results. Shape (sharp edges), size (small particles may abrade the skin due to being rubbed back and forth under the occlusive wrap), and rigidity (stiff fibers or very hard particles will be physically irritating) of solids may all enhance an irritation response and alter absorption.
 3. The degree of occlusion (in fact, the tightness of the wrap over the test site) also alters percutaneous absorption and therefore irritation. One important quality control issue in the laboratory is achieving a reproducible degree of occlusion in dermal wrappings.
 4. Both the age of the test animal and the application site (saddle of the back vs. flank) can markedly alter test outcome. Both of these factors are also operative in humans, of course, but in dermal irritation tests, the objective is to remove all such sources of variability. In general, as an animal ages, the sensitivity to irritation decreases. And the skin on the middle of the back (other than directly over the spine) tends to be thicker (and therefore less sensitive to irritations) than that on the flanks.
 5. The sex of the test animals can also alter study results, because both regional skin thickness and surface blood flow vary between males and females.
- As a generalization, there is a pattern of relative absorption rates that characterizes the different routes that are commonly employed. This order of absorption (by rate from fastest to slowest and, in a less rigorous manner, by degree of absorption from most to least) is IV > inhalation > IM > IP > SC > oral > ID > other.

5.1.3 Metabolism

Metabolism is directly influenced both by the region of the body onto which or into which a material is initially absorbed and by the resultant distribution (both the rate

TABLE 5.3 Selected Factors That May Affect Chemical Distribution to Various Tissues

-
- A. Factors relating to the chemical and its administration
1. Degree of binding of chemical to plasma proteins (i.e., agent affinity for proteins) and tissues
 2. Chelation to calcium, which is deposited in growing bones and teeth (e.g., tetracyclines in young children)
 3. Whether the chemical distributes evenly throughout the body (one-compartment model) or differentially between different compartments (models of two or more compartments)
 4. Ability of chemical to cross the blood–brain barrier
 5. Diffusion of chemical into the tissues or organs and degree of binding to receptors that are and are not responsible for the drug’s beneficial effects
 6. Quantity of chemical given
 7. Route of administration or exposure
 8. Partition coefficients (nonpolar chemicals are distributed more readily to fat tissues than are polar chemicals)
 9. Interactions with other chemicals that may occupy receptors and prevent the drug from attaching to the receptor, inhibit active transport, or otherwise interfere with a drug’s activity
 10. Molecular weight of the chemical
- B. Factors relating to the test subject
1. Body size
 2. Fat content (e.g., obesity affects the distribution of drugs that are highly soluble in fats)
 3. Permeability of membranes
 4. Active transport for chemicals carried across cell membranes by active processes
 5. Amount of proteins in blood, especially albumin
 6. Pathology or altered homeostasis that affects any of the other factors (e.g., cardiac failure, renal failure)
 7. Presence of competitive binding substances (e.g., specific receptor sites in tissues bind drugs)
 8. pH of blood and body tissues
 9. pH of urine^a
 10. Blood flow to various tissues or organs (e.g., well-perfused organs usually tend to accumulate more chemical than less well-perfused organs)
-

^a The pH of urine is usually more important than the pH of blood.

and the pattern). Rate determines whether the primary enzyme systems will handle the entire xenobiotic dose or if these are excessively saturated and overwhelmed. The pattern determines which routes of metabolism are operative.

Absorption (total amount and rate, distribution, metabolism, and species similarity in response) are the reasons for selecting particular routes in toxicology in general. In the safety assessment of pharmaceuticals, however, the route is usually dictated by the intended clinical route and dosing regimen. If this intended route of human exposure is uncertain or if there is the potential for either a number

of other routes where the human absorption rate and pattern are greater, then the most common route or the route with the greater degree of human absorption rate and more desirable pattern, the most common plan forward becomes that of the most conservative approach. This approach stresses maximizing potential absorption in the animal species (within the limits of practicality) and selecting from among those routes commonly used in the laboratory that get the most material into the animal’s system as quickly and completely as possible to evaluate the potential toxicity. Under this approach, many compounds are administered intraperitoneally in acute testing, though there is little or no real potential for human exposure by this route.

Assuming that a material is absorbed, distribution of a compound in early preclinical studies is usually of limited interest. In so-called heavy acute studies (Gad et al., 1984) where acute systemic toxicity is intensive and evaluated to the point of identifying target organs, or in range-finder-type study results, for refining the design of longer-term studies, distribution would be of interest. Some factors that alter distribution are listed in Table 5.3.

The first special case is the parenteral route, where the systemic circulation presents a peak level of the moiety of interest to the body at one time, tempered only by the results of a first pass through the liver.

The second special case arises in cases of inhalation exposures. Because of the arrangement of the circulatory system, inhaled compounds (and those administered via the buccal route) enter the full range of systemic circulation without any “first-pass” metabolism by the liver. Kerberle (1971) and O’Reilly (1972) have previously published reviews of absorption, distribution, and metabolism that are relevant to acute testing.

5.2 COMMON ROUTES

Each of the 10 routes most commonly used in safety assessment studies has its own peculiarities, and for each there are practical considerations and techniques (“tricks”) that should either be known or readily available to the practicing toxicologist.

5.2.1 Dermal Route

For all agents of concern in occupational toxicology (except therapeutic agents), the major route by which the general population is most frequently exposed is via the percutaneous (dermal) route. Brown (1980) has previously reviewed background incidence data on pesticides that shows such a route of exposure to be common. Dermal (or topical) drugs, although not as common, are certainly numerous in occurrence.

Percutaneous entry into the body is really by separate means (Marzulli, 1962; Blank and Scheuplein, 1964; Scheuplein, 1965, 1967):

- Between the cells of the stratum corneum
- Through the cells of the stratum corneum
- Via the hair follicles
- Via the sweat glands
- Via the sebaceous glands

Certain important aspects of the material of interest, as well as those of the test animals, involving application and absorption (Blank and Scheuplein, 1964) are the following:

1. Small molecules penetrate skin better than large molecules.
2. Undissociated molecules penetrate skin better than do ions.
3. Preferential solubility of the toxicant in organic solvents indicates better penetration characteristics than preferential solubility in water.
4. The less viscous or more volatile the toxicant, the greater its penetrability.
5. The nature of the vehicle and the concentration of the toxicant in the vehicle both affect absorption (vehicles are discussed later in this chapter).
6. Hydration (water content) of the stratum corneum affects penetrability.
7. Ambient temperature can influence the uptake of toxicant through the skin. The warmer it is, the greater the blood flow through the skin and, therefore, the greater the degree of percutaneous absorption.
8. Molecular shape (particularly symmetry) influences absorption (Medved and Kundiev, 1964).

There are at least two excellent texts on the subject of percutaneous absorption (Brandau and Lippold, 1982; Bronaugh and Maibach, 1985) that go into great detail.

5.2.2 Parenteral Route

The parenteral routes include three major ones—IV (intravenous), IM (intramuscular), and SC (subcutaneous)—and a number of minor routes (such as intra-arterial) that are not considered here. Administration by the parenteral routes raises a number of special safety concerns in addition to the usual systemic safety questions. These include irritation (vascular, muscular, or SC), pyrogenicity, blood compatibility, and sterility (Avis, 1985). The background of each of these, along with the underlying mechanisms and factors that influence the level of occurrence of such an effect, is discussed in Chapter 11.

The need for a rapid onset of action (and/or clearance) usually requires that an IV route be used, although at a certain stage of cardiopulmonary resuscitation, for example, the need for an even more rapid effect may require the use of an intracardiac injection. The required site of action may influence the choice of route of administration (e.g., certain radiopaque dyes are given intra-arterially near the site being evaluated; streptokinase is sometimes injected experimentally into the coronary arteries close to coronary vessel occlusion during a myocardial infarction to cause lysis of the thrombus and therefore reestablish coronary blood flow).

The characteristics of the fluid to be injected will also influence the choice of parenteral routes. The drug must be compatible with other fluids (e.g., saline, dextrose, Ringer's lactate) with which it may be combined for administration to the patient, as well as with the components of the blood itself.

There are certain clinical situations in which a parenteral route of administration is preferred to other possible routes. These include the following:

1. When the amount of drug given to a subject must be precisely controlled (e.g., in many pharmacokinetic studies), it is preferable to use a parenteral (usually IV) route of administration.
2. When the "first-pass effect" of a drug going through the liver must be avoided, a parenteral route of administration is usually chosen, although a sublingual route or dermal patch will also avoid the first-pass effect.
3. When one requires complete assurance that an uncooperative subject has actually received the drug and has not rejected it (e.g., via forced emesis).
4. When subjects are in a stupor, coma, or otherwise unable to take a drug orally.
5. When large volumes (i.e., more than a liter) of fluid are injected (such as in peritoneal dialysis, hyperalimentation, fluid replacement, and other conditions). Special consideration of fluid balance must be given to patients receiving large volumes, as well as careful consideration of the systemic effects of injection fluid components (e.g., amino acids and their nephrotoxicity).

Each of the three significant parenteral routes we are concerned with here has a specific set of either advantages and disadvantages or specific considerations that must be kept in mind.

5.2.2.1 Intravenous Route The IV route is the most common method of introducing a drug directly into the systemic circulation. It has the following advantages:

1. Rapid onset of effect
2. Usefulness in situations of poor GI absorption

3. Avoidance of tissue irritation that may be present in IM or other routes (e.g., nitrogen mustard)
4. More precise control of levels of drug than with other routes, especially of toxic drugs, where the levels must be kept within narrow limits
5. Ability to administer large volumes over time by a slow infusion
6. Ability to administer drugs at a constant rate over a long period of time

It also suffers from the following disadvantages:

1. Higher incidence of anaphylactic reactions than with many other routes
2. Possibility of infection or phlebitis at site of injection
3. Greater pain to patients than with many other routes
4. Possibility that embolic phenomena may occur—either air embolism or vascular clot—as a result of damage to the vascular wall
5. Impossibility of removing or lavaging a drug after it is given, except by dialysis
6. Inconvenience in many situations
7. Possibility that rapid injection rates may cause severe adverse reactions
8. Patient dislike of, and psychological discomfort with, the injection procedure

For IV fluids, it must be determined how the dose will be given (i.e., by bolus or slow injection, intermittent or constant infusion, or constant drip) and whether special equipment will be used to control and monitor the flow. Drugs with short half-lives are usually given by a constant drip or infusion technique. All IV fluids given immediately subsequent to an IV drug must be evaluated for their compatibility with the study drug. Suspensions are generally not given intravenously because of the possibility of blocking the capillaries.

In the IV route, anaphylactic reactions (caused by administration of an agent to an animal previously sensitized to it or to a particularly sensitive species such as a guinea pig) may be especially severe—probably because of sudden massive antigen–antibody reactions. When the drug is given by other routes, its access to antibody molecules is necessarily slower; moreover, its further absorption can be retarded or prevented at the first sign of a serious allergic reaction.

Embolism is another possible complication of the IV route. Particulate matter may be introduced if a drug intended for IV use precipitates for some reason or if a particular suspension intended for IM or SC use is inadvertently given into a vein. Hemolysis or agglutination of erythrocytes may be caused by injection of hypotonic hypertonic solutions or by more specific mechanisms (Gray, 1978).

5.2.3 Bolus versus Infusion

Technically, for all the parenteral routes (but in practice only for the IV route), there are two options for injecting a material into the body. The bolus and infusion methods are differentiated on the single basis of rate of injection, but they actually differ on a wide range of characteristics.

The most commonly exercised option is the bolus, “push,” injection, in which the injection device (syringe or catheter) is appropriately entered into the vein and a defined volume of material is introduced through the device. The device is then removed. In this operation, it is relatively easy to restrain an experimental animal, and the stress on the animal is limited. Though the person doing the injection must be skilled, it takes only a short amount of time to become so. And the one variable to be controlled in determining dosage is the total volume of material injected (assuming dosing solutions have been properly prepared). See Chapter 9 for a more complete discussion.

For infusion, a port (particularly over a period of weeks or more) (Garramone, 1986) is typically installed to minimize cumulative injection site trauma.

5.2.3.1 Subcutaneous Route Drugs given by the SC route are forced into spaces between connective tissues, as with IM injections. Vasoconstrictors and drugs that cause local irritation should not be given subcutaneously under usual circumstances, since inflammation, abscess formation, or even tissue necrosis may result. When daily or even more frequent SC injections are made, the site of injection should be continually changed to prevent local complications. Fluids given subcutaneously must have an appropriate tonicity/osmolality to prevent pain. Care must be taken to prevent injection of the drug directly into veins.

The absorption of drugs from an SC route is influenced by blood flow to the area, as with IM injections. The rate of absorption may be retarded by cooling the local area to cause vasoconstriction, adding epinephrine to the solution for the same purpose (e.g., with local anesthetics), decreasing blood flow with a tourniquet, or immobilizing the area. The opposite effect may be achieved by warming the injection region or by using the enzyme hyaluronidase, which breaks down mucopolysaccharides of the connective tissue matrix to allow the injected solution to spread over a larger area and thus increase its rate of absorption (Ballard, 1968).

Absorption from SC injection sites is affected by the same factors that determine the rate of absorption from IM sites (Schou, 1971). Blood flow through these regions is generally poorer than in muscles, so the absorption rate is generally slower.

The rate of absorption from an SC injection site may be retarded by immobilization of the limb, local cooling to cause vasoconstriction, or application of a tourniquet proximal to the injection site to block the superficial venous

drainage and lymphatic flow. In small amounts, adrenergic stimulants, such as epinephrine, will constrict the local blood vessels and, therefore, slow systemic absorption. Conversely, cholinergic stimulants (such as methacholine) will induce very rapid systemic absorption subcutaneously. Other agents may also alter their own rate of absorption by affecting local blood supply or capillary permeability.

A prime determinant of the absorption rate from an SC injection is the total surface area over which the absorption can occur. Although the SC tissues are somewhat loose and moderate amounts of fluid can be administered, the normal connective tissue matrix prevents indefinite lateral spread of the injected solution. These barriers may be overcome by agents that break down mucopolysaccharides of the connective tissue matrix; the resulting spread of injected solution leads to a much faster absorption rate.

In addition to fluids, solid forms of drugs may be given by SC injection. This has been done with compressed pellets of testosterone placed under the skin, which are absorbed at a relatively constant rate over a long period.

5.2.3.2 Intramuscular Route The IM route is frequently used for drugs dissolved in oily vehicles or for those in a microcrystalline formulation that are poorly soluble in water (e.g., procaine or penicillin G). Advantages include rapid absorption (often in under 30 min), the opportunity to inject a relatively large amount of solution, and a reduction in pain and local irritation compared with SC injections. Potential complications include infections and nerve damage. The latter usually results from the choice of an incorrect site for injection.

Although the time to peak drug concentration is often on the order of 1–2 h, depot preparations given by IM injection are absorbed extremely slowly. Numerous physiochemical properties of a material given intramuscularly will affect the rate of absorption from the site within the muscle (e.g., ionization of the drug, lipid solubility, osmolality of the solution, volume given). The primary sites used for IM injections in people are the gluteal (buttock), deltoid (upper arm), and epaxial and lateral vastus (lateral thigh) muscles, with the corresponding sites in test animals being species specific. The rate of drug absorption and the peak drug levels obtained will often differ between sites used for IM injections because of differences in blood flow between muscle groups. The site chosen for an IM injection in humans and some animals may be a critical factor in whether or not the drug exhibits an effect (Schwartz et al., 1974). Agents injected into the larger muscle masses are generally absorbed rapidly.

Blood flow through muscles in a resting animal is about $0.02\text{--}0.07\text{ mL min}^{-1}\text{ g}^{-1}$ of tissue, and this flow rate may increase many times during exercise, when additional vascular channels open. Large amounts of solution can be introduced intramuscularly, and there are usually less pain and local irritation than is encountered by the SC route.

Ordinary aqueous solutions of chemicals are usually absorbed from an IM site within 10–30 min, but faster or slower absorption is possible, depending on the vascularity of the site, the ionization and lipid solubility of the drug, the volume of the injection, the osmolality of the solution, animal temperature, and other variables. Small molecules are absorbed directly into the capillaries from an IM site, whereas large molecules (e.g., proteins) gain access to the circulation by way of the lymphatic channels. Radiolabeled compounds of widely differing molecular weights (maximum 585) and physical properties have been shown to be absorbed from rat muscle at virtually the same rate, about 16% per minute (i.e., the absorption process is limited by the rate of blood flow.)

Drugs that are insoluble at tissue pH, or that are in an oily vehicle, form a depot in the muscle tissue, from which absorption proceeds very slowly.

5.2.3.3 Intraperitoneal Route Kruger et al. (1962) demonstrated the efficiency of absorption of some chemicals injected IP, while Lukas et al. (1971) showed that compounds administered IP are absorbed primarily through the portal circulation.

A prime practical consideration in the use of the IP route for acute testing should be the utilization of aseptic techniques to preclude bacterial or viral contamination. If these are not exercised, the resulting infected and compromised animals cannot be expected to produce either valid or reproducible indications or actual chemical toxicity.

Compounds that are very lipophilic will be quickly absorbed systemically by the IP route, but not by the IM or SC route.

5.2.4 Oral Route

The oral route is the most commonly used route for the administration of drugs both because of ease of administration and because it is the most readily accepted route of administration. Although the dermal route may be as common for occupational exposure, it is much easier to accurately measure and administer doses by the oral route.

Enteral routes technically include any that will put a material directly into the GI tract, but the use of enteral routes other than oral (such as rectal) is rare in toxicology. Though there are a number of variations of technique and peculiarities of animal response that are specific to different animal species, there is also a great deal of commonality across species in methods, considerations, and mechanisms.

5.2.4.1 Mechanisms of Absorption Ingestion is generally referred to as oral or peroral (PO) exposure and includes direct intragastric exposure in experimental toxicology. The regions for possible agent action and absorption from PO absorption should receive separate consideration.

Because of the rich blood supply to the mucous membranes of the mouth (buccal cavity), many compounds can be absorbed through them. Absorption from the buccal cavity is limited to nonionized, lipid-soluble compounds. Buccal absorption of a wide range of aromatic and aliphatic acids and basic drugs in human subjects has been found to be parabolically dependent on $\log P$, where P is the octanol–water partition coefficient. The ideal lipophilic character ($\log P_0$) for maximum buccal absorption has also been shown to be in the range 4.2–5.5 (Lien et al., 1971). Compounds with high molecular weights are poorly absorbed in the buccal cavity, and, since absorption increases linearly with concentration and there is generally no difference between optical enantiomorphs of several compounds known to be absorbed from the mouth, it is believed that uptake is by passive diffusion rather than by active transport chemical moieties.

A knowledge of the buccal absorption characteristics of a chemical can be important in a case of accidental poisoning. Although an agent taken into the mouth will be voided immediately on being found objectionable, it is possible that significant absorption can occur before any material is swallowed.

Unless voided, most materials in the buccal cavity are ultimately swallowed. No significant absorption occurs in the esophagus, and the agent passes on to enter the stomach. It is a common practice in safety assessment studies to avoid the possibility of buccal absorption by intubation (gavage) or by the administration of the agent in gelatin capsules designed to disintegrate in the gastric fluid.

Absorption of chemicals with widely differing characteristics can occur at different levels in the GI tract (Hogben et al., 1959; Schranker, 1960; Bates and Gibaldi, 1970; Gad, 2007). The two factors primarily influencing this regional absorption are (1) the lipid–water partition characteristics of the undissociated toxicant and (2) the dissociation constant (pK_a) that determines the amount of toxicant in the dissociated form. There are also species differences (Schranker et al., 1957).

Therefore, weak organic acids and bases are readily absorbed as uncharged lipid-soluble molecules, whereas ionized compounds are absorbed only with difficulty and nonionized toxicants with poor lipid-solubility characteristics are absorbed slowly. Lipid-soluble acid molecules can be absorbed efficiently through the gastric mucosa, but bases are not absorbed in the stomach.

In the intestines the nonionized form of the drug is preferentially absorbed, and the rate of absorption is related to the lipid–water partition coefficient of the toxicant. The highest pK_a value for a base compatible with efficient gastric absorption is about 7.8, and the lowest pK_a for an acid is about 3.0, although a limited amount of absorption can occur outside these ranges. The gastric absorption and the intestinal absorption of a series of compounds with different carbon

chain lengths follow two different patterns. Absorption from the stomach increases as the chain lengthens from methyl to *n*-hexyl, whereas intestinal absorption increases over the range methyl to *n*-butyl and then diminishes as the chain length further increases. Share et al. (1971) and Houston et al. (1974) concluded that to explain the logic of optimal partition coefficients for intestinal absorption, it was necessary to postulate a two-compartment model with a hydrophilic barrier and a lipoidal membrane and that if there is an acceptable optimal partition coefficient for gastric absorption, it must be at least 10 times greater than the corresponding intestinal value.

Because they are crucial to the course of an organism's response, the rate and extent of absorption of biologically active agents from the GI tract also have major implications for the formulation of test material dosages and also for how production (commercial) materials may be formulated to minimize potential accidental intoxications while maximizing the therapeutic profile.

There are a number of separate mechanisms involved in absorption from the GI tract.

Passive Absorption The membrane lining of the tract has a passive role in absorption. As toxicant molecules move from the bulk water phase of the intestinal contents into the epithelial cells, they must pass through two membranes in series, one layer of water and the other the lipid membrane of the microvillar surface (Wilson and Dietschy, 1974). The water layer may be the rate-limiting factor for passive absorption into the intestinal mucosa, but it is not rate limiting for active absorption. The concentration gradient in addition to the physiochemical properties of the drug along with the lining membrane is the major controlling factor. Chemicals that are highly lipid soluble are capable of passive diffusion, and they pass readily from the aqueous fluids of the gut lumen through the lipid barrier of the intestinal wall and into the bloodstream. The interference in the absorption process by the water layer increases with increasing absorbability of the substances in the intestine (Winne, 1978).

Aliphatic carbamates are rapidly absorbed from the colon by passive uptake (Wood et al., 1978), and it is found that there is a linear relationship between $\log k_a$ and $\log P$ for absorption of these carbamates in the colon and the stomach, whereas there is a parabolic relationship between these two values for absorption in the small intestine. The factors to be considered are:

P = octanol–buffer partition coefficient

k_a = absorption rate constant

t = time

$$t^{1/2} = \text{half-life} = \frac{\ln 2}{k_a}$$

Organic acids that are extensively ionized at intestinal pH's are absorbed primarily by simple diffusion.

Facilitated Diffusion A temporary combination of the chemical with some form of "carrier" occurs in the gut wall, facilitating the transfer of the toxicant across the membranes. This process is also dependent on the concentration gradient across the membrane, and there is no energy utilization in making the translocation. In some intoxications, the carrier may become saturated, making this the rate-limiting step in the absorption process.

Active Transport As mentioned, the process depends on a carrier but differs in that the carrier provides energy for translocation from regions of lower concentration to regions of higher concentration.

Pinocytosis This process by which particles are absorbed can be an important factor in the ingestion of particulate formulations of chemicals (e.g., dust formulations, suspensions of wettable powders, etc.); however, it must not be confused with absorption by one of the aforementioned processes, where the agent has been released from particles.

Absorption via Lymphatic Channels Some lipophilic chemicals dissolved in lipids may be absorbed through the lymphatics.

Convective Absorption Compounds with molecular radii of less than 4 nm can pass through pores in the gut membrane. The membrane exhibits a molecular sieving effect.

Characteristically, within certain concentration limits, if a chemical is absorbed by passive diffusion, then the concentration of toxicant in the gut and the rate of absorption are linearly related. However, if absorption is mediated by active transport, the relationship between concentration and rate of absorption conforms to Michaelis–Menten kinetics and a Lineweaver–Burk plot (i.e., reciprocal of rate of absorption plotted against reciprocal of concentration), which graphs as a straight line.

Differences in the physiological chemistry of GI fluids can have a significant effect on toxicity. Both physical and chemical differences in the GI tract can lead to species differences in susceptibility to acute intoxication. The anthelmintic pyrinium chloride has an identical LD_{50} value when administered intraperitoneally to rats and mice ($\sim 4 \text{ mg kg}^{-1}$); when administered orally, however, the LD_{50} value in mice was found to be 15 mg kg^{-1} , while for the rat, the LD_{50} values were 430 mg kg^{-1} for females and 1550 mg kg^{-1} for males. It is thought that this is an absorption difference rather than a metabolic difference (Ritschel et al., 1974).

Most of any exogenous chemical absorbed from the GI tract must pass through the liver via the hepatic portal system (leading to the so-called first-pass effect), and, as mixing of the venous blood with arterial blood from the liver occurs,

consideration and caution are called for in estimating the amounts of chemical in both the systemic circulation and the liver itself.

Despite the GI absorption characteristics discussed earlier, it is common for absorption from the alimentary tract to be facilitated by dilution of the toxicant. Borowitz et al. (1971) have suggested that the concentration effects they observed in atropine sulfate, aminopyrine, sodium salicylate, and sodium pentobarbital were due to a combination of rapid stomach emptying and the large surface area for absorption of the drugs.

Major structural or physiological differences in the alimentary tract (e.g., species differences or surgical effects) can give rise to modifications of toxicity. For example, ruminant animals may metabolize toxicants in the GI tract in a way that is unlikely to occur in nonruminants.

The presence of bile salts in the alimentary tract can affect absorption of potential toxicants in a variety of ways, depending on their solubility characteristics.

5.2.4.2 Factors Affecting Absorption Test chemicals are given most commonly by mouth. This is certainly the most convenient route, and it is the only one of practical importance for self-administration. Absorption, in general, takes place along the whole length of the GI tract, but the chemical properties of each molecule determine whether it will be absorbed in the strongly acidic stomach or in the nearly neutral intestine. Gastric absorption is favored by an empty stomach, in which the chemical, in undiluted gastric juice, will have good access to the mucosal wall. Only when a chemical would be irritating to the gastric mucosa it is rational to administer it with or after a meal. However, the antibiotic griseofulvin is an example of a substance with poor water solubility, the absorption of which is aided by a fatty meal. The large surface area of the intestinal villi, the presence of bile, and the rich blood supply all favor intestinal absorption of griseofulvin and physiochemically similar compounds.

The presence of food can impair the absorption of chemicals given by mouth. Suggested mechanisms include reduced mixing, complexing with substances in the food, and retarded gastric emptying. In experiments with rats, prolonged fasting has been shown to diminish the absorption of several chemicals, possibly by deleterious effects upon the epithelium of intestinal villi.

Chemicals that are metabolized rapidly by the liver cannot be given for systemic effect by the enteral route because the portal circulation carries them directly to the liver. For example, lidocaine, a drug of value in controlling cardiac arrhythmias, is absorbed well from the gut, but is completely inactivated in a single passage through the liver.

The principles governing the absorption of drugs from the GI lumen are the same as for the passage of drugs across biological membranes elsewhere. A lower degree of ionization, high lipid–water partition coefficient of the nonionized form, and a small atomic or molecular radius of water-soluble

substances all favor rapid absorption. Water passes readily in both directions across the wall of the GI lumen. Sodium ion is probably transported actively from the lumen into the blood. Magnesium ion is very poorly absorbed and therefore acts as a cathartic, retaining an osmotic equivalent of water as it passes down the intestinal tract. Ionic iron is absorbed as an amino acid complex, at a rate usually determined by the body's need for it. Glucose and amino acids are transported across the intestinal wall by specific carrier systems. Some compounds of high molecular weight (polysaccharides and large proteins) cannot be absorbed until they are degraded enzymatically. Other substances cannot be absorbed because they are destroyed by GI enzymes—insulin, epinephrine, and histamine are examples. Substances that form insoluble precipitates in the GI lumen or that are insoluble either in water or in lipid clearly cannot be absorbed.

Absorption of Weak Acids and Bases Human gastric juice is very acid (about pH 1), whereas the intestinal contents are nearly neutral (actually very slightly acid). The pH difference between plasma (pH 7.4) and the lumen of the GI tract plays a major role in determining whether a drug that is a weak electrolyte will be absorbed into plasma or excreted from plasma into the stomach or intestine. For practical purposes, the mucosal lining of the GI tract is impermeable to the ionized form of a weak acid or base, but the nonionized form equilibrates freely. The rate of equilibration of the nonionized molecule is directly related to its lipid solubility. If there is a pH difference across the membrane, then the fraction ionized may be considerably greater on one side than on the other. At equilibrium, the concentration of the nonionized moiety will be the same on both sides, but there will be more total drug on the side where the degree of ionization is greater. This mechanism is known as *ion trapping*. The energy for sustaining the unequal chemical potential of the acid or base in question is derived from whatever mechanism maintains the pH difference. In the stomach, this mechanism is the energy-dependent secretion of hydrogen ions.

Consider how a weak electrolyte is distributed across the gastric mucosa between plasma (pH 7.4) and gastric fluid (pH 1.0). In each compartment, the Henderson–Hasselbalch equation gives the ratio of acid–base concentrations. The negative logarithm of the acid dissociation constant is designated here by the symbol pK_a rather than the more precisely correct pK^1 :

$$\begin{aligned} \text{pH} &= pK_a + \log \frac{[\text{base}]}{[\text{acid}]} \\ \log \frac{[\text{base}]}{[\text{acid}]} &= \text{pH} - pK_a \\ \frac{[\text{base}]}{[\text{acid}]} &= \text{antilog} (\text{pH} - pK_a) \end{aligned}$$

The implications of the preceding equations are clear. Weak acids are readily absorbed from the stomach. Weak bases are not absorbed well; indeed, they would tend to accumulate within the stomach at the expense of agent in the bloodstream. Naturally, in the more alkaline intestine, bases would be absorbed better, acids more poorly.

It should be realized that although the principles outlined here are correct, the system is dynamic, not static. Molecules that are absorbed across the gastric or intestinal mucosa are removed constantly by blood flow; thus, simple reversible equilibrium across the membrane does not occur until the agent is distributed throughout the body.

Absorption from the stomach, as determined by direct measurements, conforms, in general, to the principles outlined earlier. Organic acids are absorbed well since they are all almost completely nonionized at the gastric pH; indeed, many of these substances are absorbed well since they are all almost completely nonionized at the gastric pH; indeed, many of these substances are absorbed faster than ethyl alcohol, which had long been considered one of the few compounds that were absorbed well from the stomach. Strong acids whose pK_a values lie below 1, which are ionized even in the acid contents of the stomach, are not absorbed well. Weak bases are absorbed only negligibly, but their absorption can be increased by raising the pH of the gastric fluid.

As for bases, only the weakest are absorbed to any appreciable extent at normal gastric pH, but their absorption can be increased substantially by neutralizing the stomach contents. The quaternary cations, however, which are charged at all pH values, are not absorbed at either pH.

The accumulation of weak bases in the stomach by ion trapping mimics a secretory process; if the drug is administered systemically, it accumulates in the stomach. Dogs given various drugs intravenously by continuous infusion to maintain a constant drug level in the plasma had the gastric contents sampled by means of an indwelling catheter. The results showed that stronger bases ($pK_a > 5$) accumulated in stomach contents to many times their plasma concentrations; the weak bases appeared in about equal concentrations in gastric juice and in plasma. Among the acids, only the weakest appeared in detectable amounts in the stomach. One might wonder why the strong bases, which are completely ionized in gastric juice, whose theoretical concentration ratios (gastric juice/plasma) are very large should nevertheless attain only about a 40-fold excess over plasma. Direct measurements of arterial and venous blood show that essentially all the blood flowing through the gastric mucosa is cleared of these agents; obviously, no more chemical can enter the gastric juice in a given time period than is brought there by circulation. Another limitation comes into play when the base pK_a exceeds 7.4; now a major fraction of the circulating base is cationic, and a decreasing fraction is nonionized, so the effective concentration gradient for diffusion across the stomach wall is reduced.

The ion-trapping mechanism provides a method of some forensic value for detecting the presence of alkaloids (e.g., narcotics, cocaine, amphetamines) in cases of death suspected to be due to overdosage of self-administered drugs. Drug concentrations in gastric contents may be very high even after parenteral injection.

Absorption from the intestine has been studied by perfusing drug solutions slowly through rat intestine *in situ* and by varying the pH as desired. The relationships that emerge from such studies are the same as those for the stomach, the difference being that the intestinal pH is normally very near neutrality. As the pH is increased, the bases are absorbed better, the acids more poorly. Detailed studies with a great many drugs in unbuffered solutions revealed that in the normal intestine, acids with $pK_a > 3.0$ and bases with $pK_a < 7.8$ are very well absorbed; outside these limits the absorption of acids and bases falls off rapidly. This behavior leads to the conclusion that the “virtual pH” in the microenvironment of the absorbing surface in the gut is about 5.3; this is somewhat more acidic than the pH in the intestinal lumen is usually considered to be.

Absorption from the buccal cavity has been shown to follow exactly the same principles as those described for absorption from the stomach and intestine. The pH of human and canine saliva is usually about 6. Bases in people are absorbed only on the only on the alkaline side of their pK_a , that is, only in the nonionized form. At normal saliva pH, only weak bases are absorbed to a significant extent.

5.2.4.3 Bioavailability and Thresholds The difference between the extent of availability (often designated solely as bioavailability) and the rate of availability is illustrated in Figure 5.3, which depicts the concentration–time curve for a hypothetical agent formulated into three different dosage forms. Dosage forms A and B are designed so that the agent is put into the blood circulation at the same rate, but twice as fast as for dosage form C. The times at which agent concentrations reach a peak are identical for dosage forms A and B and occur earlier than the peak time for dosage form C. In general, the relative order of peak times following the administration of different dosage forms of the drug corresponds to the rates of availability of the chemical moiety from the various dosage forms. The extent of availability can be measured by using either chemical concentrations in the plasma or blood or amounts of unchanged chemical in the urine. The area under the blood concentration–time curve for an agent can serve as a measure of the extent of its availability. In Figure 5.3, the areas under curves A and C are identical and twice as great as the area under curve B. In most cases, where clearance is constant, the relative areas under the curves or the amount of unchanged chemical excreted in the urine will quantitatively describe the relative availability of the agent from the different dosage forms. However, even in nonlinear cases, where clearance is dose

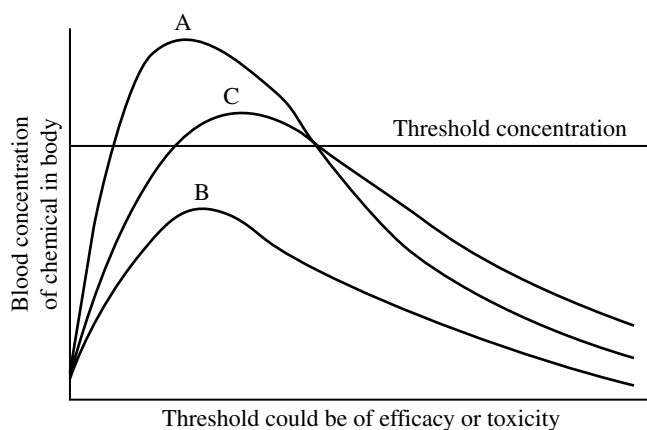


FIGURE 5.3 Three different systemic absorption curves.

dependent, the relative areas under the curves will yield a measurement of the rank order of availability from different dosage forms or from different routes of administration.

Because there is usually a critical concentration of a chemical in the blood that is necessary to elicit either a pharmacological or toxic effect, both the rate and extent of input or availability can alter the toxicity of a compound. In the majority of cases, the duration of effects will be a function of the length of time the blood–concentration curve is above the threshold concentration; the intensity of the effect for many agents will be a function of the elevation of the blood–concentration curve above the threshold concentration.

Thus, the three different dosage forms depicted in Figure 5.3 will exhibit significant differences in their levels of “toxicity.” Dosage form B requires that twice the dose be administered to attain blood levels equivalent to those for dosage form A. Differences in the rate of availability are particularly important for agents given acutely. The dosage for A reaches the target concentration earlier than chemical from dosage form C; concentrations from A reach a higher level and remain above the minimum effect concentration for a longer period of time. In a multiple dosing regimen, dosage forms A and C will yield the same average blood concentrations, although dosage form A will show somewhat greater maximum and lower minimum concentrations.

For most chemicals, the rate of disposition or loss from the biological system is independent of rate of input, once the agent is absorbed. Disposition is defined as what happens to the active molecule after it reaches a site in the blood circulation where concentration measurements can be made (the systemic circulations, generally). Although disposition processes may be independent of input, the inverse is not necessarily true, because disposition can markedly affect the extent of availability. Agents absorbed from the stomach and the intestine must first pass through the liver before reaching the general circulation (Figure 5.4). Thus, if a compound is metabolized in the liver or excreted in bile, some of the active molecule absorbed from the GI

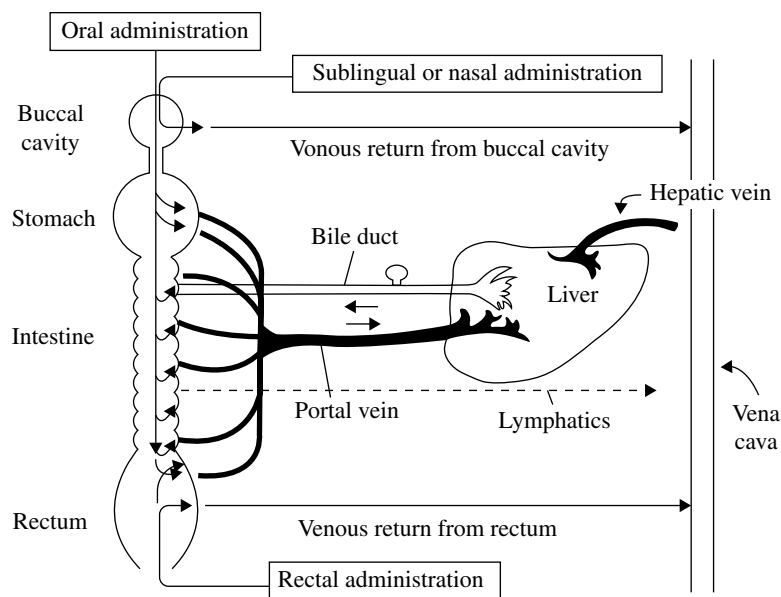


FIGURE 5.4 Path of drugs through the body after absorption by one of three routes of administration.

tract will be inactivated by hepatic processes before it can reach the systemic circulation and be distributed to its sites of action. If the metabolizing or biliary excreting capacity of the liver is great, the effect on the extent of availability will be substantial. Thus, if the hepatic blood clearance for the chemical is large, relative to hepatic blood flow, the extent of availability for this chemical will be low when it is given by a route that yields first-pass metabolic effects. This decrease in availability is a function of the physiological site from which absorption takes place, and no amount of modification to dosage form can improve the availability under linear conditions. Of course, toxic blood levels can be reached by this route of administration if larger doses are given.

It is important to realize that chemicals with high extraction ratios (i.e., greater extents of removal by the liver during first-pass metabolism) will exhibit marked intersubject variability in bioavailability because of variations in hepatic function or blood flow or both. For the chemical with an extraction ratio of 0.90 that increases to 0.95, the bioavailability of the agent will be halved, from 0.10 to 0.05. These relationships can explain the marked variability in plasma or blood drug concentrations that occurs among individual animals given similar doses of a chemical that is markedly extracted. Small variations in hepatic extraction between individual animals will result in large differences in availability and plasma drug concentrations.

The first-pass effect can be avoided, to a great extent, by use of the sublingual route and by topical preparations (e.g., nitroglycerine ointment), and it can be partially avoided by using rectal suppositories. The capillaries in the lower and middle sections of the rectum drain into the interior and

middle hemorrhoidal veins, which, in turn, drain into the inferior vena cava, thus bypassing the liver. However, suppositories tend to move upward in the rectum into a region where veins that lead to the liver predominate, such as the superior hemorrhoidal vein. In addition, there are extensive connections between the superior and middle hemorrhoidal veins, and thus probably only about 50% of a rectal dose can be assumed to bypass the liver. The lungs represent a good temporary clearing site for a number of chemical (especially basic compounds) by partition into lipid tissues, as well as serve a filtering function for particulate matter that may be given by IV injection. In essence, the lung may cause first-pass loss by excretion and possible metabolism for chemicals input into the body by the non-GI routes of administration.

Biological (test subject) factors that can influence absorption of a chemical from the GI tract are summarized in Table 5.4.

There are also a number of chemical factors that may influence absorption from the GI tract. These are summarized in Table 5.5.

5.2.4.4 Techniques of Oral Administration There are three major techniques for oral delivery of drugs to test animals. The most common way is by gavage, which requires that the material be in a solution or suspension for delivery by tube to the stomach. Less common materials may be given as capsules (particularly to dogs) or in diet (for longer-term studies). Rarely, oral studies may also be done by inclusion of materials in drinking water.

Test materials may be administered as solutions or suspensions as long as they are homogeneous and delivery is

TABLE 5.4 Test Subject Characteristics That Can Influence GI Tract Absorption^a

-
- A. General and inherent characteristics
 - 1. General condition of the subject (e.g., starved vs. well fed, ambulatory vs. supine)
 - 2. Presence of concurrent diseases (i.e., diseases may either speed or slow gastric emptying)
 - 3. Age
 - 4. Weight and degree of obesity
 - B. Physiological function
 - 1. Status of the subject's renal function
 - 2. Status of the subject's hepatic function
 - 3. Status of the subject's cardiovascular system
 - 4. Status of the subject's gastrointestinal motility and function (e.g., ability to swallow)
 - 5. pH of the gastric fluid (e.g., affected by fasting, disease, food intake, drugs)
 - 6. Gastrointestinal blood flow to the area of absorption
 - 7. Blood flow to areas of absorption for dose forms other than those absorbed through gastrointestinal routes
 - C. Acquired characteristics
 - 1. Status of the subject's anatomy (e.g., previous surgery)
 - 2. Status of the subject's gastrointestinal flora
 - 3. Timing of drug administration relative to meals (i.e., presence of food in the gastrointestinal tract)
 - 4. Body position of subject (e.g., lying on the side slows gastric emptying)
 - 5. Psychological state of subject (e.g., stress increases gastric emptying rate, and depression decreases rate)
 - 6. Physical exercise of subject may reduce gastric emptying rate
 - D. Physiological principles
 - 1. Food enhances gastric blood flow, which should theoretically increase the rate of absorption
 - 2. Food slows the rate of gastric emptying, which should theoretically slow the rate of passage to the intestines where the largest amounts of most agents are absorbed. This should decrease the rate of absorption for most agents. Agents absorbed to a larger extent in the stomach will have increased time for absorption in the presence of food and should be absorbed more completely than in fasted patients
 - 3. Bile flow and secretion are stimulated by fats and certain other foods. Bile salts may enhance or delay absorption depending on whether they form insoluble complexes with drugs or enhance the solubility of agents
 - 4. Changes in splanchnic blood flow as a result of food depend on direction and magnitude of the type of food ingested
 - 5. Presence of active (saturable) transport mechanisms places a limit on the amount of a chemical that may be absorbed
-

^a The minimization of variability due to these factors rests on the selection of an appropriate animal model, careful selection of healthy animals, and use of proper techniques.

TABLE 5.5 Chemical Characteristics of a Drug That May Influence Absorption

-
- A. Administration of chemical and its passage through the body
 - 1. Dissolution characteristics of solid dosage forms, which depend on formulation in addition to the properties of the chemical itself (e.g., vehicle may decrease permeability of suspension or capsule to water and retard dissolution and diffusion)
 - 2. Rate of dissolution in gastrointestinal fluids. Chemicals that are inadequately dissolved in gastric contents may be inadequately absorbed
 - 3. Chemicals that are absorbed into food may have a delayed absorption
 - 4. Carrier-transported chemicals are more likely to be absorbed in the small intestine
 - 5. Route of administration
 - 6. Chemicals undergo metabolism in the gastrointestinal tract
 - B. Physiochemical properties of chemicals
 - 1. Chemicals that chelate metal ions in food may form insoluble complexes and will not be adequately absorbed
 - 2. pH of dosing solutions—weakly basic solutions are absorbed to a greater degree in the small intestine
 - 3. Salts used
 - 4. Hydrates or solvates
 - 5. Crystal form of chemical (e.g., insulin)
 - 6. "Pharmaceutical" form (e.g., fluid, solid, suspension)
 - 7. Enteric coating
 - 8. Absorption of quaternary compounds (e.g., hexamethonium, amiloride) is decreased by food.
 - 9. Molecular weight of chemical (e.g., when the molecular weight of a drug is above about 1000, absorption is markedly decreased)
 - 10. pK_a (dissociation constant)
 - 11. Lipid solubility (i.e., a hydrophobic property relating to penetration through membranes)
 - 12. Particle size of chemical in solid dosage form—smaller particle sizes will increase the rate and/or degree of absorption if dissolution of the chemical is the rate-limiting factor in absorption. Chemicals that have a low dissolution rate may be made in a micronized form to increase their rate of dissolution
 - 13. Particle size of the dispersed phase in an emulsion
 - 14. Type of disintegrating agent in the formulation
 - 15. Hardness of a solid (granule, pellet, or tablet) (i.e., related to amount of compression used to make tablet) or capsule if they do not disintegrate appropriately
-

accurate. For traditional oral administration (gavage), the solution or suspension can be administered with a suitable stomach tube or feeding needle ("Popper" tube) attached to a syringe. If the dose is too large to be administered at one time, it can be divided into equal subparts with 2–4 h between each administration; however, this subdivided dosing approach should generally be avoided.

Test chemicals placed into any natural orifice exert local effects and, in many instances, systemic effects as well. The possibility of systemic effects occurring when local effects are to be evaluated should be considered.

For routes of administration in which the chemical is given orally or placed into an orifice other than the mouth, clear instructions about the correct administration of the chemical must be provided. Many cases are known of oral pediatric drops for ear infections being placed into the ear, and vice versa (eardrops being swallowed) in humans. Errors in test-article administration are especially prevalent when a chemical form is being used in a nontraditional manner (e.g., suppositories that are given by the buccal route).

Administration of drug in a capsule is a common means of dosing larger test animals (particularly dogs). It is labor intensive (each capsule must be individually prepared, though robotic systems are now available for this), but capsules offer the advantages that neat drug may be used (no special formulation need be prepared, and the questions of formulation or solution stability are avoided), the dogs are less likely to vomit, and the actual act of dosing requires less labor than using a gavage tube. Capsules may also be used with primates, though they are not administered as easily.

Incorporation of a drug in the diet is commonly used for longer-term studies (particularly carcinogenicity studies, though the method is not limited to these). Dosing by diet is much less labor intensive than any other oral dosing methodology, which is particularly attractive over the course of a long (13-week, 1-year, 18-month, or 2-year) study.

The most critical factor to dietary studies is the proper preparation of the test chemical–diet admixtures. The range of physical and chemical characteristics of test materials requires that appropriate mixing techniques be determined on an individual basis. Standard practices generally dictate the preparation of a premix, to which is added appropriate amounts of feed to achieve the proper concentrations.

Dietary preparation involving liquid materials frequently results in either wet feed in which the test article does not disperse or formation of “gumballs”—feed and test material that form discernible lumps and chemical “hot spots.” Drying and grinding of the premix to a free-flowing form prior to mixing the final diets may be required; however, these actions can affect the chemical nature of the test article.

Solid materials require special techniques prior to or during addition to diets. Materials that are soluble in water may be dissolved and added as described previously for liquids. Nonwater-soluble materials may require several preparatory steps. The test chemical may be dissolved in corn oil, acetone, or other appropriate vehicle prior to addition to the weighed diet. When an organic solvent such as acetone is used, the mixing time for the premix should be sufficient for the solvent to evaporate. Some solids may require grinding in a mortar and pestle with feed added during the grinding process.

Prior to study initiation, stability of the test chemical in the diet must be determined over a test period at least equivalent to the time period during which animals are to be exposed to a specific diet mix. Stability of test samples under the conditions of the proposed study is preferable. Labor and expense can be saved when long-term stability data permit mixing of several weeks (or a month) of test diet in a single mixing interval.

Homogeneity and concentration analysis of the test-article–diet admixture are performed by sampling at three or four regions within the freshly mixed diet (e.g., samples from the top, middle, and bottom of the mixing bowl or blender).

A variety of feeders are commercially available for rats and mice. These include various-sized glass jars and stainless steel or galvanized feed cups, which can be equipped with restraining lids and food followers to preclude significant losses of feed due to animals digging in the feeders. Slotted metal feeders are designed so that animals cannot climb into the feed, and they also contain mesh food followers to prevent digging.

Another problem sometimes encountered is palatability—the material may taste so strongly that animals will not eat it. As a result, palatability, stability in diet, and homogeneity of mix must all be ensured prior to the initiation of an actual study.

Inclusion in drinking water is rarely used for oral administration of human drugs to test animals, though it sees more frequent use for the study of environmental agents.

Physiochemical properties of the test material should be a major consideration in selecting drinking water as a dosing matrix. Unlike diet preparation or preparation of gavage dose solutions and suspensions where a variety of solvents and physical processes can be utilized to prepare a dosable form, preparations of drinking water solutions are less flexible. Water solubility of the test chemical is the major governing factor and is dependent on factors such as pH, dissolved salts, and temperature. The animal model itself sets limitations for these factors (acceptability and suitability of pH and salt-adjusted water by the animals as well as animal environmental specifications such as room temperature).

Stability of the test chemical in drinking water under study conditions should be determined prior to study initiation. Consideration should be given to conducting stability tests on test chemical–drinking water admixtures presented to some test animals. Besides difficulties of inherent stability, changes in chemical concentrations may result from other influences. Chemicals with low vapor pressure can volatilize from the water into the air space located above the water of an inverted water bottle; thus, a majority of the chemical may be found in the “dead space,” not in the water.

Certain test chemicals may be degraded by contamination with microorganisms. A primary source of these

microorganisms is the oral cavity of rodents. Although rats and mice are not as notorious as the guinea pig in spitting back into water bottles, significant bacteria can pass via the sipper tubes and water flow restraints into the water bottles. Sanitation and sterilization procedures for water bottles and sipper tubes must be carefully attended to.

Many technicians may not be familiar with terms such as sublingual (under the tongue), buccal (between the cheek and gingiva), otic, and so on. A clear description of each of these nontraditional routes (i.e., other than gavage routes) should be discussed with technicians, and instructions may also be written down and given to them. Demonstrations are often useful to illustrate selected techniques of administration (e.g., to use an inhaler or nebulizer). Some chemicals must be placed by technicians into body orifices (e.g., medicated intrauterine devices such as Progestasert).

5.2.5 Minor Routes

The minor routes do see some use in safety assessment and at least three should be briefly presented here.

5.2.5.1 Periocular Route The administration of drugs or accidental exposure of chemicals to the eyes is not commonly a concern in systemic toxicity due to the small surface area exposed and the efficiency of the protective mechanisms (i.e., blink reflex and tears). As long as the epithelium of the eyes remains intact, it is impermeable to many molecules, but, if the toxicant has a suitable polar–nonpolar balance, penetration may occur (Kondrizer et al., 1959; Swan and White, 1942).

Holmstedt (1959) and Brown and Muir (1971) have reviewed periocular absorption of pesticides. More recently, Sinow and Wei (1973) have shown that the quaternary herbicide paraquat can be lethal to rabbits if applied directly to the surface of the eyes. Parathion, in particular, is exceedingly toxic when administered via the eye—a concern that must be kept in mind for the protection of pesticide applicators.

5.2.5.2 Rectal Administration Since a number of therapeutic compounds are given in the form of suppositories, an indication of the toxicity after rectal administration is sometimes required. Toxicity studies and initial drug formulations of such compounds are usually performed by the oral route, and the rectal formulation comes late in development and marketing. In view of the difference between laboratory animals and man in the anatomy and microflora of the colon and rectum, animal toxicity studies late in drug development are of limited value. However, in a case where there is an indication of potential rectal hazard or bioavailability, the compound may be introduced into the rectum of the rat using an oral dosing needle to prevent

tissue damage. To avoid the rapid excretion of the unabsorbed dose, anesthetized animals should be used, and the dose retained with an inert plug or bung (such as a cork).

Drugs (and, therefore, test chemicals) are occasionally administered by rectum, but most are not as well absorbed here as they are from the upper intestine. Aminophylline, used in suppository form for the management of asthma, is one of the few drugs routinely given in this way. Inert vehicles employed for suppository preparations include cocoa butter, glycerinated vehicles, gelatin, and polyethylene glycol. Because the rectal mucosae are irritated by nonisotonic solutions, fluids administered by this route should always be isotonic with plasma (e.g., 0.9% NaCl).

5.2.5.3 Vaginal Administration Though not a common one, some materials do have routine exposure by this route (e.g., spermicides, tampons, douches, and antibiotics) and, therefore, must be evaluated for irritation and toxicity by this route. The older preferred models used rabbits and monkeys (Eckstein et al., 1969), but, more recently, a model that uses rats has been developed (Staab et al., 1987). McConnell (1974) clearly described the limitations, particularly of volume of test material, involved in such tests.

5.2.5.4 Nasal Administration A route that has gained increasing popularity of late for pharmaceutical administration in man is the intranasal route. The reasons for this popularity are the ease of use (and, therefore, ready patient acceptance and high compliance rate), the high degree and rate of absorption of many substances (reportedly for most substances up to 1000 molecular weight; McMartin et al., 1987), and the avoidance of the highly acid environment in the stomach and first-pass metabolism in the liver (particularly important for some of the newer peptide moieties) (Attman and Dittmer, 1971). The only special safety concerns are the potential for irritation of the mucous membrane and the rapid distribution of administered materials to the CNS.

A number of means may be used to administer materials nasally—nebulizers and aerosol pumps being the most attractive first choices. Accurate dose administration requires careful planning, evaluation of the administration device, and attention to technique.

5.2.5.5 Volume Limitations by Route In the strictest sense, absolute limitations on how much of a dosage form may be administered by any particular route are determined by specific aspects of the test species or dosage form. But there are some general guidelines (determined by issues of humane treatment of animals, accurate deliver of dose and such) that can be put forth. These are summarized in Table 5.6. Appendix E and Section 5.3 should, of course, be checked to see if there is specific guidance due to the characteristics of a particular vehicle.

TABLE 5.6 Volume Guidelines for Administration of Compounds by Route of Administration to Laboratory Animals

| Species | Route | | | | | | | | | | | | | | | |
|------------------|----------------|------------------------------|-----------------|----------------|-----------------|------|----------------|-------|--------------------------|----------|------------------|--------------------------|-----------------------------|--------------------------------------|--------------|------|
| | Intranasal | | | | IV bolus** | | | | SC*** | | | | IM mL.kg ⁻¹ * | | | |
| | µL per nostril | Gavage mL.kg ⁻¹ * | Ideal | Max | Ideal | Max | Ideal | Max | Ideal | Max | Ideal | Max | Ideal | Max | Ideal | Max |
| Vol. | | | IV bolus** | | IP*** | | SC*** | | IM mL.kg ⁻¹ * | | IV infusion** | | Continuous infusion | | Intravaginal | |
| | | | Ideal | Max | Ideal | Max | Ideal | Max | Ideal | Max | Vol. | Rate mL.kg ⁻¹ | Max ^{vi} mL (kg h) | Cath. Maint. Rate mL.h ⁻¹ | Max | Vol. |
| Mouse | 25 | 10 | 40 ^a | 5 | 25 | 5–10 | 50 | 1–5 | 20 ^f | 1 | 1 | 50 | 1 | NIA | NIA | 2 |
| Rat | 50 | 10 | 20 ^b | 1 | 20 | 5–10 | 20 | 1 | 20 ^f | 1 | 10 ^g | 50 | 1 | 5 | 0.35 | 5 |
| | | | | | | | | | | | | | | | | |
| Guinea pig | 100 | 10 | 30 | 1 | 5 | 1–5 | 20 | 1–5 | 10 ^f | 0.1 | 0.5–1 | 0.1 | 0.1 | 10 | 1 | NIA |
| Rabbit | 200 | 1 | 20 ^c | 1 ^d | 10 ^d | 3 | 5 ^e | 1–2.5 | 10 ^f | 0.1–0.5 | 1 | 0.1 | 0.1 | 20 | 1 | NIA |
| Dog | 500 | 5 | 20 | 1 | 10 | 3 | 5 | 0.5 | 2 ^f | 0.1–0.25 | 1 ^h | 0.1 | 0.1 | 20 | 5 | 5 |
| Nonhuman primate | 200 | 5 | 10 | 1 | 10 | 3 | 5 | 0.5 | 2 ^f | 0.1–0.5 | 1 ^h | 0.1 | 0.1 | 20 | 1 | 5 |
| Mini-swine | NIA | NIA | NIA | 1 | 10 | 1 | 5 | 1 | 3 ^f | 0.25 | 0.5 ⁱ | 0.1 | 0.1 | 10 | 1 | 5 |

Source: Data from Baker et al. (1979), Spector (1956), and Swenson (1977) and adapted with changes from Gad et al. (2016).

* Single dose per day except where noted otherwise.

** Solution properties such as tonicity, pH, etc. need to be taken into account when approaching the volume limits or determining the volume to be infused IV. The recommended working range for pH is 4.5–8.0. The order of degree of tolerance of pH for different dosing routes is oral>intravenous>intramuscular>subcutaneous>intraperitoneal. Animal health must also be taken into consideration, such as kidney function and cardiovascular function. These systems must be normal to handle increased fluid volumes.

*** When administering a solution IP, the viscosity, concentration, tonicity, and pH of the solution need to be taken into account.

**** When administering a solution SC, the concentration, tonicity, and pH of the solution must be taken into account.

***** NIA—no information currently available.

^a To accommodate a larger volume, the dose may be divided over time (e.g., 20 mL.kg⁻¹ administered four times per day to reach a total of 80 mL.kg⁻¹ in a 24 h period).

^b To accommodate a larger volume, the dose may be divided over time (e.g., 10 mL.kg⁻¹ administered four times per day to reach a total of 40 mL.kg⁻¹ in a 24 h period).

^c Rabbits should not be fed prior to administration. Rabbits should be fed after the completion of dose administration.

^d These volumes may also be used for intra-arterial injection.

^e Not often used.

^f If volumes greater than those cited earlier are used, the volume must be divided over multiple sites.

^g May be used if divided over multiple sites and alternating legs, maximum of 5 sites per leg. Final volume not to exceed 0.10 mL.

^h 0.05 mL total volume limit per site.

ⁱ 3 mL total volume limit per site.

^j 5 mL total volume limit per site.

^k Up to 20 mL.kg⁻¹ if divided over multiple sites.

^l 3 mL total volume limit.

^m Repeat dose 67 µL in the rabbit.

ⁿ Solution properties such as tonicity, pH, etc., need to be taken into account when determining the volume that may be infused IV. Animal health must also be taken into consideration, such as kidney function and cardiovascular function. These systems need to be normal to handle increased fluid volumes.

5.2.6 Route Comparisons and Contrasts

The first part of this chapter described, compared, and contrasted the various routes used in toxicology and presented guidelines for their use. There are, however, some exceptions to the general rules that the practicing toxicologist should keep in mind.

The relative ranking of efficacy of routes that was presented earlier in the chapter is not absolute; there can be striking exceptions. For example, though materials are usually much quicker acting and more potent when given by the oral route than by the dermal, this is not always the case. In the literature, Shaffer and West (1960) reported that tetram as an aqueous solution was more toxic when applied dermally than when given orally to rats. LD₅₀s reported were as follows (LD₅₀ (mg kg⁻¹) of tetram; 95% confidence limits):

| Rat | Oral (mg kg ⁻¹) | Percutaneous (mg kg ⁻¹) |
|--------|-----------------------------|-------------------------------------|
| Male | 9 (7–13) | 2 (1–3) |
| Female | 8 (6–11) | 2 (1–3) |

The author has in the past experienced this same phenomenon. Several materials that were found to be relatively nontoxic orally were extremely potent by the dermal route (differences in potency of more than an order of magnitude have been seen at least twice).

A general rule applicable to routes and vehicles should be presented here.

5.2.6.1 Vehicles that Can Mask the Effects of Active Ingredients Particularly for clinical signs, attention should be paid to the fact that a number of vehicles (e.g., propylene glycol) cause transient neurobehavioral effects that may mask similar short-lived (though not necessarily equally transient and reversible) effects of test materials.

5.3 FORMULATION OF TEST MATERIALS

One of the areas that is overlooked by virtually everyone in toxicology testing and research, yet is of crucial importance, is the need for formulation of candidate drugs and the use of vehicles and excipients in the formulation of test chemicals for administration to test animals (Strickley, 2008). For a number of reasons, a drug of interest is rarely administered or applied as is (“neat”). Rather, it must be put in a form that can be accurately given to animals in such a way that it will be absorbed and not be too irritating. Most laboratory toxicologists come to understand vehicles and formulation, but to the knowledge of the author, guidance on the subject is limited to a short chapter on

formulations by Fitzgerald et al. (1983). There is also a very helpful text on veterinary dosage forms by Hardee and Baggot (1998).

Table 5.7 presents an overview of typical forms of excipients for oral dosage forms. The entire process of drug development—even during the preclinical and nonclinical phase—includes a continuous development (with increased sophistication) of formulation. The use of nanoparticles (actually, developed as seeking even finer *micronized* particles in dosing formulations; it has been around for decades) is the latest approach.

Regulatory toxicology in the United States can be said to have arisen, due to the problem of vehicles and formulation, in the late 1930s, when attempts were made to formulate the new drug sulfanilamide. This drug is not very soluble in water, and a US firm called Massengill produced a clear, syrupy elixir formulation that was easy to take orally. The figures illustrate how easy it is to be misled. The drug sulfanilamide is not very soluble in glycerol, which has an LD₅₀ in mice of 31.5 g kg⁻¹, but there are other glycols that have the characteristic sweet taste and a much higher solvent capacity. Ethylene glycol has an LD₅₀ of 13.7 g kg⁻¹ in mice and 8.5 g kg⁻¹ in rats, making it slightly more toxic than diethylene glycol, which has an LD₅₀ in rats of 20.8 g kg⁻¹, similar to that for glycerol. The drug itself, which is inherently toxic, was marketed in a 75% aqueous diethylene glycol-flavored elixir. Then early in 1937 appeared the first reports of deaths, but the situation remained obscure for about 6 months until it became clear that the toxic ingredient in the elixir was the diethylene glycol. Even as late as March 1937, Haag and Ambrose were reporting that the glycol was excreted substantially unchanged in dogs, suggesting that it was likely to be safe (Hagenbusch, 1937). Within a few weeks, Holick (1937) confirmed that a low concentration of diethylene in drinking water was fatal to a number of species. Hagenbusch (1937) found that the results of necropsies performed on patients who had been taking 60–70 mL of the solvent per day were similar to those of rats, rabbits, and dogs taking the same dose of solvent with or without the drug. This clearly implicated the solvent, although some authors considered that the solvent was simply potentiating the toxicity of the drug. Some idea of the magnitude of this disaster may be found in the paper of Calvary and Klump (1939), who reviewed 105 deaths and a further 2560 survivors who were affected to varying degrees, usually with progressive failure of the renal system. It is easy to be wise after the event, but the formulator fell into a classic trap, in that the difference between acute and chronic toxicity had not been adequately considered. In passing, the widespread use of ethylene glycol itself as an antifreeze has led to a number of accidental deaths, which suggests that the lethal dose in man is around 1.4 mL kg⁻¹, or a volume of about 100 mL. In the preface to the first US Pharmacopeia (USP), published in

TABLE 5.7 Examples of Excipients in Drug Formulation (Other than Vehicles)

| | | | |
|---|---|--|--|
| <p><u>Tablets and capsules</u></p> <p>FILLER OR DILUENT <i>Increases volume or weight</i> Examples: calcium phosphate, lactose</p> <p>Binder <i>Facilitates agglomeration of powder into granules</i> Examples: povidones, starches</p> <p>Disintegrant <i>Promotes rapid disintegration to allow a drug to dissolve faster</i> Examples: sodium starch glycolate, croscopovidones</p> <p>Lubricant <i>Reduces friction between particles themselves and between particles and manufacturing equipment</i> Examples: magnesium stearate, glycerides</p> | <p>Glidant or anticaking agent <i>Promotes powder flow and reduces caking or clumping</i> Examples: talc, colloidal silicon dioxide</p> <p>Colorant <i>Produces a distinctive appearance and may protect light-sensitive ingredients</i> Examples: FC&C colors, titanium dioxide</p> <p>Capsule shell <i>Contains powders or liquids</i> Examples: gelatin, hypromellose</p> <p>Coating agent <i>May mask unpleasant tastes or odors, improve ingestion or appearance, protect ingredients from the environment, or modify release of the active ingredient</i> Examples: shellac, hypromellose</p> | <p>Flavor and fragrance <i>Improves palatability</i> Examples: peppermint, berry</p> <p>Release modifier <i>Provides extended-release capability</i> Examples: ethylcellulose, guar gum</p> <p><u>Oral liquids</u></p> <p>PH MODIFIER <i>Controls pH to improve drug stability or avoid irritation when consumed</i> Examples: citric acid and its salts, salts of phosphoric acid</p> <p>Wetting or solubilizing agent <i>Promotes dissolution of insoluble ingredients</i> Examples: sodium lauryl sulfate, polysorbates</p> | <p>Antimicrobial preservative <i>Prevents growth of bacteria, yeast, and mold</i> Examples: glycerin, benzyl alcohol</p> <p>Chelating or complexing agent <i>Stabilizes ions</i> Examples: ethylenediaminetetraacetic acid salts, cyclodextrins</p> <p>Antioxidant <i>Reduces oxidative reactions that could alter ingredients</i> Examples: ascorbic acid, butylated hydroxyanisole</p> <p>Sweetening agent <i>Improves palatability</i> Examples: sucrose, saccharin</p> <p>SOURCE: USP</p> |
|---|---|--|--|

Extracted from Kemsley (2014).

1820, there is a the statement that “It is the object of the Pharmacopoeia to select from among substances which possess medical power, those, the utility of which is most fully established and best understood; and to form from them preparations and compositions, in which their powers may be exerted to the greatest advantage.” This statement suggests that the influence that formulation and preparation may have on the biological activity of a drug (and on nonpharmaceutical chemicals) has been appreciated for a considerable time.

Available and commonly used vehicles and formulating agents are reviewed, along with basic information on their characteristics and usages, in Appendix E at the end of this book. There is a general presumption that those excipients and formulating agents listed in the Pharmacopoeia or in the *Inactive Ingredient Guide* prepared by the FDA are safe to use and without biological effect. This may not always be the case in either experimental or humans (see Weiner and Bernstein, 1989), either directly or in how they alter absorption of and response to the active ingredient.

There are some basic principles to be observed in developing and preparing test material formulations. These are presented in Table 5.8. A start to all of this should be preformulation–characterization of the chemical, physical, and

physical chemistry aspects of the drug molecule (Neervannan, 2006). Parts of this are identifying the optimal phase of the drug for formulation (Palucki et al., 2010) and useful salts (Stahl and Wermuth, 2011).

5.3.1 Preformulation

While advances in molecular biology and genomics have produced a flood of molecules with vastly improved target receptor specificity, these molecules have frequently turned out to be very very difficult to get absorbed and to the desired target tissue site.

Lipinski’s rule of five (RO5) (Lipinski et al., 2001; Lipinski, 2004) predates these recent target advances in specificity, but not the problems. In its original form, the RO5 proposed four guiding principles:

1. No more than five hydrogen bond donors (the total number of nitrogen–hydrogen and oxygen–hydrogen bonds)
2. Not more than 10 hydrogen bond acceptors (all nitrogen or oxygen atoms)
3. A molecular mass less than 500 Da
4. An octanol–water partition coefficient ($\log P$) not greater than 5

TABLE 5.8 Basic Principles to Be Observed in Developing and Preparing Test Material Formulations

-
- | | |
|----|--|
| A. | Preparation of the formulation should not involve heating of the test material anywhere near to the point where its chemical or physical characteristics are altered |
| B. | If the material is a solid and it is to be assessed for dermal effects, its shape and particle size should be preserved. If intended for use in man, topical studies should be conducted with the closest possible formulation to that to be used on humans |
| C. | Multicomponent test materials (mixtures) should be formulated so that the administered form accurately represents the original mixture (i.e., components should not be selectively suspended or taken into solution) |
| D. | Formulation should preserve the chemical stability and identity of the test material |
| E. | The formulation should be such as to minimize total test volumes. Use just enough solvent or vehicle |
| F. | The formulation should be easy to administer accurately |
| G. | pH of dosing formulations should be between 5 and 9, if possible |
| H. | Acids or bases should not be used to divide the test material (for both humane reasons and to avoid pH partitioning in either the gut or the renal tubule) |
| I. | If a parental route is to be employed, final solutions should be as nearly isotonic as possible. Do not assume a solution will remain such upon injection into the bloodstream. It is usually a good idea to verify that the drug stays in solution upon injection by placing some drops into plasma |
| J. | Particularly if use is to be more than a single injection, steps (such as filtration) should be taken to ensure suitable sterility |
-

While Lipinski *prima facie* applies to oral route drugs, it also is useful for other routes.

For clinically useful drugs (and therefore for drugs proceeding through preclinical and nonclinical evaluation and development), there are a number of desirable attributes:

- A simple structure
- Simple and efficient synthesis
- Nonhygroscopic
- Avoidance of chiral centers
- Lack of mutagenicity
- Crystalline with solid-state stability
- No strong odors, colors, or (of oral) tastes
- Compatible with standard excipient
- Stable at ambient temperatures and at physiologic pH's (Niazi, 2007; Adeyeye and Brittain, 2008; Gibson, 2009)

Bioavailability is defined as the fraction of the dose reaching either the therapeutic target organ or tissue or the systemic circulation as unchanged compound following administration by any route. For an agent administered orally, bioavailability may be less than unity, for several

reasons. The molecule may be incompletely absorbed. It may be metabolized in the gut, the gut wall, the portal blood, or the liver prior to entry into the systemic circulation (see Figure 5.4). It may undergo enterohepatic cycling with incomplete reabsorption following elimination into the bile. Biotransformation of some chemicals in the liver following oral administration is an important factor in the pharmacokinetic profile, as will be discussed further. Bioavailability measures following oral administration are generally given as the percentage of the dose available to the systemic circulation.

As the components of a mixture may have various physicochemical characteristics (solubility, vapor pressure, density, etc.), great care must be taken in preparing and administering any mixture so that what is actually tested is the mixture of interest.

Examples of such procedures are making dilutions (not all components of the mixture may be equally soluble or miscible with the vehicle) and generating either vapors or respirable aerosols (not all the components may have equivalent volatility or surface tension, leading to a test atmosphere that contains only a portion of the components of the mixture).

By increasing or decreasing the viscosity of a formulation, the absorption of a toxicant can be altered (Ritschel et al., 1974). Conversely, the use of absorbents to diminish absorption has been used as an antidote therapy for some forms of intoxication. Using the knowledge that rats cannot vomit, there have been serious attempts at making rodenticides safer to nontarget animals by incorporating emetics into the formulations, but this has had only a limited degree of success. Gaines used *in vivo* liver perfusion techniques to investigate the apparent anomaly that the carbamate Isolan was more toxic when administered to rats percutaneously than when administered orally (Gaines, 1960). It has been shown that these results, a manifestation of different formulations, have been used for the two routes of exposure (oral and percutaneous) in estimating the LD₅₀ values using a common solvent, *n*-octanol. It was found that Isolan was significantly more toxic by the oral route than by the percutaneous route; by regression analysis it was found that at no level of lethal dose values was the reverse correct.

Although the oral route is the most convenient for most therapeutic uses, there are numerous factors that make it unpredictable, particularly for drug molecules that have very limited water solubility (Liu, 2008). Absorption by this route is subject to significant variation from animal to animal and even in the same individual animal at different times. Considerable effort has been spent by the pharmaceutical industry to develop drug formulations with absorption characteristics that are both effective and dependable. Protective enteric coatings for pharmaceuticals were introduced long ago to retard the action of gastric fluids and then disintegrate and dissolve after passage of a

tablet into the human intestine. The purposes of these coatings for drugs are to protect the active ingredient, which would be degraded in the stomach, to prevent nausea and vomiting caused by local gastric irritation (also a big problem in rodent studies, where over a long time period gastric irritation frequently leads to forestomach hyperplasia), to obtain higher local concentrations of the active ingredient intended to act locally in the intestinal tract, to produce a delayed biological effect, or to deliver the active ingredient to the intestinal tract for optimal absorption there. Such coatings are generally fats, fatty acids, waxes, or other such agents, and all of these intended purposes for drug delivery can readily be made to apply for some toxicity studies. Their major drawback, however, is the marked variability in time for a substance to be passed through the stomach. In humans, this gastric emptying time can range from minutes to as long as 12 h. One would expect the same for animals, as the limited available data suggest is the case. Similar coating systems, including microencapsulation (see Melnick et al., 1987), are available for, and currently used in, animal toxicity studies.

The test chemical is unlikely to be absorbed or excreted unless it is first released from its formulation. It is this stage of the process that is the first and most critical step for the activity of many chemicals. If the formulation does not release the chemical, the rest of the process becomes somewhat pointless.

It might be argued that the simplest way around the formulation problem is to administer any test as a solution in water, thereby avoiding the difficulties altogether. However, since multiple, small, accurately measured doses of a chemical are required repeatedly, reproducible dilutions must be used. Also, the water itself is to be regarded as the formulation vehicle, and the test substance must be water soluble and stable in solution and many are not. If one takes into account this need for accuracy, stability, and optimum performance *in vivo*, the whole problem becomes very complex.

It is uncommon for direct connections between observed toxicity and formulation components to be made. Indeed, it is usually assumed that vehicles and other nontest chemical components are innocuous or have only transitory pharmacological effects. Historically, however, this has certainly not been the case. Even lactose may have marked toxicity in individual test animals (or humans) who are genetically incapable of tolerating it.

The initial stage of drug release from the formulation, both in terms of the amount and the rate of release, may exercise considerable influence at the clinical response level. A close consideration of the formulation parameters of any chemical is therefore essential during the development of any new drug, and, indeed, there are examples where formulations of established drugs also appear to require additional investigation.

The effects of formulation additives on chemical bioavailability from oral solutions and suspensions have been previously reviewed by Swarbrick (2006). He pointed out how the presence of sugars in a formulation may increase the viscosity of the vehicle. However, sugar solutions alone may delay stomach-emptying time considerably when compared to solutions of the same viscosity prepared with celluloses, which may be due to sugar's effect on osmotic pressure. Sugars of different types may also have an effect on fluid uptake by tissues, and this, in turn, correlates with the effect of sugars such as glucose and mannitol on drug transport.

Surfactants have been explored widely for their effects on drug absorption, in particular using experimental animals (Gibaldi and Feldman, 1970; Gibaldi, 1976). Surfactants alter dissolution rates (of lipid materials), surface areas of particles and droplets, and membrane characteristics, all of which affect absorption.

Surfactants may increase the solubility of the drug via micelle formation, but the amounts of material required to increase solubility significantly are such that at least orally the laxative effects are likely to be unacceptable. The competition between the surfactant micelles and the absorption sites is also likely to reduce any useful effect and make any prediction of net overall effect difficult. However, if a surfactant has any effect at all, it is likely to be in the realm of agents that help disperse suspensions of insoluble materials and make them available for solution. Natural surfactants, in particular bile salts, may enhance absorption of poorly soluble materials.

The effective surface area of an ingested chemical is usually much smaller than the specific surface area that is an idealized *in vitro* measurement. Many drugs whose dissolution characteristics could be improved by particle-size reduction are extremely hydrophobic and may resist wetting by GI fluids. Therefore, the GI fluids may come in intimate contact with only a fraction of the potentially available surface area. The effective surface area of hydrophobic particle can often be increased by the addition of a surface-active agent to the formulation, which reduces the contact angle between the solid and the GI fluids, thereby increasing effective surface area and dissolution rate.

Formulations for administering dermally applied toxicants present different considerations and problems. The extent of penetration and speed with which a biologically active substance penetrates the skin or other biological membrane depend on the effect that the three factors—vehicle, membrane, and chemical—exert on the diffusion process. It is now accepted that they together represent a functional unit that controls the penetration and location of the externally applied chemicals in the deeper layers of the skin or membrane layer. The importance of the vehicle for the absorption process has been neglected until recently. One of the few requirements demanded of the vehicle has been that it acts as an inert medium that incorporates the test

chemical in the most homogeneous distribution possible. In addition, chemical stability and good cosmetic appearance have been desirable characteristics. Most formulations in toxicology are based on empirical experience.

The chemical incorporated in a vehicle should reach the surface of the skin at a suitable rate and concentration. If the site of action lies in the deeper layers of the epidermis or below, the substance must cross the stratum corneum, if the skin is intact. Both processes, diffusion from the dosage form and diffusion through the skin barriers, are inextricably linked. They should be considered simultaneously and can be influenced by the choice of formulation.

The thesis that all lipid-soluble compounds basically penetrate faster than water-soluble ones cannot be supported in this absolute form. A lipophilic agent can penetrate faster or slower or at the same rate as a hydrophilic agent, depending on the vehicle used.

Disregarding such chemical-specific properties as dissociation constants (in the case of ionic compounds), particle size, and polymorphism, as well as side effects of viscosity, binding to vehicle components, complex formulation, and the like, the following formulation principles arise:

- a. Optimization of the concentration of chemical capable of diffusion by testing its maximum solubility
- b. Reduction of the proportion of solvent to a degree that is adequate to keep the test material still in solution
- c. Use of vehicle components that reduce the permeability barriers

These principles lead to the conclusion that each test substance requires an individual formulation. Sometimes different ingredients will be required for different concentrations to obtain the maximum rate of release. No universal vehicle is available for any route, but a number of primary approaches are. Any dosage preparation lab should be equipped with glassware, a stirring hot plate, a sonicator, a good homogenizer, and a stock of the basic formulating material, as detailed at the end of this chapter.

5.3.2 Dermal Formulations

Preparing formulations for application to the skin has special considerations associated with it, which, in the case of human pharmaceuticals, has even led to a separate book (Barry, 1983).

The physical state of the skin is considerably affected by external factors such as relative humidity, temperature, and air movement at the skin surface. If this contact is broken (e.g., by external applications of ointments or creams), it is reasonable to assume that the new skin will change in some way, sometimes to an extent that creates new conditions of

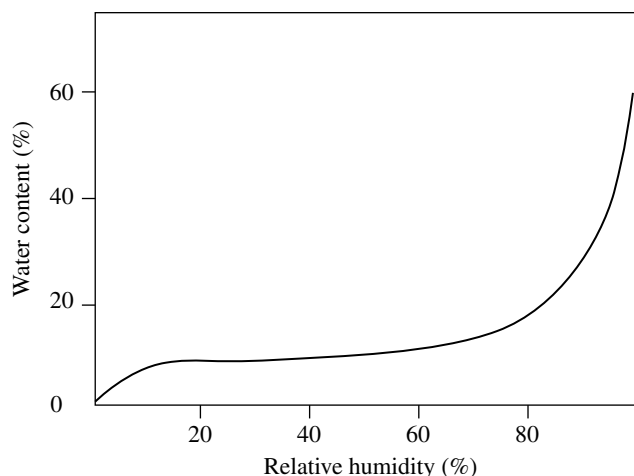


FIGURE 5.5 Course of moisture absorption of skin.

permeability for the test material. This would be the case, for example, if the stratum corneum becomes more hydrated than normal due to the topical delivery form. Temperature might also have an effect, as is the case when any constituents of the vehicle affect the inner structure of the skin through interactions with endogenous skin substances. Often several of these processes occur together. Figure 5.5 shows the relationship between water content (%) and relative humidity (%).

Since this is a question of interactions between the vehicle and the skin (and the latter cannot be viewed as an inert medium), the composition of the vehicle itself may be altered (e.g., by incorporation into skin constituents or through loss of volatile components).

The first contact between vehicle and skin occurs on the skin's surface. The first phase of interaction undoubtedly begins with the lipid mantle, in the case of so-called normal skin. If the skin has been damaged by wounds, the surface can form a moist milieu of serous exudate, resulting in abnormal wetting properties. Normally it is impregnated with oily sebaceous secretions and horny fat, presenting a hydrophobic surface layer. Water will not spread out as a film but will form droplets, while bases with a high affinity to the skin surface constituents spread spontaneously into a film and can wet. In the case of a base of a low viscosity, the degree of wetting can often be determined by measuring the angle of contact. If the preparation wets the skin surface, is drawn by capillary action from the visible area into the large inner surface of the stratum corneum, and is transported away into the interior, then it is said that the ointment or cream penetrates well. Spreading and wetting are purely surface phenomena, not penetration in the strict sense. If the skin shows a high content of its own lipids, spreading is limited. It is also reduced if the value of the surface tension of the skin (σ_s) decreases compared to the value of the interfacial force between the skin and subject liquid ($\gamma_{s/l}$) and the surface tension of the subject liquid (σ_l), as is the case with

aqueous bases. Addition of amphiphilic compounds decreases σ_1 and $y_{s/1}$ and thus spreadability increases.

How much the endogenous emulsifying substances of the fatty film, such as cholesterol esters and fatty acid salts, affect this spreading process is not clear. They can probably promote the emulsification of hydrophobic substances with water. Whether the sebaceous and epidermal lipids alone are sufficient to emulsify water and so form a type of emulsive film remains controversial. However, it is assumed that they, together with appropriate vehicle components, improve the spreading of the applied vehicle and that this effect can be potentiated by mechanical means such as intensive rubbing in. A good spreadability ensures that the active ingredient is distributed over a large area.

High local concentrations are avoided, and, at the same time, close contact is made between the chemical and the upper layers of the skin.

In grossly simplified terms, hydrogels, suspensions, and water emulsions behave on the skin surface similarly to aqueous solutions. By contrast, pastes and water-oil emulsions act like oil. The ability of an organic solvent to stick or wet depends on its specific properties (e.g., its viscosity and its surface tension).

At present, the information concerning alterations in vehicle composition on the skin surface is sparse. However, two possible extremes are conceivable. On the one hand, if the vehicle has a high vapor pressure, it often completely evaporates shortly after application. On the other hand, the vehicle may remain on the skin surface in an almost completely unchanged composition, (e.g., highly viscous Vaseline or similar thick covering systems). Between these two extremes lie the remaining types of vehicles.

The first situation applies for the short-chain alcohols, acetone, or ether. After their evaporation, the drug remains finely dispersed on or in the skin at 100% concentration.

If individual components evaporate, the structure of the vehicle changes and, under certain circumstances, also the effective drug concentration. Oil-water emulsions lose water rapidly, giving rise to the well-known cooling effect. If evaporation continues, the dispersed oil phase coalesces and forms a more or less occlusive film on the skin, together with the emulsifier and the drug. Of course, it is possible that a certain hydrophilic proportion of the drug is then present in suspended form or at least can react with charged molecules and is thus removed from the diffusion process at the start; at the same time, it is to be expected that soluble constituents of the skin are incorporated so that a new system can be formed on the surface and the adjoining layers of skin. Comparable transformations probably also occur after application of water-oil emulsions, providing one realizes that the water evaporates more slowly, the cooling effect is less strong, and, due to the water-oil character of the molecule, the occlusive effect can be more marked because of the affinity of the oily components for the skin.

By contrast, Vaseline and similar highly viscous, lipid bases from the outset form an impenetrable layer, virtually unaffected by external factors or effects emanating from the skin itself. Interactions with the skin lipids are only likely at the boundary between ointment and skin.

The evaporation of the water from the skin into the atmosphere is a continuous process. It can be increased or decreased by the use of suitable vehicles. An evaporation increase will always occur if the water vapor from the vehicle is taken away more quickly than water can diffuse from the deeper layers into the stratum corneum. This applies in principle to all hydrophilic bases, particularly for systems with an oil-water character. After loss of most of their own water, hydrophilic bases develop a true draining effect that can lead to the drying out of the underlying tissue. How much the penetration of hydrophilic drugs can be proved with the help of oil-water systems depends on the solution properties of the rest of the components in the skin. Generally, such compounds can only seldom reach deeper layers. It is equally difficult to show an adequate release of water from hydrophilic systems to a dry skin. If any such effects do occur, they are short term and are quickly overtaken by opposing processes. The same seems to apply to most of the traditional moisturizers such as glycerin and propylene glycol (Powers and Fox, 1957; Rieger and Deems, 1974). They can also cause a large rise in the rate of evaporation, depending on the relative humidity, and thus increase the transepidermal loss of water. It is probably impossible to prevent this drying out without preparations having some occlusive properties.

In contrast, vehicles that are immiscible with water and those with a high proportion of oils have occlusive effects. They reduce both insensible perspiration and the release of sweat. The sweat collects as droplets at the opening of the glands, but does not spread as a film between the hydrophobic skin surface and the lipophilic base because the free surface energy of the vehicle-skin interface is smaller than that between water and skin. If a lipophilic layer of vehicle is present, this is not spontaneously replaced by the water-skin layer if sweat is secreted.

The horny layer consists of about 10% extracellular components such as lipids, proteins, and mucopolysaccharides. Around 5% of the protein and lipids form the cell wall. The majority of the remainder is present in the highly organized cell contents, predominantly as keratin fibers, which are generally assigned an α -helical structure. They are embedded in a sulfur-rich amorphous matrix, enclosed by lipids that probably lie perpendicular to the protein axis. Since the stratum corneum is able to take up considerably more water than the amount that corresponds to its volume, it is assumed that this absorbed fluid volume is mainly located in the region of these keratin structures.

Some insight into where on the relative humidity continuum water molecules are absorbed can be gained from

equilibrium isotherms (Ziegenmeyer, 1982) (Figure 5.5), which show a characteristic sigmoidal shape. At low relative humidity, water is first absorbed at specific skin sites, probably in the region of the peptide compounds and the various polar side chains. At higher moisture content, layers of water form on the skin. By using Zimm–Lundberg cluster theory (Zimm and Lundberg, 1956), additional information can be obtained about the nature of the absorbed water.

Because of thick intertwining protein fibers in the cell and in the area of the cell membrane, cell structure is rigid and remains so but is altered by the osmotic effect of the penetrating water. The uptake of water entails a continual shifting of the cell matrix, which gradually develops elastic opposing forces that increasingly resist further expansion. Eventually an equilibrium is reached if both forces balance each other. In the case of water, it takes quite a long time to completely hydrate the cell. This process can, however, be shortened if there are components present with a solvent effect diffusing out of a basic vehicle. The duration and degree of swelling depend on the affinity of all the dissolved substances for the tissue and on the size of the maximum possible elastic reaction, which stabilizes cell structure.

5.3.3 Interactions between Skin, Vehicle, and Test Chemical

The diffusion coefficient of the hydrated stratum corneum is larger than that of dry skin. Therefore, hydration increases the rate of passage of all substances that penetrate the skin. If the hydrated keratin complex is represented by a biphasic system, then it can be considered to exist as a continuous region covered with layers of water and intervening layers of lipids. Nonpolar compounds are predominantly dissolved in the nonpolar lipid matrix and diffuse through it. Polar substances, by contrast, pass through the aqueous layers. The diffusion of water and low molecular weight hydrophilic molecules through these layers of water is more difficult than a corresponding free diffusion in an aqueous solution. This could, under certain circumstances, be due to a higher degree of organization of water in the protein structures (than in plasma or the free state), in the sense that this water is only available as the driving force of the diffusion process to a limited degree.

The degree of hydration can be controlled by the choice of vehicle. Lipophilic paraffin bases are available, but vehicles such as water–oil emulsions are more acceptable since they are less occlusive and offer ease of formulation.

In principle, temperature can also have an effect on penetration, which may be exerted on the basic vehicle if it contains temperature-sensitive components (e.g., nonionic entities or thermotherapeutic agents; Groves, 1966). Room and body temperatures can be enough to change the hydrophilic–lipophilic balance and thus possibly change the entire system. It has long been known that increasing temperature

can considerably reduce diffusional resistance and thereby increase the rate of penetration of substances. In practice, however, this effect is of no importance. Of course, skin temperature will be increased a few degrees by occlusion because of the prevention of sweating and restriction of heat radiation. However, compared to the increase in penetration achieved by the simultaneous hydration process, this effect is insignificant.

Additives aimed at accelerating penetration always attempt to enable diffusion of pharmacologically active compounds into or through the stratum corneum without damaging it and without causing undesirable systemic effects. Although attempts have been made to limit these effects, this goal has not been achieved as yet. There are numerous substances that decrease the diffusional resistance of the skin, such as propylene glycol, tensides, parotid substances such as urea, dimethyl sulfoxide (DMSO), dimethylformamide (DMF), and various other organic solvents, mostly of medium chain length. They all improve the penetration of dissolved agents, but only at the cost of the integrity of skin structure, raising the question of the degree of damage and reversibility.

If the substances have passed the stratum corneum, they also generally diffuse into the living part of the epidermis, reach the circulation, and then have systemic effects depending on the amount absorbed. Because these are often constituents of formulations, one generally expects them to have little direct influence on skin penetration. However, their amphiphilic properties allow them to form new systems with the body's constituents and even to change the physical state of water in the skin. By this means, a pathway is cleared for other hydrophilic substances to gain entry into the general circulation.

Most of a permeability enhancer (such as tenside) is bound to the stratum corneum. It is assumed that the underlying mechanism of the process involves interactions with keratin structures. Positively and negatively charged ionic groups of proteins have been suggested as binding sites for ionic substances. Ion pairs can also form. On the other hand, hydrophobic areas are present that can bind with the uncharged part of the enhancers. The total free binding energy of molecules to keratin is made up of the contributions arising from electrostatic and nonpolar interactions. Nonpolar interactions increase with chain length of the molecule. This would be the reason why predominantly anionic molecules of medium chain length exert stronger effects on the keratin structure than those of shorter chain length (Dominguez et al., 1977).

In order to reach the interior of the tightly enmeshed keratin, the molecule must overcome the elastic energy of the polypeptide matrix. The energy necessary to do this is proportional to the volume of the penetrating molecule. The larger the volume, the more difficult it will be for the molecule to approach the various binding sites of proteins in the

interior of the keratin complex. Thus, the size of the penetrating molecule is subject to certain limitations. If more molecules are present than can become bound, it is possible that a few of them will reach the living layers of the epidermis, as has been described for several anionic, mostly medium-chain enhancer molecules such as tensides. It remains unclear whether this is a consequence of pure saturation or if other interactions are involved (e.g., with structural lipids or hydrophilic materials from the intercellular lipids).

The extent to which the vehicle can affect the entire diffusion process can be shown by an example. In a four-component system of 40% oil, 40% water, and 20% of an emulsifying agent and coemulsifier, alteration of only the proportion of emulsifier to coemulsifier leads to systems of completely different colloidal-chemical structures, which can be labeled as either creams, gels, or microemulsions.

Dermal administration presents fewer logistic difficulties than oral administration. Liquids can be administered as supplied, and powders or solids can be moistened with saline to form a thick paste or slurry, or can be applied dry and moistened with saline. Solid materials (sheets or plastic, fabric, etc.) can also be administered dermally. Liquid materials or slurries are applied directly to the skin, taking care to spread the material evenly over the entire area or as much of the area as can reasonably be covered and then covering with a strip of gauze. If a large amount of material is being administered and the abdominal skin will be exposed, it is sometimes necessary to apply material to the gauze and to the skin. Dry materials are weighed out, then placed on the gauze strip, and moistened with physiological saline (generally 15 mL) so that they will adhere to the gauze. The gauze is then wrapped around the animal. This porous gauze dressing is then held in place by an additional wrapping, generally of an impervious material, to create an "occlusive" covering. This occlusion enhances penetration and prevents ingestion or evaporation of the test material.

Another recently developed approach is the use of plastic containment capsules (modified Hilltop Chambers) for administration of well-measured doses in a moisturized microenvironment (Derelanko et al., 1987).

Finally, it should be noted that for some agents (contrary to the general rule), decreasing the concentration of chemical in a vehicle may increase its apparent intrinsic toxicity.

5.3.4 Oral Formulations

The physical form of a material destined for oral administration often presents unique challenges. Liquids can be administered as supplied or diluted with an appropriate vehicle, and powders or particulates can often be dissolved or suspended in an appropriate vehicle. However, selection of an appropriate vehicle is often difficult. Water and oil (such as the vegetable oils) are used most commonly.

Materials that are not readily soluble in either water or oil can frequently be suspended in a 1% aqueous mixture of methylcellulose. Occasionally, a more concentrated methylcellulose suspension (up to 5%) may be necessary. Materials for which appropriate solutions or suspensions cannot be prepared using one of these three vehicles often present major difficulties.

Limited solubility or suspendability of a material often dictates preparation of dilute mixtures that may require large volumes to be administered. The total volume of liquid dosing solution or suspension that can be administered to a rodent is limited by the size of its stomach. However, because rats lack a gagging reflex and have no emetic mechanism, any material administered will be retained. Guidelines for maximum amounts to be administered are given in Table 5.6.

Limitations on total volume, therefore, present difficulties for materials that cannot easily be dissolved or suspended. The most dilute solutions that can be administered for a limit-type test (5000 mg kg^{-1}), using the maximum volumes shown in Table 5.6, generally are 1% for aqueous mixtures and 50% for other vehicles.

Although vehicle control animals are not required for commonly used vehicles (e.g., water, oil, methylcellulose), most regulations require that the biological properties of a vehicle be known and/or that historical data be available. Unfortunately, the best solvents are generally toxic and, thus, cannot be used as vehicles. Ethanol and acetone can be tolerated in relatively high doses but produce effects that may complicate interpretation of toxicity associated with the test material alone. It is sometimes possible to dissolve a material in a small amount of one of these vehicles and then dilute the solution in water or in oil.

Gels and resins often present problems because of their viscosity at room temperature. Warming these materials in a water bath to a temperature of up to 50°C will frequently facilitate mixing and dosing. However, it is important to ascertain that no thermal degradation occurs and that actually administered formulations be at or near body temperature.

Other possibilities for insoluble materials are to mix the desired amount of material with a small amount of the animal's diet or to use capsules. The difficulty with the diet approach is the likelihood that the animal will not consume all of the treated diet or that it may selectively not consume chunks of test material. Use of capsules, meanwhile, is labor intensive. In rare cases, if all of these approaches fail, it may not be possible to test a material by oral administration. In capsules, particle size is generally inversely related to solubility and bioavailability. However, milling of solids may adversely affect their chemical nature and/or pose issues of safety.

If necessary, the test substance should be dissolved or suspended as a suitable vehicle, preferably in water, saline, or an aqueous suspension such as 0.5% methylcellulose in

water. If a test substance cannot be dissolved or suspended in an aqueous medium to form a homogeneous dosage preparation, corn oil or another solvent can be used. The animals in the vehicle control group should receive the same volume of vehicle given to animals in the highest-dose group.

The test substance can be administered to animals at a constant concentration across all dose levels (i.e., varying the dose volume) or at a constant dose volume (i.e., varying the dose concentration). However, the investigator should be aware that the toxicity observed by administration in a constant concentration may be different from that observed when given in a constant dose volume. For instance, when a large volume of corn oil is given orally, GI motility is increased, causing diarrhea and decreasing the time available for absorption of the test substance in the GI tract. This situation is particularly true when a highly lipid-soluble chemical is tested.

If an organic solvent is used to dissolve the chemical, water should be added to reduce the dehydrating effect of the solvent within the gut lumen. The volume of water or solvent–water mixture used to dissolve the chemical should be kept low, since excess quantities may distend the stomach and cause rapid gastric emptying. In addition, large volumes of water may carry the chemical through membrane pores and increase the rate of absorption. Thus, if dose-dependent absorption is suspected, it is important that the different doses are given in the same volume of solution.

Large volumes than those detailed earlier may be given, although nonlinear kinetics seen under such circumstances may be due to solvent-induced alteration of intestinal function. The use of water-immiscible solvents such as corn oil (which are sometimes used for gavage doses) should be avoided, since it is possible that mobilization from the vehicle may be rate limiting. Magnetic stirring bars or homogenizers can be used in preparing suspensions. Sometimes a small amount of a surfactant such as Tween 80, Span 20, or Span 60 is helpful in obtaining a homogeneous suspension.

A large fraction of such a material may quickly pass through the GI tract and remain unabsorbed. Local irritation by a test substance generally decreases when the material is diluted. If the objective of the study is to establish systemic toxicity, the test substance should be administered in a constant volume to minimize GI irritation that may, in turn, affect its absorption. If, however, the objective is to assess the irritation potential of the test substance, then it should be administered undiluted.

5.3.5 Parenteral Formulations

Parenteral dose forms include aqueous, organic, and oily solutions, emulsions, suspensions, and solid forms for implantation. These parenterals need to be sterile and pyrogen-free. Additionally they are, if possible, buffered as

close to normal physiological pH and preferably are isotonic with the body fluids.

The preparation of parenteral dosage forms of approved and potential drugs for animals is the same as for humans. Turco and King (1974) provide a comprehensive review of the subject, which were written with human therapeutics in mind and contains very little that is not applicable to animals. Sterility, lack of pyrogenicity, blood compatibility, and low to no irritation at the point of injection are biological requirements; there are also a corresponding set of physico-chemical requirements.

Parenteral products are usually given to humans when an immediate effect is needed, when a patient is unable to accept medication by the oral route, or when the drug will be ineffective by the oral route. These conditions apply to animals used in safety evaluation.

Parenteral products can be easily administered to confined or restrained animals, leaving no doubt that the animal received its medication.

To be acceptable, an SC or IM formulation should cause only a minimum amount of irritation and no permanent damage to the tissues and be systemically distributed and active when administered by this route. The ideal parenteral product is an aqueous solution isotonic with the body fluids with a pH between 7 and 8. When the drug lacks sufficient aqueous solubility, a suspension may be considered; however, in most cases, the bioavailability of the drug may be affected, and encapsulation by the body at the site of injection is extremely likely. The solubility of the drug in water may be improved by the addition of cosolvents such as alcohol, propylene glycol, polyethylene glycol, dimethylacetamide, DMSO, or DMF. The resulting solution must have additional tolerance for water so that the drug will not precipitate at the site of injection when the solution is diluted by body fluids. If precipitation occurs at the site of injection, the absorption of the drug may be delayed or even completely inhibited.

Water-miscible solvents alone can be used when the drug is chemically unstable in the presence of any water. The number of solvents available for this purpose is extremely limited. The classic review of this subject was made in 1963 (Stiegeland and Noseworthy, 1963), and some 30 years later, no additional solvents are available. This is unlikely to change in the near future due to the extensive effort necessary to determine the safety of a solvent used as a vehicle. When a nonaqueous vehicle is used, one can invariably expect some degree of pain upon injection, and subsequent tissue destruction is possible. This damage may be due in part at least to the heat of solution as vehicle mixes with body fluids; it may be associated with tissues rejecting the solvent; or it may be an inherent property of the solvent.

Fixed oils of vegetable origin and their esters may be used as parenteral vehicles for some drugs, particularly steroidal hormones. While an oleaginous vehicle may delay or impair

absorption of the drug, this characteristic has been used to advantage with some drugs where a small dose is desired over a long period of time. The formulator must know which species will receive the formulation and the type of equipment used in its administration. A product intended for a dog or primate is usually given to a single animal at a time. Conventional glass or disposable syringes will be used with a 20 or 22 gauge needle, which may impede the flow of the liquid, especially when an oleaginous vehicle is used. Impedance is usually compensated for by using small animals, since the volume of injection is small and no more than one injection is normally given at one time.

The viscosity of the solution will influence its acceptability when automatic injection equipment is used. If many animals are injected at one time, a viscous solution that requires a great deal of force to eject will rapidly tire the user. When the automatic injector is refilled from a reservoir, a viscous solution will be slow to fill the volumetric chamber. The subjective aspect of measuring the ease of expelling a dose can be eliminated by constructing an apparatus that will measure the pressure needed to expel a dose. An objective means of measuring ease will allow the formulator to vary the composition of the injection and measure any improvement in injectability. For example, the addition of a wetting agent can be investigated, and, if improvement is seen, the level of use can be optimized.

A parenteral product in a multidose vial must contain a preservative to protect the contents of the vial against contamination during repeated withdrawal of dose aliquots.

5.4 DOSING CALCULATIONS

One of the first things a new technician (or graduate student) must learn is how to calculate dose. Generally, administered doses in systemic toxicity studies are based on the body weight of the animal (expressed as either weight or volume—for liquids—of the test substance per kilogram of body weight of the animal), although some would maintain that surface area may be a more appropriate basis on which to gauge individual dose. The weight (or dose) of the test substance is often expressed in milligrams or grams of active ingredient if the test substance is not pure (i.e., if it is not 100% active ingredient).

Ideally, only the 100% pure sample should be tested; however, impurity-free samples are difficult to obtain, and preparation of formulations (as previously discussed) is frequently essential. The toxicity of impurities or formulation components should be examined separately if the investigator feels that they may contribute significantly to the toxicity of the test substance.

If the test substance contains only 75% active ingredient and the investigator chooses a constant dose volume of 10 mL kg⁻¹ bw across all dose levels, it will be more

convenient to prepare a stock solution such that when 10 mL kg⁻¹ of this stock solution is given to the animal, the dose will be the desired one (say, 500 mg kg⁻¹ of active ingredient). The concentration of this stock solution would be (500 mg/10 mL)/0.75 = 66.7 mg of the test substance per milliliter of diluent.

Aliquots of the test substance for other dose levels can then be prepared by dilution of the stock solution. For example, the solution concentration for a 250 mg kg⁻¹ dose level is (200 mg/10 mL)/0.75 = 26.7 mg of the test substance per milliliter of diluent.

This solution can be prepared by diluting the stock solution 25 times; that is, for each mL of the 26.7 mg mL⁻¹ solution to be prepared,

$$\frac{(26.7 \text{ mg mL}^{-1})(1 \text{ mL})}{66.7 \text{ mg mL}^{-1}} = 0.400 \text{ mL of the stock solution}$$

This amount should be diluted to a final volume of 1 mL with the vehicle.

The other way to express a relative dose in animals or humans is to do so in terms of body surface area. There are many reasons for believing that the surface area approach is more accurate for relating doses between species (Schmidt-Nielson, 1984)—and especially between test animals and humans—but this is still a less common approach in safety assessment, although it is the currently accepted norm in a couple of areas—carcinogenesis and chemotherapy, for example.

5.5 CALCULATING MATERIAL REQUIREMENTS

One of the essential basic skills for the efficient design and conduct of safety assessment studies is to be able to accurately project compound requirements for the conduct of a study. In theory, this simply requires plugging numbers into a formula such as

$$(A \times B \times C \times D) 1.1 = \text{total compound requirement}$$

where

A = number of animals in each study group

B = the *sum* of doses of the dose groups (such as 0.1 + 0.3 + 1.0 mg kg⁻¹ = 1.4 mg kg⁻¹)

C = the number of doses to be delivered (usually the length of the study in days)

D = the average body weight per animal (assuming dosing is done on a per body weight basis)

1.1 = a safety factor (in effect, 10%) to allow for spillage, wall loss on glassware and containers, and computation errors.

TABLE 5.9 Standardized Total Compound Requirements for Rodent Diet Studies^{a,b}

| Length of Study (Days) | Total Compound Requirement (g) per Dose (mg kg ⁻¹ day ⁻¹) | | | | | |
|------------------------|--|------|------|------|------------------|------------------|
| | 1 | 3 | 10 | 30 | 100 | 300 |
| Rat^c | | | | | | |
| 2 weeks | 0.2 | 0.4 | 1.2 | 4 | 10.6 | 32 |
| 4 weeks | 0.43 | 0.7 | 2.5 | 7.5 | 25 | 75 |
| 13 weeks | 0.8 | 2.6 | 8.5 | 25.5 | 85 | 260 |
| 52 weeks | 7 | 21 | 70 | 210 | 0.7 ^d | 2.1 ^d |
| 2 years | 15 | 45 | 150 | 450 | 1.5 ^d | 4.5 ^d |
| Mouse | | | | | | |
| 2 weeks | 0.03 | 0.06 | 0.22 | 0.65 | 2.2 | 6.4 |
| 4 weeks | 0.08 | 0.14 | 0.8 | 1.4 | 8 | 14 |
| 13 weeks | 0.14 | 0.42 | 1.4 | 4.2 | 14 | 42 |
| 18 months | 0.85 | 2.5 | 8.5 | 25 | 85 | 250 |

^a Based on 10 animals per sex per group for the length of the study that are 6–8 weeks old at study initiation. Animals are weighed to determine body weights.

^b See Katdare and Chaubal (2006).

^c Sprague–Dawley rats (body weights and compound requirements for Fischers would be less).

^d In kilograms.

As an example of this approach, consider a study that calls for 10 dogs/sex/group ($A = 10 \times 2 = 20$) to receive 0, 10, 50, or 150 mg kg⁻¹ day⁻¹ ($B = 10 + 50 + 150 = 210$ mg kg⁻¹) for 30 days ($C = 30$). On average, the dogs of the age range used weigh 10 kg ($D = 10$ kg). Our compound need is then $(20 \times 210 \text{ mg kg}^{-1} \times 30 \times 10 \text{ kg}) 1.1 = 1.386$ kg.

The real-life situation is a bit more complicated, since animal weights change over time, diet studies have doses dependent on daily diet consumption, the material may be a salt but dosage should be calculated on the basis of the parent compound, and not all animals may be carried through the entire study.

For rats and mice (where weight change is most dramatic and diet studies most common), Table 5.9 presents some reliable planning values for compound requirements during diet studies.

5.6 EXCIPIENTS

Excipients are usually thought of as being inert substances (such as gum arabic and starch) that form the vehicle or bulk of the dosage form of a drug. They are, of course, both much more complicated than this and not necessarily inert. A better definition would be that of the USP and National Formulary (USP, 2014), which defined excipients as any component, other than the active substances (i.e., drug substances or DS) intentionally added to the formulation of a dosage form. These substances serve a wide variety of purposes: enhancing

stability, adding bulking, increasing and/or controlling absorption, providing or masking flavor, coloring, and serving as a lubricant in the manufacturing process. They are, in fact, essential for the production and delivery of marketed DP. As will soon be made clear, they are regulated both directly and as part of the DP. For the pharmaceutical manufacturers, the use of established and accepted excipients (such as can be found in Smolinske (1992) or APA (1994)—though these lists are not complete) is much preferred. However, both pharmaceutical manufacturers and the companies which supply excipients must from time to time utilize (and therefore develop, evaluate for safety, and get approved) new excipients.

In the last 10 years, the use of nanosuspensions as components of formulations has presented a new set of opportunities—and potential problems (Rabinow, 2004).

While for the purpose of nonclinical formulation, our concerns are generally limited to vehicles, their formulation components can be important. Table 5.7 lists examples of this.

5.6.1 Regulation of Excipients

Table 5.10 lists the relevant sections of CFR 21 which govern excipients. Under Section 201(g)(1) of the Federal Food, Drug, and Cosmetic Act (FD&C Act; 1), the term *drug* is defined as:

(A) articles recognized in the official *United States Pharmacopeia*, official *Homeopathic Pharmacopeia of the United States*, or official *National Formulary*, or any supplement to any of them; and (B) Articles intended for use in the diagnosis, cure, mitigation, treatment, or prevention of disease in man or other animals; and (C) Articles (other than food) intended to affect the structure of any function of the body of man or other animals; and (D) Articles intended for use as a component of any articles specified in clause (A), (B), or (C).

An excipient meets the definitions as listed in (A) and (D) above.

In 21 CFR § 210.3(b)(8)(2), an “inactive ingredient means any component other than an active ingredient.” According to the CFR, the term inactive ingredient includes materials in addition to excipients. 21 CFR § 201.117 states the following:

Inactive ingredients: A harmless drug that is ordinarily used as an inactive ingredient, such as a coloring, emulsifier, excipient, flavoring, lubricant, preservative, or solvent in the preparation of other drugs shall be exempt from Section 502(f)(1) of the Act. This exemption shall not apply to any substance intended for a use which results in the preparation of a new drug, unless an approved new-drug application provides for such use.

TABLE 5.10 US Code of Federal Register References to Excipients

| Subject | Reference | Content |
|--|------------------------------|---|
| General | 21 CFR § 210.3(b)(8) | Definitions |
| | 21 CFR § 201.117 | Inactive ingredients |
| | 21 CFR § 210.3(b)(3) | Definitions |
| Over-the-counter drug products | 21 CFR § 330.1(e) | General conditions for general recognition as safe, effective, and not misbranded |
| | 21 CFR § 328 | Over-the-counter drug products intended for oral ingestion that contain alcohol |
| Drug Master Files | 21 CFR § 314.420 | Drug master files |
| Investigational New Drug (IND) Application | 21 CFR § 312.23(a)(7) | IND content and format |
| New Drug Application | 21 CFR § 312.31 | Information amendments |
| | 21 CFR § 314.50(d)(1)(ii)(a) | Content and format of an application |
| | 21 CFR § 314.70 | Supplements and other changes to an approved application |
| Abbreviated New Drug Application | 21 CFR § 314.94(a)(9) | Content and format of an abbreviated application |
| | 21 CFR § 314.127 | Refusal to approve an abbreviated new drug application |
| | 21 CFR § 314.127(a)(8) | Refusal to approve an abbreviated new drug application |
| Current Good Manufacturing Practice | 21 CFR § 211.84(d) | Testing an approval or rejection of components, drug product containers, and closures |
| | 21 CFR § 211.165 | Testing and release for distribution |
| | 21 CFR § 211.180(b) | General requirements |
| | 21 CFR § 211.80 | General requirements |
| | 21 CFR § 211.137 | Expiration dating |
| Listing of drugs | 21 CFR § 207 | Registration of procedures of drugs and listing of drugs in commercial distribution |
| | 21 CFR § 207.31(b) | Additional drug listing information |
| | 21 CFR § 207.10(e) | Exceptions for domestic establishments |
| Labeling | 21 CFR § 201.100(b)(5) | Prescription drugs for human use |
| | 21 CFR § 201.20 | Declaration of presence of FD&C Yellow No. 5 and/or FD&C Yellow No. 6 in certain drugs for human use |
| | 21 CFR § 201.21 | Declaration of presence of phenylalanine as a component of aspartame in over-the-counter and prescription drugs for human use |
| | 21 CFR § 201.22 | Prescription drugs containing sulfites; required warning statements |

Excipients also meet the definition of component in the Good Manufacturing Practice (GMP) regulations in 21 CFR § 210.3(b)(3): “Component means any ingredient intended for use in the manufacture of a drug product, including those that may not appear in such drug product.”

The *NF* Admissions Policy in the *USP 30/National Formulary 25* defines the word *excipient* (3): “An excipient is any component other than the active substance(s), intentionally added to the formulation of a dosage form. It is not defined as an inert commodity or an inert component of a dosage form.”

Similar to all other drugs, excipients must comply with the adulteration and misbranding provisions of the FD&C Act. Under Section 501(a), an excipient shall be deemed to be adulterated if it consists in whole or in part of any filthy, putrid, or decomposed substance, or if it has been prepared, packed, or held under unsanitary conditions whereby it may have been contaminated with filth, or whereby it may have been rendered injurious to health. An excipient is adulterated

if the methods used in, or the facilities or controls used for its manufacture, processing, packing, or holding do not conform to or are not operated or administered in conformity with current GMP to assure that such drug meets the requirements of the act as to safety and has the identity and strength and meets the quality and purity characteristics which it purports or is represented to possess. In addition, under Section 501(b), an excipient shall be deemed to be adulterated if it purports to be or is represented as a drug the name of which is recognized in an official compendium, and its strength differs from, or its quality or purity falls below, the standards set forth in such compendium.

In 2005, the US FDA promulgated new guidance on the selection and use of excipients in nonclinical and clinical studies. US FDA compliance officials require the use of inactive ingredients that meet compendial standards when standards exist and either have previous use in US FDA-approved pharmaceuticals or that they be qualified as “novel” excipients (with studies as summarized in Table 5.11)

TABLE 5.11 Summary of Toxicological Studies Recommended for New Pharmaceutical Excipients Based on Route of Exposure

| Tests | Oral | Mucosal | Transdermal | Dermal/topical | Parenteral | Inhalation/intranasal | Ocular |
|--|------|---------|-------------|----------------|------------|-----------------------|--------|
| Appendix 1—base set | R | R | R | R | R | R | R |
| Acute oral toxicity | R | R | R | R | R | R | R |
| Acute dermal toxicity | C | C | C | C | C | R | C |
| Acute inhalation toxicity | R | R | R | R | R | R | R |
| Eye irritation | R | R | R | R | R | R | R |
| Skin irritation | R | R | R | R | R | R | R |
| Skin sensitization | R | R | R | R | R | R | R |
| Acute parenteral toxicity | — | — | — | — | R | — | — |
| Application site evaluation | — | — | R | R | R | R | — |
| Pulmonary sensitization | — | — | — | — | — | R | — |
| Phototoxicity/photoallergy | — | — | R | R | — | — | — |
| Ames test | R | R | R | R | R | R | R |
| Micronucleus test | R | R | R | R | R | R | R |
| ADME-intended route | R | R | R | R | R | R | R |
| 28-day toxicity (2 species)- intended route | R | R | R | R | R | R | R |
| Appendix 2 | | | | | | | |
| 90-day toxicity (most appropriate species) | R | R | R | R | R | R | R |
| Developmental toxicity (rat and rabbit) | R | R | R | R | R | R | R |
| Additional assays | C | C | C | C | C | C | C |
| Genotoxicity assays | R | R | R | R | R | R | R |
| Appendix 3 | | | | | | | |
| Chronic toxicity (rodent, nonrodent) | C | C | C | C | C | C | C |
| Photocarcinogenicity | — | — | C | C | — | — | — |
| Carcinogenicity | C | C | C | C | C | C | — |

C, conditionally required; R, required.

(FDA, 2005). The FDA/CDER maintains an Inactive Ingredient Committee whose charter includes the evaluation of the safety of inactive ingredients on an as-needed basis, preparation of recommendations concerning the types of data needed for excipients to be declared safe for inclusion in a drug product, and other related functions.

From a regulatory standpoint, the FDA's concern regarding safety involves the toxicity, degradants, and impurities of excipients, as discussed in other chapters in this book. In addition, other chapters of this book address types of toxicity concerns, toxicity testing strategies, and exposure and risk assessment of excipients.

Excipients must be safe for their intended use. Under 21 CFR § 330.1(e), over-the-counter (OTC) human drugs that are generally recognized as safe and effective and not misbranded may only contain inactive ingredients if they are suitable and if the amounts administered are safe and do not interfere with the effectiveness of the drug or with required tests or assays. Color additives may be used in accordance with the provisions of the FD&C Act and the regulations of 21 CFR Parts 70–82. The FDA proposed that to make it clear that, to be considered as suitable within the meaning of 21 CFR § 330.1(e), each inactive ingredient in an OTC human drug product should perform a specific function (5). The

proposed regulation defined *safe and suitable* to mean that the inactive ingredient meets various conditions as mentioned in the foregoing. OTC drug manufacturers are responsible for assuring that these conditions are met. There is no formal approval mechanism.

In the United States, the safety and suitability of excipients used in new drugs are considered as part of the New Drug Application (NDA) process. There is no separate and independent review and approval system for excipients. There are no specific regulations or guidelines that specify the requirements needed to gain approval of a new drug that contains a new excipient. Generally, pharmaceutical companies choose excipients that previously have been approved for commercial use in other NDAs. The FDA's *Inactive Ingredient Guide* (6), discussed later in this chapter, contains a listing of inactive ingredients present in approved drug products. There is currently no way of gaining a listing for an excipient in the guide independent of the NDA route. The FDA reviews the status of an excipient in food as information to support its use in drug products. Factors relative to the use of an excipient, such as dosing regimen and route of administration, are also reviewed. Advances in excipient technology and drug dosage from technology have created a need for a separate regulatory approval process for new

excipients. The *USP* published IPEC's Excipient Safety Evaluation Guidelines as Information Chapter 1074, Excipient Biological Safety Evaluation Guideline.

Information on existing or new excipients can be described and provided to the FDA in an NDA directly. Alternatively, the manufacturers of excipients may prepare and submit type IV Drug Master Files (DMF) to support the use of an excipient in one or more NDAs. The DMFs are discussed in FDA's regulations under 21 CFR § 314.420 and the FDA-issued Guidance for Drug Master Files (8). When authorized by the DMF submitter (i.e., the excipient manufacturer) and cross-referenced by an NDA submitter, the FDA reviews the DMF to make determinations on the safety, manufacture, and quality of the excipient use in the new drug that is the subject of the then pending NDA. The DMF becomes active when reviewed in conjunction with the review and approval of an NDA.

The *USP/NF* provides a listing of excipients by categories in a table according to the function of the excipient in a dosage form, such as tablet binder, disintegrant, and such. An excellent reference for excipient information is the *Handbook of Pharmaceutical Excipients* (Rowe et al., 2012). Additionally, Gad et al. (2016) provide an excellent and extensive database of nonclinical formulation components and either acceptable maximum usage levels by species route and duration of study.

Excipients have historically not been subjected to extensive safety testing because they have been considered a priori to be biologically inactive and therefore nontoxic. Many, if not most, excipients used are approved food ingredients, the safety of which has been assured by a documented history of safe use or appropriate animal testing. Some of the excipients are generally recognized as safe (GRAS) food ingredients. The excipient is an integral component of the finished drug preparation and, in most countries, is evaluated as part of this preparation. There has been no apparent need to develop specific guidelines for the safety evaluation of excipients, and most developed countries do not have specific guidelines. However, as drug development has become more complex and/or new dosage forms have developed, improved drug bioavailability has become more important. It was noted that the available excipients were often inadequate, new pharmaceutical excipients specifically designed to meet the challenges of delivering new drugs were needed, and these are being developed. The proper safety evaluation of new excipients has now become an integral part of drug safety evaluation.

In the absence of official regulatory guidelines, the Safety Committees of the IPEC in the United States, Europe, and Japan developed guidelines for the proper safety evaluation of new pharmaceutical excipients (IPEC, 2012). The committees critically evaluated guidelines for the safety evaluation of food ingredients, cosmetics, and other products, as well as textbooks and other appropriate materials. Before

initiating a safety evaluation program for a new pharmaceutical excipient, it is advisable to address the following:

1. Chemical and physical properties and functional characterization of the test material (Hawley, 1971)
2. Analytical methods that are sensitive and specific for the test material and that can be used to analyze for the test material in animal food used in the feeding studies or in the vehicle used for other studies
3. Available biological, toxicological, and pharmacological information on the test material and related materials (which involves a thorough search of the scientific literature)
4. Intended conditions of use, including reasonable estimates of exposure
5. Potentially sensitive segments of the population

As discussed in Chapters 1 and 3, a comprehensive and critical search of the scientific literature on the test material and related materials is essential before the start of any testing program.

As pharmaceutical excipients are assumed to be biologically nonreactive, dose–response relations cannot always be established. An acceptable alternative is to use a maximum attainable or maximum feasible dose. This is the highest dose possible that will not compromise the nutritional or health status of the animal. Table 5.12 summarizes the maximum or limit doses for various types of studies by different routes of exposure. For example, 2000 mg kg⁻¹ bw of an orally administered test material is the maximum dose recommended for a testing strategy that has been developed for new pharmaceutical excipients that takes into consideration the physicochemical nature of the product and the potential route(s) and duration of exposures, both through its intended use as part of a drug product and through workplace exposure during manufacturing. The number and types of studies recommended in this tiered approach are based on the duration and routes of potential human exposure. Thus, the longer the exposure to the new pharmaceutical excipient, the more studies are necessary to assure safety. Table 5.11 summarizes the entire set of toxicological studies recommended for new pharmaceutical excipients (Weiner and Katkoskie, 1999; IPEC, 2012).

Tests have been outlined for each exposure category to assure safe use of the time period designated. The tests for each exposure category assure the safe use of the new pharmaceutical excipient of the time frame specified for the specific exposure category. Additional tests are required for longer exposure times.

The base set required for all excipients is detailed in Table 5.13. These are sufficient, however, only for those excipients intended for use for up to 2 weeks in man.

TABLE 5.12 Limit Doses for Toxicological Studies

| Nature of Test | Species | Limit Dose ^a |
|--|---------------------|--|
| Acute oral | Rodent | 2000 mg kg ⁻¹ bw |
| Acute dermal | Rabbit | 2000 mg kg ⁻¹ bw |
| | Rat | |
| Acute inhalation ^b | Rat | 5 mg L ⁻¹ air for 4 h or maximum attainable level under conditions of study |
| Dermal irritation | Rabbit | 0.5 mL liquid 0.5 g solid |
| Eye irritation | Rabbit | 0.1 mL liquid 100 mg solid |
| 14-day/28-day oral repeated dosing; 90-day subchronic | Rodent, nonrodent | 1000 mg kg ⁻¹ bw day ⁻¹ |
| 14-day/28-day oral repeated dosing; 90-day subchronic | Rat, rabbit | 1000 mg kg ⁻¹ bw day ⁻¹ |
| Chronic toxicity, carcinogenicity | Rats, mice | 5% maximum dietary concentration for nonnutrients |
| Reproduction | Rats | 1000 mg kg ⁻¹ bw day ⁻¹ |
| Developmental toxicity (teratology) | Mice, rats, rabbits | 1000 mg kg ⁻¹ bw day ⁻¹ |

^a mg kg⁻¹ bw, milligrams of test material dosed per kilogram of body weight to the test species.

^b Acute inhalation guidelines that indicate this limit dose are US Environmental Protection Agency Toxic Substance Health Effect Test Guidelines, Oct., 1984; (PB82-232984) Acute Inhalation Toxicity Study; the OECD Guidelines of the Testing of Chemicals, Vol. 2, Section 4; Health Effects, 403, Acute Inhalation Toxicity Study, May 12, 1982; and the *Official Journal of the European Communities*, L383A, Vol. 35, Dec. 29, 1992, Part B.2 (adapted from Weiner and Katkoskie (1999)).

If exposure to the new pharmaceutical excipient is expected to occur for longer than 2 but no more than 6 weeks, additional toxicological studies are required, as shown in Table 5.14. The longer the expected human exposure, the more extensive will be the toxicological studies to assure safety. A tiered approach assures that those tests necessary to ensure safety for the expected duration of human exposure are conducted. Thus, to assure safe use for greater than 2 weeks, but no more than 6 weeks in humans, subchronic toxicity and developmental toxicity studies are required. To assure safe use for greater than six continuous weeks, chronic or oncogenicity studies are conditionally required, as per Table 5.14. This means long-term studies should be considered for prolonged human exposures, but may not be absolutely required. A thorough scientific review of the data generated in the base set and Appendix 2 studies should be undertaken. From a critical evaluation by a competent toxicologist, the results of the physicochemical properties of the test material, the 28-day and 90-day tests, the ADME–PK acute and repeated-dose tests, and the

TABLE 5.13 Base Set Studies for a Single Dose up to 2-Week Exposure in Humans

| Test | Purpose |
|--------------------------------|---|
| Acute oral toxicity | To determine the potential acute toxicity/ lethality following a single oral dose |
| Acute dermal toxicity | To determine the potential acute toxicity/ lethality following a single dermal dose |
| Acute inhalation toxicity | To determine the potential acute toxicity/ lethality following a single 4 h inhalation exposure to a test atmosphere containing the new pharmaceutical excipient (aerosol, vapor, or particles) |
| Eye irritation | To determine the potential to produce acute irritation or damage to the eye |
| Skin irritation | To determine the potential to produce acute irritation or damage to the skin |
| Skin sensitization | To determine the potential to induce skin sensitization reactions |
| Ames test | To evaluate potential mutagenic activity in a bacterial reverse mutation system with and without metabolic activation |
| Micronucleus test | To evaluate the clastogenic activity in mice using polychromatic erythrocytes |
| ADME-intended route | To determine the extent of absorption, distribution, metabolism, and excretion by the intended route of exposure following a single dose and repeated doses |
| 28-day toxicity-intended route | To assess the repeated-dose toxicity in male and female animals of two species following dosing for 28 days by the intended route of exposure |

TABLE 5.14 Appendix 3 Studies for Repeated Chronic Exposure in Humans

| Test | Purpose |
|-----------------------------|--|
| Chronic toxicity | To assess the toxicity following chronic (lifetime) exposure by the route of intended exposure |
| Oncogenicity | To assess the potential to induce tumors by the intended route of exposure |
| One-generation reproduction | To assess the potential reproductive and developmental toxicity in males and females by the intended route of exposure |

developmental toxicity test(s), a final determination can be made on the value of chronic toxicity or oncogenicity studies.

For example, if no toxicity is observed at a limit dose of 1000 mg kg⁻¹ bw per day following the 90-day toxicity study, no genotoxicity was found, and the ADME–PK profile indicates that the material is not absorbed and is completely excreted unchanged in the feces, then it is likely that a chronic study is not necessary. The decision to conduct

chronic studies should be determined on a case-by-case basis using scientific judgment. It will be interesting to observe how this scheme may change in light of ICH.

REFERENCES

- Adeyeye MC, Brittain HG. (Eds.) (2008) Preformulation in Solid Dosage Form Development. In: *Drugs and the Pharmaceutical Sciences*, Vol. 178. Informa Healthcare USA, New York.
- Attman PL, Dittmer DS. (Eds.) (1971) *Respiration and Circulation*. FASEB, Bethesda, MD, pp. 56–59.
- Avis KE. (1985) Parenteral preparations. In: *Remington's Pharmaceutical Sciences* (Gennaro AR, Ed.). Mack Publishing Company, Easton, PA, pp. 1518–1541.
- Baker HJ, Lindsey JR, Weisbroth SH. (1979) *The Laboratory Rat*, Vol. 1. Academic Press, New York, pp. 411–412.
- Ballard BE. (1968) Biopharmaceutical considerations in subcutaneous and intramuscular drug administration. *J. Pharm. Sci.* 57:357–378.
- Barry BW. (1983) *Dermatological Formulations*. Marcel Dekker, New York.
- Bates TR, Gibaldi M. (1970) Gastrointestinal absorption of drugs. In: *Current Concepts in the Pharmaceutical Sciences: Biopharmaceutics* (Swarbrick J, Ed.). Lea & Febiger, Philadelphia, pp. 58–99.
- Blank HI, Scheuplein RJ. (1964) The epidermal barrier. In: *Progress in the Biological Sciences in Relation to Dermatology* (Rook A and Champion RH, Eds.), Vol. 2. Cambridge University Press, Cambridge.
- Boersen N, Lee T, Hui H-W. (2014) Development of preclinical formulations for toxicology studies. In: *A Comprehensive Guide to Toxicology in Preclinical Drug Development* (Faqi AS, Ed.). Academic Press, San Diego, pp. 69–86.
- Borowitz JL, Moore PF, Yim GKW, Miya TS. (1971) Mechanism of enhanced drug effects produced by dilution of the oral dose. *Toxicol. Appl. Pharmacol.* 19:164–168.
- Brandau R, Lippold BH. (1982) *Dermal and Transdermal Absorption*. Wissenschaftliche, Verlagsgesellschaft mbH, Stuttgart.
- Bronaugh RL, Maibach HI. (1985) *Percutaneous Absorption*. Marcel Dekker, New York.
- Brown VK. (1980) *Acute Toxicity in Theory and Practice*. John Wiley & Sons, New York.
- Brown VK, Muir CMC. (1971) Some factors affecting the acute toxicity of pesticides to mammals when absorbed through skin and eyes. *Int. Pest. Control.* 13:16–21.
- Calvary HO, Klump TG. (1939) The toxicity for human beings for diethylene glycol with sulfanilamide. *Southern Med. J.* 32:1105.
- Derelanko MJ, Gad SC, Gavigan FA, Babich PC, Rinehart WE. (1987) Toxicity of hydroxylamine sulfate following dermal exposure: variability with exposure method and species. *Fund. Appl. Toxicol.* 8(3):583–594.
- Dominguez JG, Parra JL, Infante MR, Pelejero CM, Balaguer F, Sastre T. (1977) A new approach to the theory of adsorption and permeability of surfactants on keratinic proteins: the specific behavior of certain hydrophobic chains. *J. Soc. Cosmet. Chem.* 28:165–182.
- Eckstein P, Jackson MCN, Millman N, Sobrero AJ. (1969) Comparison of vaginal tolerance tests and spermicidal preparations in rabbits and monkeys. *J. Reprod. Fertil.* 20:85–93.
- FDA (May 2005) Guidance for Industry: Nonclinical Studies for the Safety Evaluation of Pharmaceutical Excipients. Center for Drug Evaluation and Research and Center for Biologics Evaluation and Research, Rockville, MD.
- Fitzgerald JM, Boyd VF, Manus AG. (1983) Formulations of insoluble and immiscible test agents in liquid vehicles for toxicity testing. In: *Chemistry for Toxicity Testing* (Jameson CW and Walters DB, Eds.). Butterworth, Boston, pp. 83–90.
- Gad SC. (2007) *Target Organ Toxicity: The GI Tract*. Taylor & Francis, Philadelphia, PA.
- Gad SC. (Ed.) (2008) *Preclinical Development Handbook: ADME and Biopharmaceutical Properties*. John Wiley & Sons, Hoboken, NJ.
- Gad SC, Smith AC, Cramp AL, Gavigan FA, Derelanko MJ. (1984) Innovative designs and practices for acute systemic toxicity studies. *Drug Chem. Toxicol.* 7(5):423–434.
- Gad SC, Spainhour CB, Shoemaker C, Stackhouse Pallman DR, Stricker-Krongrad A, Downing PA, Seals RE, Eagle LA, Polhamus K, Daly J. (2016) Tolerable levels of nonclinical vehicles and formulations used in studies by multiple routes in multiple species with notes on methods to improve utility. *Int. J. Toxicol.*, 35(2):95–178.
- Gaines TB. (1960) The acute toxicity of pesticides to rats. *Toxicol. Appl. Pharmacol.* 2:88–99.
- Garramone JP. (1986) *Vascular Access Port Model SLA Users Manual*. Norfolk Medical Products, Skokie, IL.
- Gibaldi M. (1976) Biopharmaceutics. In: *The Theory and Practice of Industrial Pharmacy* (Lachman L, Lieberman HA and Kanig JL, Eds.), 2nd Ed. Lea & Febiger, Philadelphia, PA.
- Gibaldi M, Feldman S. (1970) Mechanisms of surfactant effects on drug absorption. *J. Pharm. Sci.* 59:579.
- Gibson M. (Ed.) (2009) Pharmaceutical Preformulation and Formulation. In: *Drugs and the Pharmaceutical Sciences*, 2nd Ed, Vol. 199. Informa Healthcare USA, New York.
- Gray JE. (1978) Pathological evaluation of injection injury. In: *Sustained and Controlled Release Drug Delivery System* (Robinson J, Ed.). Marcel Dekker, New York, p. 35105.
- Groves M. (1966) The influence of formulation upon the activity of thermotherapeutic agents. *Reps. Progr. Appl. Chem.* 12:51–151.
- Hagenbusch OE. (1937) Elixir of sulfanilamide messengill. *J. Am. Med. Assoc.* 109:1531.
- Hardee GE, Baggot JD. (Eds.) (1998) Development and formulation of veterinary dosage forms. In: *Drugs and the Pharmaceutical Sciences*, 2nd Ed, Vol. 88. CRC Press, New York.
- Hawley GG. (1971) *The Condensed Chemical Dictionary*. Van Nostrand Reinhold, New York.
- Hogben CAM, Tocco DJ, Brodie BB, Schranker LS. (1959) On the mechanism of intestinal absorption of drugs. *J. Pharmacol. Exp. Ther.* 125(4):275–282.
- Holick HGO. (1937) Glycerine, ethylene glycol, propylene glycol and diethylene glycol. *J. Am. Med. Assoc.* 109:1517.

- Holmstedt B. (1959) Pharmacology of organophosphorus cholinesterase inhibitors. *Pharmacol. Rev.* 11:567–688.
- Houston JB, Upshall DG, Bridges JW. (1974) The re-evaluation of the importance of partition coefficients in the gastrointestinal absorption of nutrients. *J. Pharmacol. Exp. Ther.* 189:244–254.
- IPEC. (2012) *The Joint IPEC—PQG Guide for Pharmaceutical Excipients*. Pharmaceutical Quality Group, London.
- Katdare A, Chaubal MV. (2006) *Excipient Development for Pharmaceutical Biotechnology and Drug Delivery Systems*. Informa Healthcare, New York.
- Kemsley J. (2014) Eyes on excipients: inactive but essential drug ingredients are getting checked out with modern rigor to ensure their quality. *Chemical & Engineering News*, 92(5):9–11. Available at: <http://cen.acs.org/articles/92/i5/Eyes-Excipients.html> (accessed April 7, 2015).
- Kerberle H. (1971) Physicochemical factors of drugs affecting absorption, distribution and excretion. *Acta Pharmacol. Toxicol.* 29(Suppl 3):30–47.
- Kondrizer AA, Mayer WH, Zviblis P. (1959) Removal of sarin from skin and eyes. *Arch. Indust. Health.* 20:50–52.
- Kruger S, Greve DW, Schueler FW. (1962) The absorption of fluid from the peritoneal cavity. *Arch. Int. Pharmacodyn.* 137:173–178.
- Lien E, Koda RT, Tong GL. (1971) Buccal and percutaneous absorptions. *Drug Intell. Clin. Pharm.* 5:38–41.
- Lipinski CA. (2004) Lead- and drug-like compounds: the rule-of-five revolution. *Drug Discov. Today Technol.* 1(4):337–341.
- Lipinski CA, Lombardo F, Dominy BW, Feeney PJ. (2001) Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. *Adv. Drug Deliv. Rev.* 46(1–3):3–26.
- Liu R. (Ed.) (2008) *Water-Insoluble Drug Formulation*, 2nd Ed. CRC Press, Boca Raton, FL.
- Lukas G, Brindle SD, Greengard P. (1971) The route of absorption of intraperitoneally administered compounds. *J. Pharmacol. Exp. Ther.* 178:562566.
- Marzulli FN. (1962) Barriers to skin penetration. *J. Invest. Derm.* 39:387–393.
- McConnell RF. (1974) Special requirements for testing spermicides. *Acta Endocrinol. Suppl. (Copenh.)*. 185:373–386.
- McMartin C, Hutchinson LEF, Hyde R, Peters GE. (1987) Analysis of structural requirements for the absorption of drugs and macromolecules from the nasal cavity. *J. Pharm. Sci.* 76:535–540.
- Medved LL, Kundiev YI. (1964) On the methods of study of penetration of chemical substances through the intact skin (Russian). *Gig. Sanit.* 29:71–76.
- Melnick RL, Jameson CW, Goehl TJ, Kuhn GO. (1987) Application of microencapsulation for toxicology studies. *Fund. Appl. Toxicol.* 11:42531.
- Monkhouse DC, Rhodes CT. (1998) *Drug Products for Clinical Trials*. Marcel Dekker, New York.
- Neervannan S. (2006) Preclinical formulations for discovery and toxicology: physicochemical challenges. *Expert Opin. Drug Metab. Toxicol.* 2(5):715–731.
- Niazi SK. (2007) *Handbook of Preformulation: Chemical, Biological, and Botanical Drugs*. Informa Healthcare USA, New York.
- O'Reilly WJ. (1972) Pharmacokinetics in drug metabolism and toxicology. *Can. J. Pharm. Sci.* 7:6–77.
- Palucki M, Higgins JD, Kwong E, Templeton AC. (2010) Strategies at the interface of drug discovery and development: early optimization of the solid state phase and preclinical toxicology formulation for potential drug candidates. *J. Med. Chem.* 53(16):5897–5905.
- Powers DH, Fox C. (1957) A study of the effect of cosmetic ingredients, creams and lotions on the rate of moisture loss. *Proc. Sci. Sect. Toilet Goods Assoc.* 28:21–26.
- Rabinow BE. (2004) Nanosuspensions in drug delivery. *Nat. Rev. Drug Discov.* 3(9):785–796.
- Racy I. (1989) *Drug Formulation*. John Wiley & Sons, New York.
- Rieger MM, Deems DE. (1974) Skin moisturizers. II. The effects of cosmetic ingredients on human stratum corneum. *J. Soc. Cosmet. Chem.* 25:253.
- Ritschel WA, Siegel EG, Ring PE. (1974) Biopharmaceutical factors influencing LDP: Pt. 1 Viscosity. *Arzneim.-Forsch.* 24:907–910.
- Rowe RC, Sheskey PJ, Cook WG, Fenton ME. (2012) *Handbook of Pharmaceutical Excipients*, 7th Ed. American Pharmaceutical Association, Washington, DC.
- Scheuplein RJ. (1965) Mechanism of percutaneous absorption. (i) Routes of penetration and the influence of solubility. *J. Invest. Derm.* 45:334–346.
- Scheuplein RJ. (1967) Mechanism of percutaneous absorption. (ii) Transient diffusion and the relative importance of various routes of skin penetration. *J. Invest. Derm.* 48:79–88.
- Schmidt-Nielsen K. (1984) *Scaling: Why Is Animal Size So Important?* Cambridge University Press, New York.
- Schou J. (1971) Subcutaneous and intramuscular injection of drugs. In: (Brodie BB and Gillette JR, Eds.) *Handbook of Experimental Pharmacology*. Springer-Verlag, Berlin, Ch. 4.
- Schranker LS. (1960) On the mechanism of absorption of drugs from the gastrointestinal tract. *J. Med. Pharm. Chem.* 2:343–359.
- Schranker LS, Shore PA, Brodie BB, Hogben CAM. (1957) Absorption of drugs from the stomach, I. The rat. *J. Pharmacol. Exp. Ther.* 120:528.
- Schwartz ML, Meyer MB, Covino BG, Narang RM. (1974) Antiarrhythmic effectiveness of intramuscular lidocaine: influence of different injection sites. *J. Clin. Pharmacol.* 14:77–83.
- Shaffer CB, West B. (1960) The acute and subacute toxicity of technical *o,o*-diethyl *s*-2-diethyl-aminoethyl phosphorothioate hydrogen oxalate (TETRAM). *Toxicol. Appl. Pharmacol.* 2:1–13.
- Share PA, Brodie BB, Hogben CAM. (1971) The gastric secretion of drugs: a pH partition hypothesis. *J. Pharmacol. Exp. Ther.* 119:361.
- Sinow J, Wei E. (1973) Ocular toxicity of paraquat. *Bull. Environ. Contam. Toxicol.* 9:163–168.

- Smolinske SC. (1992) *Handbook of Food, Drug and Cosmetic Excipients*. CRC Press, Boca Raton, FL.
- Spector WS. (1956) *Handbook of Biological Data*. W. B. Saunders, Philadelphia, PA.
- Stiegeland AJ, Noseworthy MM. (1963) Use of nonaqueous solvents in parenteral products. *J. Pharm. Sci.* 52:917–927.
- Staab RJ, Palmer MA, Auletta CS, Blaszcak DL, McConnell RF. (1987) A relevant vaginal irritation/subacute toxicity model in the rabbit and ovariectomized rat. *Toxicologist*. 7A:1096.
- Stahl PH, Wermuth CG. (Eds.) (2011) *Handbook of Pharmaceutical Salts: Properties, Selection and Use*, 2nd Ed. Wiley-VCH & VHCA, Zurich.
- Strickley RG. (2008) Formulation in drug discovery. In: *Annual Report on Medicinal Chemistry* (Wood A and Desai M, Eds.). Elsevier, Oxford, pp. 419–451.
- Swan KC, White NG. (1942) Corneal permeability: (1) factors affecting penetration of drugs into the cornea. *Am. J. Ophthalmol.* 25:1043–1058.
- Swarbrick J. (Ed.) (2006) *Current Concepts in Pharmaceutical Sciences: Dosage Form, Design and Bioavailability*. Lea & Febiger, Philadelphia, PA.
- Swenson MJ. (Ed.) (1977) *Duke's Physiology of Domestic Animals*. Comstock, Ithaca, NY, p. 178.
- Turco S, King RE. (1974) *Sterile Dosage Forms*. Lea & Febiger, Philadelphia, PA.
- USP (2014) United States Pharmacopoeia 37/NF 32, Washington, DC.
- Weiner M, Bernstein IL. (1989) *Adverse Reactions to Drug Formulation Agents*. Marcel Dekker, New York.
- Weiner ML, Katkoskie LA. (1999) *Excipient Toxicity and Safety*. Marcel Dekker, New York.
- Wilson FA, Dietschy JM. (1974) The intestinal unstirred layer-its surface area and effect on active transport kinetics. *Biochem. Biophys. Acta*. 363:112–126.
- Winne D. (1978) Dependence of intestinal absorption *in vivo* in the unstirred layer. *Nauyn-Schmiedelberg's Arch. Pharmacol.* 304:175–181.
- Wood SG, Upshall DG, Bridges JW. (1978) The absorption of aliphatic carbamates from the rat colon. *J. Pharm. Pharmacol.* 30:638–641.
- Ziegenmeyer J. (1982) The influence of the vehicle on the absorption and permeation of drugs. In: *Dermal and Transdermal Absorption* (Brandan R and Reisen P, Eds.). Wissenschaftliche Verlagsgesellschaft, Stuttgart, pp. 73–89.
- Zimm BH, Lundberg JL. (1956) Sorption of vapors by high polymers. *J. Phys. Chem.* 60:425.

NONCLINICAL MANIFESTATIONS, MECHANISMS, AND END POINTS OF DRUG TOXICITY

A fundamental precept in toxicology (and in clinical medicine) is that the place to start an evaluation of adverse effects or disease must be a knowledge and understanding of what is normal (the normative or healthy state). We use concurrent controls in nonclinical toxicology studies to provide assurance that the specific experimental units (for the purposes of this chapter, experimental animals from the same source and selected as a random sample from a single lot) treated are evaluated in reference to members of the same population. But what represents a change from normal depends on a learned background knowledge from observation and practice and from databases of laboratory and other measurements taken over time from the pool of animals of interest, and what is normal varies in specific ways between species.

In both toxicology and clinical medicine, adversity is nonrandom (significant) changes in one (or sometimes one of two possible) directions which threaten the health (“normality”) of the exposed organism. Detecting these changes and understanding the significance of (and potential for causality of) the effect are the domain of descriptive toxicology and the heart of regulatory toxicology. These changes we measure are the manifestations of toxicology (Table 6.1). What causes toxicity—the underlying mechanisms generalized in Table 6.2—are the realm of investigative and mechanistic toxicology.

Our traditional systemic (“general”) toxicology studies have their genesis in the 1930s, when the practice of clinical medicine added a host of objective tools for the diagnosis and monitoring of health. All these tools were updated for use in animal studies; we came to look for everything that might help find and understand adverse effects—the shotgun approach. We still use this approach, adding more tools as they become available.

6.1 MANIFESTATIONS

There is a vast and expanding literature on manifestations of toxicity—the product of vast (and continuing) efforts in descriptive toxicology, much of it done to meet regulatory needs. Recommended texts which present these in broad categories would include Ballantyne et al. (2009), Hayes (2014), Klaassen (2013), and Smart and Hodgson (2009). Chapter 3 expands on the available sources of such information.

But, at its core, the list of categories of such manifestations (Table 6.1) is rather limited.

In recent years, five more manifestations more specifically or commonly associated with the effects of drugs have come to be added to this list:

- (i) *Mitochondrial dysfunction*. Increased uptake of calcium (because of ATP depletion) by mitochondria activates phospholipases, resulting in accumulation of free fatty acids. These cause changes in the permeability of mitochondrial membranes, such as the *mitochondrial permeability transition*.
- (ii) *Progressive loss of phospholipids*. Increased degradation by endogenous phospholipases and inability of the cell to keep up with the synthesis of new phospholipids (reacylation, an ATP-dependent process).
- (iii) *Cytoskeletal abnormalities*. Activated proteases lyse cytoskeletal elements and cell swelling causes detachment of cell membrane from cytoskeleton; stretching of the cell membrane results in increased membrane damage.
- (iv) *Reactive oxygen species*. Produced within the cell and by infiltrating neutrophils and macrophages,

TABLE 6.1 Manifestations of Toxicity

| |
|--|
| Lethality |
| Local tissue intolerance/corrosion |
| Immune effects |
| Reproductive effects |
| Developmental effects |
| Neurotoxicity |
| Target organ effects |
| Hyper of unintended pharmacology |
| Carcinogenesis |
| Nonspecific nonlethal organism level effects (body weight changes) |

especially after restoration of blood flow to an area (reperfusion injury). Cell injury triggers release of a number of inflammatory cytokines and chemokines which amplify the host immune response and attract neutrophils to the site.

- (v) *Lipid breakdown products*. Unesterified free fatty acids, acylcarnitine, and lysophospholipids. These have a detergent effect on membranes and may exchange with membrane phospholipids, causing permeability changes.

6.2 MECHANISMS OF TOXICITY

Just as manifestations of toxicology assessed in regulatory safety assessment, studies can be considered as belonging in a limited number of categories, so the mechanisms that cause/underlie these measurable and observable effects can be considered as belonging to a limited number of mechanistic causes (summarized in Table 6.2). These now have become more critical in understanding drug toxicity.

The presentation of the concepts of manifestation and mechanisms of toxicity (not the subject of this volume) allows us to consider the actual components measured and evaluated in our “shotgun” regulatory general toxicity studies.

6.3 END POINTS MEASURED IN GENERAL TOXICITY STUDIES

All general toxicology studies capture some or all of a series of common sets of measurements intended to capture as much information as possible without compromising the health or functionality of the main study test animals. Table 6.3 summarizes these categories for both “pilot” (non-GLP, used to properly set dose levels for the longer GLP IND-enabling studies) and for all the follow-on (regulatory required) GLP general toxicology studies. It should be pointed out that new technologies become available (imaging in smaller animals and use of telemetry, for example—both the subject of later chapters in this book).

TABLE 6.2 Mechanisms of Toxicity

| |
|-------------------------------|
| Cytotoxicity |
| Cellular dysregulation |
| Disruption of cellular repair |
| Genotoxicity |
| Membrane effects |
| Receptor effects |
| Cellular organelle effects |
| Biochemical/metabolic effects |
| Genomic effects |

6.3.1 Clinical Observations

The value of clinical observations in general toxicology studies depends heavily on both the training (and experience) of the observers/scorers and of the rigor instilled by the data collection method. Appendix A presents a widely accepted (and used) lexicon of clinical observation terms. How these observations correlate with specific target organs—and what other potentially correlating measures may be—is presented in Table 6.4.

While mortality (on study deaths) and morbidity (on study occurrences of severe clinical signs) are obviously critical parameters in good toxicology studies, they are generally collected and integrated with other clinical observations.

6.3.2 Body Weights

Change in body weight (and the associated calculated parameter of body weight gain) is a nonspecific, broad screen for adverse systemic toxicity. Animals are initially assigned to groups based on a randomization scheme that have each group vary insignificantly from one another in terms of body weight. In most cases (always in rodent studies), the animals are young at study starts (really juveniles) and in a predictable near-log growth phase. Weights are measured prior to the initial dose, then typically 1, 3, 5, 7, 11, and 14 days thereafter, then weekly or biweekly after this in longer studies. The frequency of measurement of weights goes down as the study proceeds—after 2 weeks, weighing is typically weekly through 6 weeks, then every other week through 3 months, and monthly thereafter. Because the animals used in these studies are young adults in the early log phase of their growth, decreases in the rate of gain relative to control animals are a very sensitive (albeit nonspecific) indicator of systemic toxicity.

6.3.3 Food and Water Consumption

Food consumption is typically measured with one or two uses in mind. First, it may be explanatory in the interpretation of reductions (either absolute or relative) in body weight. In cases where administration of the test compound is via diet, it is essential to be able to adjust dietary content so as to

TABLE 6.3 Common Evaluation Parameters for Toxicity Studies

| Parameter | Time Points Evaluated | Pilot Studies | GLP Repeat-Dose Studies (IND Enabling) ^a | Comments |
|-----------------------------------|-------------------------------------|---------------|---|---|
| Mortality/morbidity | 1–2× daily | ✓ | ✓ | |
| Clinical observations | Daily | ✓ | ✓ | |
| FOB | After first and last dose | | ✓ | |
| Body weights | 1–2×/week | ✓ | ✓ | In 90 days and longer studies, after the first 30 days, measured monthly |
| Food and water consumption | 1–2×/week | | ✓ | Typically calculated as a per cage average for rodents |
| Clinical chemistry and hematology | At necropsy | ✓ | ✓ | Prestudy and interim samples are also often collected for large animals |
| Urinalysis | 1–2 days before necropsy | | ✓ | |
| Plasma drug levels | Day 1 and last day of dosing in DRF | ✓ | ✓ | Still commonly requires satellite groups for rodents because of the limited blood volume available for sampling |
| Ophthalmology | Prestudy and before each necropsy | | ✓ | |
| Electrocardiography | Prestudy and before each necropsy | | ✓ | Typically conducted only for nonrodents (dogs, minipigs, and primates) |
| Gross necropsy | At necropsy | ✓ | ✓ | |
| Organ weights | At necropsy | | ✓ | |
| Histopathology | Postnecropsy | | ✓ | |

^a 14 and 28 days studies; longer-term studies may add additional interim time points.

accurately maintain dose levels. Additionally, the actual parameter itself is a broad and nonspecific indicator of systemic toxicity. Food consumption is usually measured over a period of several days, first weekly and then on a once-a-month basis. Water consumption, which is also sometimes measured, is similar in interpretation and use. Additionally, particularly in nonrodents, loss of appetite is an early indicator of adverse effects in animals.

6.3.4 Clinical Signs

Clinical signs are generally vastly underrated in value, probably because insufficient attention is paid to care in their collection. Two separate levels of data collection are actually involved here. The first is the morbidity and mortality observation, which is made twice a day. This generally consists of a simple cage-side visual assessment of each animal to determine if it is still alive and, if so, whether it appears in good (or at least stable) health. Historically, this regulatory required observation was intended to ensure that tissues from intoxicated animals were not lost for meaningful histopathologic evaluation due to autolysis (Arnold et al., 1990).

The second level of clinical observation is the detailed hands-on examination analogous to the human physical examination. It is usually performed against a checklist (see Gad (2009) for an example), and evaluation is of the incidence of observations of a particular type in a group of treated animals compared to controls. Observations range from being indicative of nonspecific systemic toxicity to

fairly specific indicators of target organ toxicity. These more detailed observations are typically taken after the first week of a study and on a monthly basis thereafter.

Ophthalmologic examinations are typically made immediately prior to initiation of a study (and thus serve to screen out animals with preexisting conditions) and toward the end of a study.

Particularly when the agent under investigation either targets or acts via a mechanism likely to have a primary effect on a certain organ for which functional measures are available, an extra set of measurements of functional performance should be considered. The organs or organ systems that are usually of particular concern are the kidneys, liver, cardiovascular, nervous, and immune. Special measures (such as creatinine clearance as a measure of renal function) are combined with other data already collected (organ weights, histopathology, clinical pathology, etc.) to provide a focused “special” investigation or evaluation of adverse effects on the target organ system of concern. In larger animals (dogs and primates), some of these measures (such as ECGs) are made as a matter of course in all studies.

6.3.5 Clinical Chemistry and Pathology

Clinical pathology covers a number of biochemical and morphological evaluations based on invasive and noninvasive sampling of fluids from animals that are made periodically during the course of a subchronic study. These evaluations are sometimes labeled as clinical (as opposed to anatomical)

TABLE 6.4 (Continued)

| Clinical Signs | Potential Target of Toxicity | Potential Correlating Data |
|---|---|----------------------------|
| Straub tail | Central nervous system | Physical examination |
| Tremors | | FOB evaluation |
| Convulsions | | Gross necropsy |
| Seizures | | Clinical chemistry data |
| Stereotypy | | Hematology data |
| Aggression | | Radiographic data |
| Loss of motor function | | Ultrasound data |
| Abnormal gait | | Imaging data |
| Head tilt | | |
| Circling | | |
| Quivering | | |
| Trembling | | |
| Ataxia | | |
| Heightened sensitivity to noise | Autonomic nervous system or nervous system/ nonspecific | Physical examination |
| Lacrimation | | FOB observations |
| Salivation | | Gross necropsy data |
| Rhinorrhea | | Radiographic data |
| Aggression | | Ultrasonographic data |
| Passivity | | Imaging data |
| Resistance to handling | | |
| Vocalization | | |
| Hyperactivity | | |
| Dry eye | | |
| Abnormal response to light stimuli | Ophthalmic or visual nervous system | Physical examination |
| Ptosis | | Complete ophthalmic exam |
| Blinking | | Gross necropsy observation |
| Blepharospasm | | Ultrasonographic data |
| Miosis | | |
| Mydriasis | | |
| Fixed pupils | | |
| Blindness | | |
| Dry eye | | |
| Hyphema | | |
| Cloudy cornea | | |
| Strabismus | | |
| Bulging eye(s) | | |
| Loss of motor function | | Clinical chemistry data |
| Ataxia | | Gross necropsy |
| Generalized muscular weakness | Muscular/neuro muscular systems | Urine data |
| Inability to rise | | Radiographic data |
| Inability or reluctance to move | | Ultrasound data |
| Generalized muscular pain upon palpation | | Imaging data |
| Ataxia | | |
| Inability to rise | Otic system | Clinical chemistry data |
| Inability or reluctance to move | | Hematology |
| Head shaking | | Radiographic data |
| Head tilt | | Ultrasound data |
| Scratching at ear(s) | | Imaging data |
| Anuria | Kidney and ureters and urinary bladder | Relative kidney weights |
| Dysuria | | Gross necropsy |
| Hematuria | | Clinical chemistry data |
| Polyuria | | Urine chemistry data |
| Stranguria | | Ultrasound data |
| Calculi | | Renal clearance |

(Continued)

TABLE 6.4 (Continued)

| Clinical Signs | Potential Target of Toxicity | Potential Correlating Data |
|--|---|--|
| Pallor | Heart and cardiovascular or hematopoietic system including spleen | Physical examination |
| Hypothermia | | Relative heart weight |
| Cyanosis | | Gross necropsy observation |
| Syncope | | Clinical chemistry data |
| Prolonged capillary refill time | | Hematology data |
| Abnormal heart rate | | ECG evaluation |
| Abnormal heart rhythm | | Telemetric evaluation |
| Weak pulse | | Ultrasound data |
| Difference between carotid and femoral pulse intensity | | Specifically targeted clinical chemistry data |
| Bruising | | Coagulation testing profiling |
| Bleeding from orifices | | |
| Fluid-filled abdomen | Hepatic system | Gross necropsy |
| Yellow discoloration to skin, sclera, and mucous membranes | | Clinical chemistry data |
| Abdominal pain | | Hematology data |
| Enlarged +/- liver | | Radiographic data |
| | | Ultrasonographic data |
| | | Imaging data |
| Absent/few feces (droppings) | Gastrointestinal tract | Specifically targeted clinical chemistry data |
| Constipation | | Physical examination |
| Diarrhea | | Gross necropsy |
| Tenesmus | | Clinical chemistry data |
| Bloody stool | | Hematology data |
| Rapid breathing | | Radiographic data |
| Shallow breathing | | Ultrasonographic data |
| Abdominal guarding | | Imaging data |
| Borborygmi | | Evaluation of stomach contents |
| Painful abdomen on palpation | | |
| Gas-filled abdomen | | |
| Noisy breathing | Respiratory system | Physical examination |
| Irregular breathing | | Gross necropsy observation |
| Change in rate of respiration | | Clinical chemistry data |
| Change in rhythm of respiration | | Hematology data |
| Change in depth of breathing | | Radiographic data |
| Change in pattern of breathing to abdominal breathing | | Ultrasonographic data |
| Gasping | | Imaging data |
| Panting | | Evaluation of respiratory parameters such as tidal volume etc. |
| Chattering of teeth | | |
| Labored breathing | | |
| Coughing | | |
| Wheezing | | |
| Restrictive pattern | | |
| Obstructive pattern | | |
| Inspiration difficulty | | |
| Expiration difficulty | | |
| Foamy froth about mouth + lips + nose | | |

TABLE 6.4 (Continued)

| Clinical Signs | Potential Target of Toxicity | Potential Correlating Data |
|---|------------------------------|---|
| Broken bones Generalized malaise Generalized weakness Ataxia | Endocrine | Physical examination Gross necropsy Clinical chemistry data Targeted clinical chemistry or in specific in vitro testing to evaluate specific gland dysfunction |
| Excessive drinking Excessive urination Abnormal circadian rhythm Reproductive dysfunction Immune dysfunction Infection(s) Skin infection(s) | Immune system | Targeted clinical chemistry or in vitro tests to evaluate specific endocrine gland function |

Appendix B and Gad and Chengelis (1999) should be consulted for a lexicon of clinical signs.

TABLE 6.5 Clinical Pathology Measures

| Clinical Chemistry | Hematology | Urinalysis |
|--|-----------------------------------|------------------|
| Albumin | Erythrocyte count (RBC) | Chloride |
| Alkaline phosphatase (ALP) | Hemoglobin (HGB) | Bilirubin |
| Blood urea nitrogen (BUN) | Hematocrit (HCT) | Glucose |
| Calcium | Mean corpuscular hemoglobin (MCH) | Occult blood |
| Chloride | Mean corpuscular volume (MCV) | pH |
| Creatine | Platelet count | Phosphorus |
| Creatine phosphokinase (CPK) | Prothrombin time | Potassium |
| Direct bilirubin | Reticulocyte count | Protein |
| Gamma glutamyl transaminase (GGT) | White cell count (WBC) | Sodium |
| Globulin | White cell differential count | Specific gravity |
| Glucose | | Volume |
| Lactic dehydrogenase (LDH) | | |
| Phosphorus | | |
| Potassium | | |
| Serum glutamic oxaloacetic transaminase (SGOT) | | |
| Serum glutamic pyruvic transaminase (SGPT) | | |
| Sodium | | |
| Total bilirubin | | |
| Total cholesterol | | |
| Total protein | | |
| Triglycerides | | |

pathology determinations and are now commonly the result of evaluations by automated analysis (Thomas, 1979). Table 6.5 presents a summary of the parameters measured under the headings of clinical chemistry, hematology, and urinalysis using samples of blood and urine collected at predetermined intervals during the study. Conventionally, these intervals are typically at three points evenly spaced (chronologically) over the course of the study, with the first being 1 month after study initiation and the last being immediately prior to termination of the test animals. For a 3-month

study, this means that samples of blood and urine would be collected at 1, 2, and 3 months after study initiation (i.e., after the first day of dosing of the animals). There are some implications of these sampling plans that should be considered when the data are being interpreted. Many of the clinical chemistry (and some of the hematologic) markers are really the result of organ system damage that may be transient in nature (see Table 6.6 for a summary of interpretations of clinical chemistry findings and Table 6.7 for a similar summary for hematologic findings). The samples on which

TABLE 6.6 Association of Changes in Biochemical Parameters with Actions at Particular Target Organs

| Parameter | Blood | Heart | Lung | Kidney | Liver | Bone | Intestine | Pancreas | Notes |
|-------------------|-------|-------|------|--------|-------|------|-----------|----------|---|
| Albumin | | | | ↓ | ↓ | | | | Produced by the liver. Very significant reductions indicate extensive liver damage |
| ALP | | | | | ↑ | ↑ | ↑ | | Elevations usually associated with cholestasis. Bone alkaline phosphatase tends to be higher in young animals |
| Bilirubin (total) | ↑ | | | | ↑ | | | | Usually elevated due to cholestasis, either due to obstruction or hepatopathy |
| BUN | | | | ↑ | ↓ | | | | Estimates blood-filtering capacity of the kidneys. Doesn't become significantly elevated until the kidney function is reduced to 60–75% |
| Calcium | | | | ↑ | | | | | Can be life threatening and result in acute death |
| Cholinesterase | | | | ↑ | ↓ | | | | Found in plasma, brain, and RBC |
| CPK | | ↑ | | | | | | | Most often elevated due to skeletal muscle damage but can also be produced by cardiac muscle damage. Can be more sensitive than histopathology |
| Creatinine | | | | ↑ | | | | | Also estimates blood-filtering capacity of kidney as BUN does |
| Glucose | | | | | | | | ↑ | Alterations other than those associated with stress are uncommon and reflect an effect on the pancreatic islets or anorexia |
| GGT | | | | | ↑ | | | | Elevated in cholestasis. This is a microsomal enzyme and levels often increase in response to microsomal enzyme induction |
| HBDH | | ↑ | | | ↑ | | | | Increase usually due to skeletal muscle, cardiac muscle, or liver damage. Not very specific |
| LDH | | ↑ | ↑ | ↑ | ↑ | | | | Absolute alterations are usually associated with decreased production (liver) or increased loss (kidney). Can see increase in case of muscle “wasting” (catabolism) |
| Protein (total) | | | | ↓ | ↓ | | | | |
| SGOT | | ↑ | | ↑ | ↑ | | | ↑ | Present in skeletal muscle and heart and most commonly associated with damage to these |
| SGPT | | | | | ↑ | | | | Elevations usually associated with hepatic damage or disease |
| SDH | | | | | ↑↓ | | | | Liver enzyme which can be quite sensitive but is fairly unstable. Samples should be processed as soon as possible |

↑, increase in chemistry values; ↓, decrease in chemistry values.

ALP, alkaline phosphatase; BUN, blood urea nitrogen; CPK, creatinine phosphokinase; GGT, gamma glutamyl transferase; HBDH, hydroxybutyric dehydrogenase; LDH, lactic dehydrogenase; RBC, red blood cells; SDH, sorbitol dehydrogenase; SGOT, serum glutamic oxaloacetic transaminase (also called aspartate aminotransferase (AST); SGPT, serum glutamic pyruvic transaminase (also called alanine aminotransferase (ALT)).

TABLE 6.7 Some Probable Conditions Affecting Hematological Changes

| Parameter | Elevation | Depression | Parameter | Elevation | Depression |
|-----------------------------|---|--|-------------|---|---|
| Red blood cells | <ol style="list-style-type: none"> 1. Vascular shock 2. Excessive diuresis 3. Chronic hypoxia 4. Hyperadrenocorticism | <ol style="list-style-type: none"> 1. Anemias <ol style="list-style-type: none"> a. Blood loss b. Hemolysis c. Low RBC production | Platelets | | <ol style="list-style-type: none"> 1. Bone marrow depression 2. Immune disorder |
| Hematocrit | <ol style="list-style-type: none"> 1. Increased RBC 2. Stress 3. Shock <ol style="list-style-type: none"> a. Trauma b. Surgery 4. Polycythemia | <ol style="list-style-type: none"> 1. Anemias 2. Pregnancy 3. Excessive hydration | Neutrophils | <ol style="list-style-type: none"> 1. Acute bacterial infections 2. Tissue necrosis 3. Strenuous exercise 4. Convulsions 5. Tachycardia 6. Acute hemorrhage | |
| Hemoglobin | <ol style="list-style-type: none"> 1. Polycythemia (increase in production of RBC) | <ol style="list-style-type: none"> 1. Anemias 2. Lead poisonings | Lymphocytes | <ol style="list-style-type: none"> 1. Leukemia 2. Malnutrition 3. Viral infections | |
| Mean cell volume | <ol style="list-style-type: none"> 1. Anemias 2. B12 deficiency | <ol style="list-style-type: none"> 1. Iron deficiency | Monocytes | <ol style="list-style-type: none"> 1. Protozoal infections | |
| Mean corpuscular hemoglobin | <ol style="list-style-type: none"> 1. Reticulocytosis | <ol style="list-style-type: none"> 1. Iron deficiency | Eosinophils | <ol style="list-style-type: none"> 1. Allergy 2. Irradiation 3. Pernicious anemia 4. Parasitism | |
| White blood cells | <ol style="list-style-type: none"> 1. Bacterial infections 2. Bone marrow stimulation | <ol style="list-style-type: none"> 1. Bone marrow depression 2. Cancer chemotherapy 3. Chemical intoxication 4. Splenic disorders | Basophils | <ol style="list-style-type: none"> 1. Lead poisoning | |

RBC, red blood cell.

analysis is performed are from fixed points in time, which may miss transient changes (typically, increases) in some enzyme levels. Thrall et al. (2012) provide excellent detailed coverage of veterinary hematology and clinical chemistry.

There is now a concerted effort to improve the predictive value of nonclinical safety assessment studies for major classes of issues that are “discovered” in clinical trials and marketed use of new drugs. In particular, there are efforts to identify biomarkers which can be measured in animal studies and have relevance and more similarity for clinical findings (Gupta, 2014).

Particular focus has been on biomarkers for the heart (see Braunwald, 2008), kidney, and liver. The kidney guidance is quite recent with EMEA and FDA announcing in late May 2008 that they have accepted seven biomarkers (Kim-1, albumin, total protein, cystatin C, β 2-microglobulin, urinary clusterin, and urinary trefoil factor 3) for use in preclinical studies. The liver guidance is also recent (EMEA, 2008), though this reflects an ongoing effort (see Kaplowitz, 2005) to try and reduce the occurrence of “idiosyncratic” liver toxicity occurrences in the clinic. Table 6.8 presents a list of available clinical chemistry measures that are considered preclinical alerts for the three major categories of hepatic toxicity.

It would seem likely that rather than these being adapted whole cloth into the standard design of repeat-dose studies, they are likely to be used in a more considered measure when there is cause for specific concern. Most CROs (as of this writing) have not added these to their standard designs or, indeed, offered them as validated, waiting for client demand for same.

A difficulty with detecting potential hepatic toxicities in humans (DLTs—delayed liver toxicities) (Kumar et al., 2015) is that their appearance in limited parts of a patient population is typically after some months of use—the results of first an (innate) immunity inflammatory response followed

by a longer-term adaptive immune response. According to the FDA, drug-induced DLTs are the leading cause of acute liver failure (ALF). Nonclinical studies, particularly the 30-day studies that enable INDs and FIM studies, are just not long enough.

6.3.6 Hematology

In general toxicity studies, we primarily measure the formed blood elements (red and white cells, including their subsets, and platelets), hemoglobin levels in red cells, and the functionality of coagulation. Collectively, these serve to transport oxygen to the tissues of the body, as essential components of the immune system, and to respond to hemorrhagic events. The actual target organ for formed blood element toxicity are the stem cells located in specific tissues of the body (Weiss and Wardrop, 2010; Kaushansky et al., 2011).

6.3.7 Gross Necropsy and Organ Weights

The end of the “in-life” phase of general toxicity studies is the euthanasia of the animals (usually immediately preceded or accompanied by the collection of urine and blood samples) followed by visual examination and preservation of tissues (see list in Table 6.9), with weighing of the selected organs (as denoted in Table 6.9). Observed abnormalities seen on visual examination are investigated in the histopathological examination of the preserved tissues, with attempts to correlate the two results. Organ weights are usually statistically evaluated as ratios using total body weight as their weights change as body weights change (the exception being the brain).

These changes in relative body weights are sensitive indicators of effects on the specific organs of the body and indicators of target organ toxicity.

TABLE 6.8 Clinical Chemistry Measures that are Considered Useful in Identifying Liver Toxicity

| Parameters | Hepatocellular | Hepatobiliary | Mitochondrial |
|------------------------------------|----------------|---------------|---------------|
| Alanine aminotransferase (ALT) | X | | |
| Aspartate aminotransferase (AST) | X | | |
| Sorbitol dehydrogenase (SDH) | X | | |
| Glutamate dehydrogenase (GLDH) | X | | X |
| Total bile acids (TBA) | X | X | |
| Alkaline phosphatase (ALP) | | X | |
| Gamma-glutamyl transferase (GGT) | | X | |
| 5'-Nucleotidase (5-NT) | | X | |
| Total bilirubin (TBILI) | | X | |
| <i>Potential ancillary markers</i> | | | |
| Lactate | | | X |
| Lactate dehydrogenase (LHD) | X | | |
| Ornithine carbamoyltransferase | X | | X |
| Unconjugated bilirubin (UBILI) | X | | |

TABLE 6.9 Tissues for Histopathology

| | |
|--------------------------------------|--|
| Adrenals ^a | Mainstream Bronchi |
| Body and cervix | Major salivary gland |
| Brain, all three levels ^a | Mesenteric lymph nodes |
| Cervical lymph nodes | Ovaries and tubes |
| Cervical spinal cord | Pancreas |
| Duodenum | Pituitary |
| Esophagogastric junction | Prostate |
| Esophagus | Skeletal muscle from proximal hindlimb |
| Eyes with optic nerves | Spleen ^a |
| Femur with marrow | Sternebrae with marrow |
| Heart | Stomach |
| Ileum | Testes with epididymides ^a |
| Kidneys ^a | Thymus and mediastinal contents ^a |
| Large bowel | Thyroid with parathyroid ^a |
| Larynx with thyroid and parathyroid | Trachea |
| Liver ^a | Urinary bladder |
| Lungs ^a | Uterus including horns |

^aOrgans to be weighed.

6.3.8 Histopathology

Histopathology is generally considered the single most significant portion of data to come out of a repeat-dose toxicity study. It actually consists of three related sets of data (gross pathology observations, organ weights, and microscopic pathology) that are collected during the termination of the study animals. At the end of the study, a number of tissues are collected immediately after termination from all animals that survived until the end of the study (test and control). Organ weights and terminal body weights are recorded at study termination so that absolute and relative (to body weight) values can be statistically evaluated.

These tissues, along with the organs for which weights are determined, are listed in Table 6.9. All tissues collected are typically processed for microscopic observation, but only those from the high-dose and control groups are necessarily minimally evaluated microscopically. If a target organ is discovered in the high-dose group, then successively lower-dose groups are examined until a “clean” (devoid of effect) level is discovered (Haschek et al., 2013). A fundamental understanding of what is normal in tissue structure for a specific organ from a specific mammalian species is essential in such evaluations and is the result of years of study and experience (Minckler, 1971).

In theory, all microscopic evaluations should be performed blind (without the pathologist knowing from which dose group a particular animal came), but this is difficult to do in practice and such an approach frequently degrades the quality of the evaluation. Like all the other portions of data in the study, proper evaluation benefits from having access to all data that addresses the relevance, severity, timing, and potential

mechanisms of a specific toxicity. Blind examination is best applied in peer review or consultations on specific findings.

In addition to the “standard” set of tissues specified in Table 6.9, observations during the course of the study or in other previous studies may dictate that additional tissues be collected or special examinations (e.g., special stains, polarized light or electron microscopy, immunocytochemistry, or quantitative morphometry) be undertaken to evaluate the relevance of, or understand the mechanisms underlying, certain observations.

Histopathology testing is a terminal procedure, and, therefore, sampling of any single animal is a one-time event (except in the case of a tissue collected by biopsy). Because it is a regulatory requirement that the tissues from a basic number of animals be examined at the stated end of the study, an assessment of effects at any other time course (most commonly, to investigate recovery from an effect found at study termination) requires that satellite groups of animals be incorporated into the study at start-up. Such animals are randomly assigned at the beginning of the study and otherwise treated exactly the same as the equivalent treatment (or control animals).

6.3.9 Ophthalmology

Ophthalmological examination of all animals in study (particularly nonrodents) should be performed both before study initiation and at the completion of the period at which the drug is administered. This should be performed by an experienced veterinary ophthalmologist.

6.3.10 Cardiovascular Function

Particularly in light of recent concerns with drug-induced arrhythmias, careful consideration must be given to incorporate adequate evaluation of drug-induced alterations on cardiovascular function. This is usually achieved by measuring blood pressure, heart rate, and an ECG prestudy and periodically during the course of the study (usually at least one intermediate period and at the end of the study) in the nonrodent species being employed.

6.3.11 Neurotoxicology

Table 6.10 presents the FDA’s current draft criteria (FDA, 1993, 2007) for end points to be incorporated in studies as a screen for neurotoxicity. In practice, a functional observation battery is employed at several end points (usually 1 and 3 months into the study) to fill these requirements.

6.3.12 Immunotoxicology

In response to concerns about potential effects of drugs on the immune system, ICH and FDA (2006) have promulgated a

TABLE 6.10 FDA Criteria for a Neurotoxicity Screen as a Component of Acute, Pilot, and Subchronic Studies

| |
|--|
| Histopathological examination of tissues representative of the nervous system, including the brain, spinal cord, and peripheral nervous system |
| Quantitative observations and manipulative test to detect neurological, behavioral, and physiological dysfunctions. These may include: |
| General appearance |
| Body posture |
| Incidence and severity of seizure |
| Incidence and severity of tremor, paralysis, and other dysfunction |
| Level of motor activity and arousal |
| Level of reactivity to stimuli |
| Motor coordination |
| Strength |
| Gait |
| Sensorimotor response to primary sensory stimuli |
| Excessive lacrimation or salivation |
| Piloerection |
| Diarrhea |
| Ptoxis |
| Other signs of neurotoxicity deemed appropriate |

TABLE 6.11 FDA Draft Recommendation for Type I Immunotoxicity Test That Can Be Included in Repeated Dose Toxicity Studies

| Type I Test | |
|---|---------------------------|
| Hematology | Clinical chemistry |
| White blood cell counts | Total serum production |
| Differential white blood cell counts | Albumin |
| Lymphocytosis | Albumin-to-globulin ratio |
| Lymphopenia | Serum transaminases |
| Eosinophilia | |
| Histopathology | |
| Lymphoid tissues | |
| Spleen | |
| Lymph nodes | |
| Thymus | |
| Peyer's patches in gut | |
| Bone marrow | |
| Cytology (if needed) ^a | |
| Prevalence of activated macrophages | |
| Tissue prevalence and location of lymphocytes | |
| Evidence of B-cell germinal centers | |
| Evidence of T-cell germinal centers | |
| Necrotic or proliferative changes in lymphoid tissues | |

^aMore comprehensive cytological evaluation of the tissues would not be done unless there is evidence of potential immunotoxicity from the preceding evaluation.

guidance calling for a basic set of potential indicators of immunotoxicity (Table 6.11) be evaluated and considered in standard repeat-dose studies. Most of these end points are, it should be noted, already collected in traditional subchronic designs.

6.4 COMPLICATIONS

A vast number of potential complications can plague the interpretation of general toxicology studies.

Health of the animals and variability of technician procedures and practices in handling of animals are two potentially “lurking” variables not readily visible but quite capable of skewing the results of a study.

Stress as a confounding variable is well recognized but not generally well understood or characterized, but it can have profound effects on animals and alter (or potentially mask) the observation of toxicity, acting primarily by modulating the immune system (both innate and humoral, and noted as such in ICH S8). The June 2013 issue of *Toxicologic Pathology* provides an excellent overview in the form of contributed articles.

REFERENCES

- Arnold DL, Grice HC, Krewski DR. (1990) *Handbook of In Vivo Toxicity Testing*. Academic Press, San Diego, CA.
- Ballantyne B, Marrs T, Syverson T. (2009) *General and Applied Toxicology*, 3rd Ed., 6 volumes. Macmillan, New York.
- Braunwald E. (2008) Biomarkers in heart failure. *NEJM*. 358:2148–2159.
- European Medicines Agency (EMA). (2008) *Non-Clinical Guideline on Drug-Induced Hepatotoxicity*. UK, London.
- FDA. (1993) Chapter V. Additional recommended studies. In: *Toxicological Principles for the Safety Assessment of Direct Food Additives and Color Additives Used in Food, Redbook II*. Center for Food Safety and Applied Nutrition, FDA, Washington, DC.
- FDA. (2006) *Guidance for Industry S8 Immunotoxicity Studies for Human Pharmaceuticals*. Center for Drug Evaluation and Research and Center for Biologics Evaluation and Research, Rockville, MD.
- FDA. (2007) Chapter IV. Guidelines for toxicity studies. In: *Toxicological Principles for the Safety of Food Ingredients, Redbook 2000*. Center for Food Safety and Applied Nutrition, FDA, Washington, DC.
- Gad SC. (2009) *Drug Safety Evaluation*, 2nd Ed. John Wiley & Sons, Inc, Hoboken, NJ.
- Gad SC, Chengelis CP. (1999) *Acute Toxicology Testing*, 2nd Ed. Academic Press, San Diego, CA.
- Gupta RC. (2014) *Biomarkers in Toxicology*. Academic Press, San Diego, CA.
- Haschek WM, Rousseaux CG, Wallig MA. (Eds.) (2013) *Haschek and Rousseaux's Handbook of Toxicologic Pathology*, 3rd Ed. Academic Press, San Diego, CA.
- Hayes AW. (2014) *Principles and Methods of Toxicology*, 6th Ed. Taylor & Francis, Philadelphia, PA.
- Kaplowitz N. (2005) Idiosyncratic drug hepatotoxicity. *Nat. Rev. Drug. Discov.* 4:489–499.

- Kaushansky K, Lichtman MA, Beutler E, Kipps TJ, Seligsohn U, Prchal JT. (2011) *Williams Hematology*, 8th Ed. McGraw-Hill Medical, New York.
- Klaassen CD. (2013) *Casarett & Doull's Toxicology*, 8th Ed. McGraw-Hill, New York.
- Kuman V, Abbas AK, Aster JC. (2015) *Robbins & Cotran Pathologic Basis of Disease*, 9th Ed. Elsevier, Philadelphia, PA, pp. 840–841.
- Minckler J. (1971) *Pathology, an Introduction*. C.V. Mosby Co., St. Louis, MO.
- Smart RC, Hodgson E. (2009) *Molecular and Biochemical Toxicology*, 4th Ed. John Wiley & Sons, Inc, Hoboken, NJ.
- Thomas HC. (1979) *Handbook of Automated Electronic Clinical Analysis*. Reston Publishing Company, Reston, VA.
- Thrall MA, Weiser G, Allison RW, Campbell TW. (2012) *Veterinary Hematology and Clinical Chemistry*, 2nd Ed. Wiley-Blackwell, Philadelphia, PA.
- Weiss DJ, Wardrop KJ. (2010) *Schalm's Veterinary Hematology*, 6th Ed. Wiley-Blackwell, Philadelphia, PA.

PILOT TOXICITY TESTING IN DRUG SAFETY EVALUATION: MTD AND DRF

7.1 INTRODUCTION

Acute toxicity testing serves to detect, identify, and assess the dose–response relationship of a single (or single day’s) administered doses of a drug or chemical entity. Such effects are produced by a single or a few doses over the course of a day such as *bid* or *tid* (in the case of continuously infused intravenous formulation, in a 24-h course of treatment) of a drug and expressed in no more than 14 days (though they most commonly are seen within 3–4 days). Historically, the main focus of these tests has been lethality determinations and the identification of overt signs and symptoms of over-dosage. For a complete historical perspective, see Andrew (2009), Gad and Chengelis (1999), Auletta (1998), Piegorsch (1989), or Deichmann and Gerade (1969). A more enlightened and modern view holds that, especially for pharmaceutical agents, lethality in animals is a relatively poor predictor of hazard (other than lethality) in man (Gad and Chengelis, 1999) and is of limited direct utility for drugs. The current trend is either to generate increasing amounts of more sophisticated data from these tests (“expanded acutes”), usually for “phase 0” studies, or to do only a limited maximum tolerated dose (MTD¹) study to identify doses for the repeat-dose pilot phase (USFDA, 2006; Chan and Hayes, 2015; Wilson et al., 2015). The various types of acute study designs, their utility in pharmaceutical product testing, their limitation, and the resultant sample data are discussed in

this chapter. The most common use of these is as part of a screening program of potential candidates. What is seen in these studies, it must be kept in mind, is route, test species, and formulation dependent. Results are due to the inherent toxicity of the drug, systemic toxicity of what actually gets absorbed, and local tissue tolerance effects at the site of administration (Andrew, 2009).

For new product approvals (and first-in-human clinical trials), single-dose toxicity studies are required by regulatory authorities, though this requirement is being challenged in the European communities (Osterberg, 1983; OECD, 1991; USFDA (CDER), 1996; USFDA, 2006).

Another use, now almost abandoned except for in natural product-derived drugs (Pendergast, 1984), is in quality control testing or batch release testing. The latter was once a mandated part of the standardization process for antibiotics, digoxin, and insulin in the US Pharmacopeia. While, perhaps, this type of testing is part of a broad safety picture, it is not typically part of a “preclinical” safety package used to make decisions on whether to further develop a NCE or on what the allowable clinical dosage shall be. Therefore, these uses are not discussed here. The emphasis in this volume, rather, is on tests used to elucidate the toxicity of new chemical entities, not the safety of finished drug preparations. These tests fall into three general categories: (i) range-finding studies, used primarily to set dosages for initial subchronic or acute testing; (ii) complete “heavy” or expanded acute toxicity tests, used to thoroughly describe the single-dose toxicity of a chemical or to support the opening of an explanatory or phase zero IND, and (iii) screening tests, used to select candidates for development.

¹MTD refers to the highest dose of a drug or treatment that does not cause unacceptable side effects.

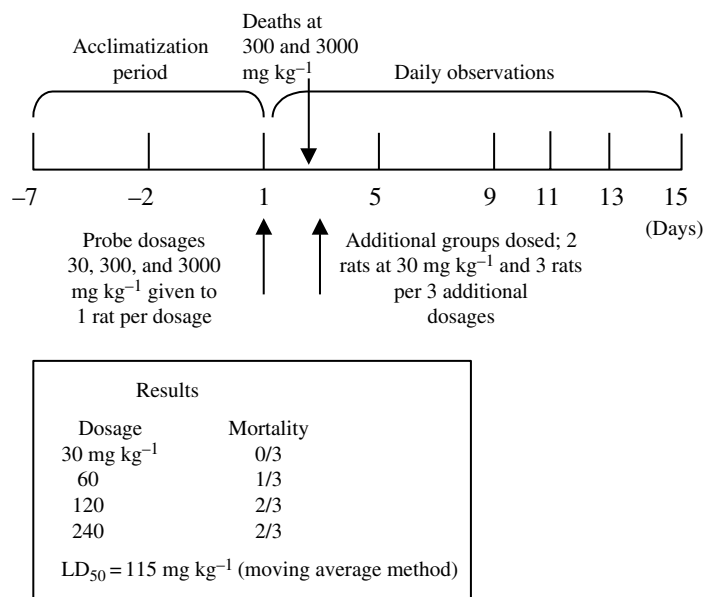


FIGURE 7.1 Example of typical dosage probe protocol.

7.2 RANGE-FINDING STUDIES

Range finders or pilots (now also commonly called DRFs²) are not required to be or normally conducted completely under the auspices of the Good Laboratory Practice Act. They are not used to generate data to support decisions on human safety; rather, they are used to provide sufficient data to allow successful dose selection for definitive (GLP regulatory) toxicity studies. These dosage-level determinations can be for use in acute studies, in *in vivo* genotoxicity studies, or subchronic studies. As discussed by Gad and Chengelis (1999), however, there can be a great deal of difference between the single toxic dosage and subchronic repeat-dose toxicity of a drug. Therefore, acute range-finding studies currently and most commonly are the initial component of a two-phase study where a second set ("B" phase) of animals will receive a short-term treatment (up to 7 days) with the drug in question. These studies seek to identify a MTD for dose setting in subsequent studies.

7.2.1 Lethality Testing

Often, in range-finding tests, the end point is simply to determine what the maximum dosage that can be given without killing an animal is for the drug. There are numerous designs available for obtaining this information that minimizes the expenditure of animals and other resources.

The minimum lethal dose (MLD), also sometimes called the LD_{LO}, will usually be determined in such studies.

7.2.1.1 Classical LD₅₀ The LD₅₀ test has a rich and controversial history (it is one of a number of tests that raises the ire of the animal welfare movement) (Trevan, 1927; Rowan, 1981; LeBeau, 1983). In pharmaceutical development, however, there is rarely a need or requirement for an LD₅₀ (Office of Science Coordination, FDA, 1984). In general, a complete, precisely calculated LD₅₀ consumes more resources than is generally required for range-finding purposes. The reader is referred to Gad and Chengelis (1999) for a complete discussion of this test.

7.2.1.2 Dose Probes Dose probe protocols (see Figure 7.1) are of value when one needs the information supplied by a traditional protocol but has no preliminary data from which to choose dosages. In this acute protocol, one animal is dosed at each of three widely spaced dosages, where the top dosage is generally the maximum amount deliverable. The method works best if the dosages are separated by constant multiples (e.g., 3000, 300, and 30 mg kg⁻¹—a logarithmic progression). Subsequent dosages are selected on the basis of the results from these probe animals. If none of these animals die, the protocol defaults to a limit test (described later), and two more animals are dosed at the top dosage to confirm the limit.

A dose probe can develop into a more thorough lethality determination. If one or two animals die, then two additional dosages between the lethal and nonlethal dosages are chosen and three animals treated per dosage for defining acute lethality. Selection of these dosages is often a matter of

²DRF studies are studies which are necessary to set doses for GLP studies required by regulation to be able to initiate initial clinical studies. Their results help to select doses, and also to justify that dose selection and identify need for modification in study design.

TABLE 7.1 Dosage Selection for the Two-Step Dose-Probing Protocol Design

| Mortality by Dose ^a | | | Dosages (mg kg ⁻¹) for the Definitive Experiment as Determined by the Results of the Probe | | | |
|--------------------------------|-------------------------|--------------------------|--|-------------------|------|------|
| 10 mg kg ⁻¹ | 100 mg kg ⁻¹ | 1000 mg kg ⁻¹ | | | | |
| 0/3 | 0/3 | 0/3 | | 1600 | 2900 | 5000 |
| 0/3 | 0/3 | 1/3 | 600 | 1000 ^b | 1600 | 2900 |
| 0/3 | 0/3 | 2/3 | 200 | 400 | 800 | 1600 |
| 0/3 | 0/3 | 3/3 | 140 | 225 | 370 | 600 |
| 0/3 | 1/3 | 3/3 | 50 | 100 ^b | 200 | 400 |
| 0/3 | 2/3 | 3/3 | 20 | 40 | 80 | 160 |
| 0/3 | 3/3 | 3/3 | 15 | 25 | 40 | 60 |
| 1/3 | 3/3 | 3/3 | 5 | 10 ^b | 20 | 40 |
| 2/3 | 3/3 | 3/3 | 2 | 4 | 8 | 16 |
| 3/3 | 3/3 | 3/3 | 1 | 2 | 4 | 8 |

Source: Data from Lörke (1983).

^aNumber of animals that died/number of animals used.

^bThe results from the probe are inserted for these doses.

personal judgment. If, for example, one wishes to apply the moving average method of calculation, these subsequent dosages can be either even fractions of the top dosage or even multiples of the low dosage. In either case, two to three animals are dosed at the initial dose and three to four animals are dosed at each of the two to three new dosages. The results should be three to four groups of three to four animals each, which should probably provide sufficient data for calculating the LD₅₀ and the slope of the curve. Probing can also be used to define the dosages for subchronic tests. Instead of selecting additional doses for an acute study, one can use the results from the probe to select two dosages for a short (e.g., 5 days) daily dosing regimen (see Section 7.2.1.7).

In a few instances, all the animals may die following the first day of dosing. In that case, the probe activity continues on day 2 with two more animals dosed at two widely spaced lower dosages (i.e., 3 and 0.3 mg kg⁻¹). This regimen could continue daily until a nonlethal dosage is identified. Unless one has grossly misestimated the toxicity of the test substance, it is unlikely that the probing process would take more than 3 days. Carrying our example into 3 days of dosing would have resulted in probing the 3 µg kg⁻¹ to the 3 g kg⁻¹ range, and it is a rare chemical that is lethal at less than 3 µg kg⁻¹. Once a nonlethal dosage is identified, additional animals and/or dosages can be added, as discussed earlier.

There are two disadvantages to dose probe studies. First, delayed deaths pose difficulties. Hence, all animals should be observed for at least 7 days after dosing (though most deaths occur within 3 and very few after 4 days). Second, if the follow-up dosages are not lethal, the next decision point is ill defined. Should more animals be dosed at some different dosage? The resulting data sets may be cumbersome and difficult to analyze by traditional statistical methods. Alternatively (and this is true regardless of protocol design), if no “partial response” (mortality >0 but <100%) dosage is identified, one can simply conclude that the LD₅₀ is between

two dosages, but the data do not permit the calculation of the LD₅₀ or the slope of the curve. This can happen if the dosage response is fairly steep.

Lörke (1983) has developed a similar protocol design. His probe (or dose range) experiment consists of three animals per dosage at 10, 100, and 1000 mg kg⁻¹. The results of the experiment dictate the dosages for the second round of dosing, as shown in Table 7.1. Animals were observed for 14 days after dosing. Lörke (1983) compared the results obtained when one to five animals were used per dosage group for the second test. He concluded that using only one animal per group has unreliable results in only 7% of chemicals tested. Hence, the Lörke design can produce reasonable estimates of lethal dosages using 14 or fewer animals. Schultz and Fuchs (1982) have proposed a dose probe protocol that adequately deals with delayed deaths (Figure 7.2). All animals are observed for 7 days before subsequent dosages are given. Dosing is initiated at two widely delivered dosages using one rate for each dosage. A third probe dosage is determined pending the outcome of the first two probes. A fourth may also be used. After that groups of three to four animals are used at subsequent dosages either as part of a “para-acute” dosing regimen to select or confirm dosages for a subchronic study or to continue with the definition of an acute lethality curve.

7.2.1.3 Up/Down Method Using classical or traditional acute lethality protocols, 15–30 animals per curve would be required to calculate a single LD₅₀ (Berkson, 1955). This is because the method relies on the analysis of group responses. The up/down method can provide lethality information by analyzing the responses on an individual animal basis using appropriate statistical maximum likelihood methods (Bruce, 1985). Deichmann and LeBlanc (1943) published an early method that provided an estimate of lethality using no more than six animals. All animals were dosed at the same

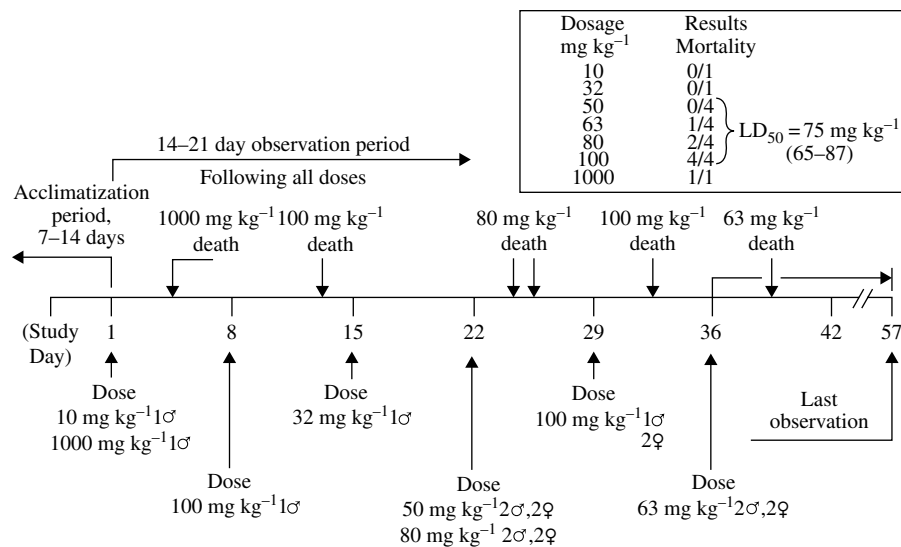


FIGURE 7.2 Example of dose probe method with delayed deaths. *Source:* Adapted from Schultz and Fuchs (1982).

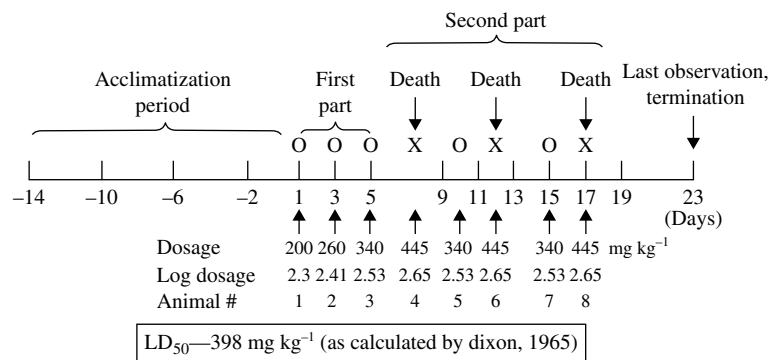


FIGURE 7.3 Example of typical up/down acute lethality protocol.

time. The dosage range was defined as $1.5 \times$ a multiplication factor (e.g., 1.0, 1.5, 2.2, 3.4, 5.1 mL kg⁻¹). The approximate lethal dose (ALD), as they defined it, was the highest dose that did not kill the recipient animal. The resultant ALD differed from the LD₅₀ (as +calculated by the probit method from more complete data sets) by -22 to +33%.

The Deichmann method proved to be too imprecise (Muller and Kley, 1982). Later, Dixon and Wood (1948), followed by Brownlee et al. (1953), developed the method in which one animal was exposed per dosage, but subsequent dosages were adjusted up or down by some constant factor depending on the outcome of the previous dosage. In this method (Figure 7.3), which has been developed more extensively by Bruce (1985), individual animals are dosed at different dosages on successive days. If an animal dies, the dosage for the next animal is decreased by a factor of 1.3. Conversely, if an animal lives, the next dosage is increased by a factor of 1.3. The process is continued until five animals have been dosed after a reversal of the first observation. Alternatively, one can use the tables developed by Dixon (1965). This design can be

used not only for range-finding purposes but also to define an LD₅₀ if this value is needed. In general, only six to nine animals are required—unless the initial dosages are grossly high or low. When compared to the LD₅₀ obtained by other more classical protocols, excellent agreement is obtained with the up/down method (Bruce, 1985). As with classical protocols, sexes should be tested separately. However a further reduction in the numbers of animals used can be accomplished if one is willing to accept that females are of the same or increased sensitivity as males, as is the case approximately 85–90% of the time (Gad and Chengelis, 1999).

There are three main disadvantages to using the up/down method. The first is regulatory, the second procedural, and the third scientific. First, many regulatory guidelines simply have a requirement for the use of traditional protocols. Some also specify the method of calculation. Second, the sequential dosing design is inappropriate for substances that cause delayed deaths. As reported by various authors (Gad et al., 1984; Bruce, 1985), delayed deaths (beyond 2 days after dosing) are rare but not known. They are most prevalent

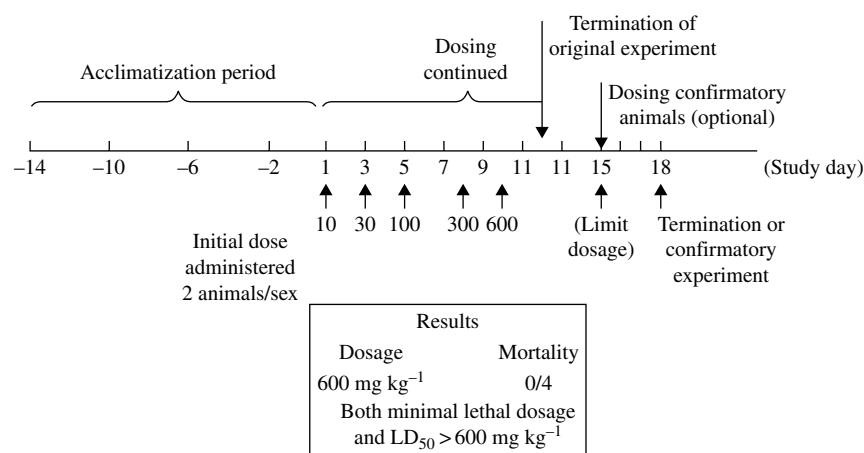


FIGURE 7.4 Example of typical pyramiding dose protocol.

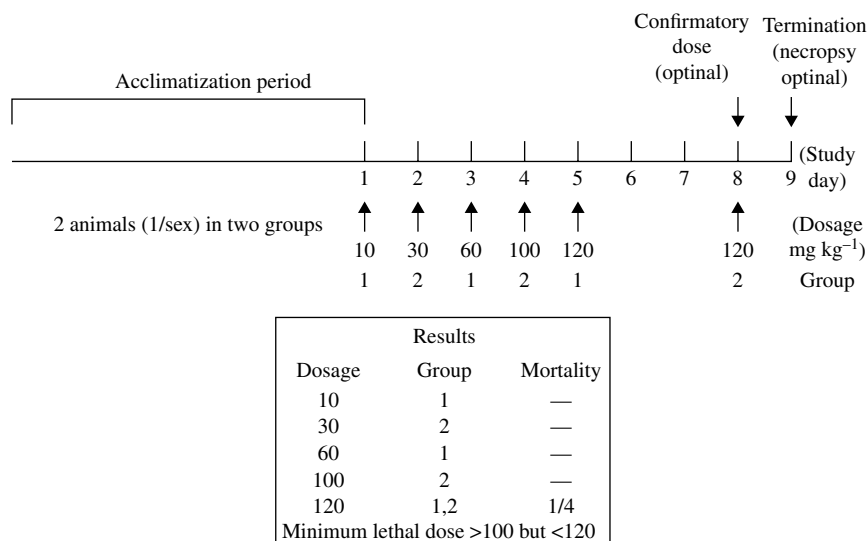


FIGURE 7.5 Example of typical “leapfrog” dosing protocol.

when animals are dosed by the intraperitoneal route with a chemical that causes peritonitis. Death secondary to severe liver or gastrointestinal damage may also take over 2 days to occur. To guard against possible spurious results, all animals should be maintained and observed for at least 7 days after dosing. If delayed deaths occur, the original data set must be corrected and the LD₅₀ recalculated. A substantial number of delayed deaths could result in a data set from which an LD₅₀ cannot be calculated, in which case the test should be rerun.

7.2.1.4 “Pyramiding” Studies Using this type of design (Figure 7.4), one can obtain information about lethality with the minimum expenditure of animals. A minimum of two animals are dosed throughout the study, usually on alternate days (e.g., Monday, Wednesday, and Friday), but the dosage at session may be 1, 3, 10, 30, 100, 300, 1000, and 3000 mg kg⁻¹ or 10, 20, 40, 80, 160, 320, 640, and

1280 mg kg⁻¹. One is literally stepping up, or pyramiding, the lethality–dosage curve. Dosing continues in this fashion until one or both animals exhibit severe toxicity (a nontolerant dose) or until some practical upper limit is reached. For drugs, there is no longer a need to go higher than 1000 mg kg⁻¹ for rodents or nonrodents. An alternative, but similar, design is the “leapfrog” study (Figure 7.5). This consists of two groups of two animals each. They are dosed on alternating days (oral 3-day intervals), but the dosages are increased each day. Extending the example of the pyramiding regiment, group 1 would receive 10, 60, and 120 mg kg⁻¹, while group 2 would be given 30, 100, and 120 mg kg⁻¹. This design is of the value when one has to complete the range-finding activity in a short period of time. Because these designs utilize few animals, they are commonly used for assessing severe toxicity in nonrodent species. An exploratory study typically uses an animal of each sex.

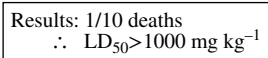


FIGURE 7.6 Example of typical limit test protocol.

There are three conclusions that can be reached on the basis of data from a pyramiding dosage study. First, if none of the animals die or exhibit severe toxicity, then both the threshold or minimum/maximum lethal/toxic dosage (MLD/MTD) and the LD₅₀ are greater than the top or limit dosage. Second, if all animals die at the same dosage, then both the MLD and the LD₅₀ are reported as being between the last two dosages given. This is not an uncommon finding as an indication that the lethality curve has a steep slope. Third, one animal may die at one dosage, and remaining deaths occur at a subsequent dosage. In this case, the MLD is between the lowest nonlethal dosage and the dosage at which the first death occurred, while the LD₅₀ is reported as being between this latter dosage and the dosage at which the last animal dies. A frequently employed variation with nonrodents is, if severe toxicity is not observed, the animals are dosed for five or seven consecutive days at the highest observed tolerated dose. This "phase B" study portion serves to provide more confidence in selecting the top dose in subsequent repeat-dose studies.

There are some disadvantages to the pyramiding dose protocol. First, it cannot produce a true lethality or dose-response curve or provide for the calculation of an LD₅₀. Second, this method cannot identify delayed deaths. If an animal, for example, dies 1 h after the second dosage, one has no way of determining whether it was actually the second dosage or a delayed effect of the first. For this reason it is of little value to observe the animals for any more than a few days after the last dosage. Third, if the test article has an unusually long half-life, bioaccumulation can lead to an underestimation of the acute lethal dosage. By contract, the pharmacological accommodation can lead to a spuriously high estimate of lethality. Depending on the importance of the finding, one may wish to confirm that the results obtained at the highest dosage administered were obtained while dosing two naïve animals at the same dosage. Fortunately, the minimum 48-h period between dosing sessions will minimize such effects. Because of this design feature, it may take as long as 3 weeks to complete the dosing sequence. However, in these types of studies as there is generally no need for a 1–2-week postdosing observation or holding

period, the actual study may not take significantly more time than a test of more traditional design.

Keep in mind that the objective of such studies is to gain information about lethality and gross tolerance for use in setting doses in more extensive studies. For nonrodents (especially monkeys), if none of the animals die or demonstrate obvious signs of toxicity, little would be gained by euthanization and necropsy of such animals. They can be saved and used again, following a reasonable "washout" period, to assess the lethality, toxicity, or safety pharmacology of a different chemical. Indeed, given the emphasis on moving smoothly through development, while necropsy and organ weights may provide useful data in an acceptable time frame, the same is not true for histopathology. In the hands of a skilled toxicologist, such adaptive reuse of animals is a cost-effective way to minimize overall usage.

7.2.1.5 Limit Tests There are relatively innocuous drugs that are simply not potentially lethal. The limit test (Figure 7.6) provides the simplest protocol for determining the lethality of such substances. The limit test is designed to obtain clearance at a specific dosage based on the assumption that what may occur at a higher dosage is not of practical relevance. Thus, one dosage only is studied. This limit “dosage” can be set on the basis of the chemical or physical properties of the test article (or vehicle) or on the basis of an upward safety margin. If the preparation is highly acidic ($\text{pH} < 3$), a large intravenous dose would be expected to cause systemic acidosis as well as local irritation but will yield little relevant toxicology information, as such a preparation would never be approved for clinical use. Alternately, if the anticipated human dosage of a drug is 0.3 mg kg^{-1} , there is probably little reason to test dosages in excess of 300 mg kg^{-1} ($1000\times$ the expected human dosage). In general, there is never any reason (or regulatory requirement) to use dosages of more than 1.5 g kg^{-1} .

There are three possible outcomes to a limit test. If none of the animals die, then the conclusion is that the MLD is greater than the limit dosage. If fewer than 50% of the animals die, then the conclusion is that the LD₅₀ is greater than the limit dosage. If more than 50% of the animals die, then

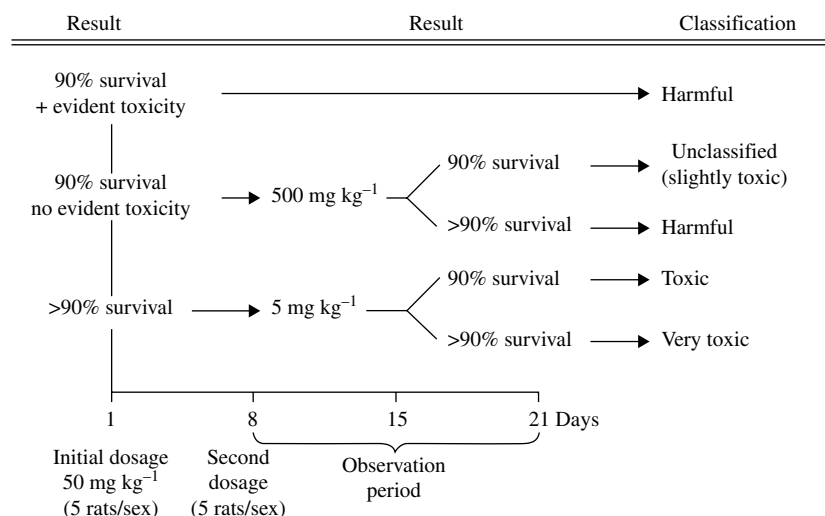


FIGURE 7.7 British Toxicology Society fixed-dose procedure.

one has a problem. Depending on the reasons for performing the test, one could reset the limit and repeat the study, or one could assess lethality by a different protocol. Alternatively, the change in the limit could reflect a change in the chemical or biological properties of the test substance that should be evaluated further.

7.2.1.6 Fixed-Dose Procedure The fixed-dose design (Figure 7.7) was proposed by the British Toxicology Society (1984). It is designed to supply the data needed for classification or labeling purposes. It is essentially a three-step limit test.

Five rats per sex are given 50 mg kg⁻¹. If survival is less than 90%, a second group of animals is given 5 mg kg⁻¹. If survival is again less than 90%, the substance is classified as “very toxic”; otherwise, it is classified as “toxic.”

If, after the 50 mg kg⁻¹ dose, survival is 90% but there is evident toxicity, no further dosages are given and the substance is classified as “harmful.” If, on the other hand, there is no evident toxicity at 50 mg kg⁻¹, another group of rats is given 500 mg kg⁻¹. If there is again 90% survival and no evident toxicity, the substance is given “unclassified” or “slightly toxic” status.

The fixed-dose procedure is relatively new and apparently results in a large decrease in animal usage. It is also noteworthy in that it utilizes not only lethality but also “evident toxicity,” which, in all likelihood, refers to obvious signs of CNS effects, such as seizures or prostration. It remains to be established whether or not this protocol design becomes widely accepted by various regulatory.

The potential utility of the fixed-dose procedure was demonstrated in an international validation study in which the acute oral toxicity of 20 different chemicals was evaluated using both the fixed-dose and classical LD₅₀ procedures. Thirty-three laboratories in eleven different countries were

involved in the validation project, and the results have been published (van den Heuvel et al., 1990). The results demonstrated that the fixed-dose procedure produced consistent evaluations of acute toxicity that were not subject to significant interlaboratory variation and provided sufficient information for hazard identification and risk assessment based on signs of toxicity (clinical signs, time to onset, duration, outcome, etc.). The fixed-dose procedure used fewer animals than the classical LD₅₀ tests and generally required less time to complete. Because of the emphasis on toxicity (rather than mortality) and the use of fewer animals, the fixed-dose procedure could be considered a more “humane” or animal-sparing design than the classical LD₅₀ test. When the results of the fixed-dose and LD₅₀ tests were compared for hazard ranking purposes (Table 7.2), comparable results were obtained. Thus, it would appear that the fixed-dose procedure has utility and has been recommended late in 2000 for broad regulatory adaptation by ICCVAM.

7.2.1.7 “Rolling” Acute Test The rolling acute test is a combination protocol that is designed to find a tolerated dose to use for a subchronic toxicity test. The first segment can be either a dose probe or an up/down or pyramiding type of study to define the MLD. In the second segment, three to five animals are dosed for as short period of time—5 to 7 days. The objective of this design is to compensate for the fact that cumulative toxicity can occur at substantial differences in acute and subchronic toxic dosages. One can be easily misled by selecting subchronic dosages based entirely on acute lethality data. An example is a drug tested where it was found that 360 mg kg⁻¹ was acutely nonlethal and the MLD was 970 mg kg⁻¹. The dosages selected for the 4-week subchronic study were 50, 100, 200, and 400 mg kg⁻¹ day⁻¹. The top-dose animals all died within a week. Substantial mortality occurred at 200 mg kg⁻¹ and evident toxicity was

TABLE 7.2 Comparison of Toxicity Classification Based on LD₅₀ versus Fixed-Dose Procedure

| Test Chemical | Toxicity Classification Based on LD ₅₀ | Fixed-Dose: Number of Laboratories Classifying Chemical | | | |
|---------------------------|---|---|-------|---------|--------------|
| | | Very Toxic | Toxic | Harmful | Unclassified |
| Nicotine | Toxic | — | 23 | 3 | — |
| Sodium pentachlorophenate | Harmful | — | 1 | 25 | — |
| Ferrocene | Harmful/unclassified | — | — | 3 | — |
| 2-Chloroethyl alcohol | Toxic | — | 19 | 7 | — |
| Sodium arsenite | Toxic | — | 25 | 1 | — |
| Phenyl mercury acetate | Toxic | 2 | 24 | — | — |
| <i>p</i> -Dichlorobenzene | Unclassified | — | — | — | 26 |
| Fentin hydroxide | Toxic | — | 8 | 17 | 1 |
| Acetanilide | Harmful | — | — | 4 | 22 |
| Quercetin dihydrate | Unclassified | — | — | — | 26 |
| Tetrachlorvinphos | Unclassified | — | — | 1 | 25 |
| Piperidine | Harmful | — | 2 | 24 | — |
| Mercuric chloride | Toxic | — | 25 | 1 | — |
| 1-Phenyl-2-thiourea | Toxic/harmful | 12 | 12 | 2 | — |
| 4-Aminophenol | Harmful | — | — | 17 | 9 |
| Naphthalene | Unclassified | — | — | — | 26 |
| Acetonitrile | Harmful | — | — | 4 | 22 |
| Aldicarb (10%) | Very toxic | 22 | — | — | — |
| Resorcinol | Harmful | — | — | 25 | 1 |
| Dimethyl formamide | Unclassified | — | — | — | 26 |

Source: Data from van der Heuvel et al. (1990).

present at 50 mg kg⁻¹. A no-effect dosage was not identified, so the entire test had to be repeated with a different dosage structure. The rolling acute structure is a quick and relatively simple “sanity” check that permits one to avoid making such mistakes.

7.2.2 Using Range-Finding Lethality Data in Drug Development: The Minimum Lethal Dose

Range-finding data are often used early in drug development to make preliminary safety estimates. The LD₅₀ is simply a calculated point on a curve. The shape or slope of this curve is also an important characteristic of the test substance. However, unless one does a great deal of acute toxicity testing, the difference between a slope of 1.5 and a slope of 4 has very little meaning. Further, for safety considerations, the dosage that kills 50% of the animals is not as important as the dosage at which lethality first becomes apparent (i.e., the threshold dosage or MLD). For example, if the oral LD₅₀s of two different drugs (A and B) were 0.6 and 2.0 g kg⁻¹, respectively, what would we conclude about the relative safety of these compounds? Further, let us assume that the estimated human dosage of drug A is 0.5 mg kg⁻¹ and of drug B is 5 mg kg⁻¹. Do our conclusions concerning the relative safety of these two drugs change? In fact, the LD₅₀s of both drugs are so high that both are considered only slightly toxic (0.5–5.0 g kg⁻¹). One can also compute the lethality safety margin or index (LSI, equal to LD₅₀/EHD, where EHD = the estimated human dose) for these two drugs; both indices are

so large (1200 for A and 400 for B) that there is still no toxicologically relevant difference between the two drugs. Let us now assume that the lethality curve for substance A is very steep, such that 0.4 g kg⁻¹ causes death in a very small percentage of animals—it is, in fact, the lowest dose administered that causes death. This is the MLD or estimated MLD (EMLD). Let us now assume that the lethality curve for B is very shallow, such that its MLD is also 0.4 g kg⁻¹. Does this change our safety considerations of these two drugs? One can calculate a new more conservative safety index (MLD/EHD) of 800 for A and 80 for B. As a very general rule of thumb, an index for lethality of less than 100 is cause for mild concern, one less than 10 is cause for caution, and one less than 1 should be cause for extreme caution. In the case of our two hypothetical drugs, the development of drug B should be approached with more caution than that of drug A, despite the fact that B has a higher LD₅₀. This is demonstrated in Figure 7.8. There are drugs sold over the counter, however, that have LSI of less than 10. For example, the MLD of indomethacin in rats is 3.7 mg kg⁻¹ (from data reported by Schiantarelli and Cadel (1981)), while the maximum recommended human dose is 200 mg (2.9 mg kg⁻¹ for a 70-kg person); hence, indomethacin has an LSI of 1.3. Such a finding is only cause for some caution but does not in and of itself justify restricting the use or sale of a drug. Hence, because it results in a more conservative safety factor and also takes into consideration the slope of the lethality curve, the use of the MLD rather than the LD₅₀ is recommended in calculating acute safety indices.

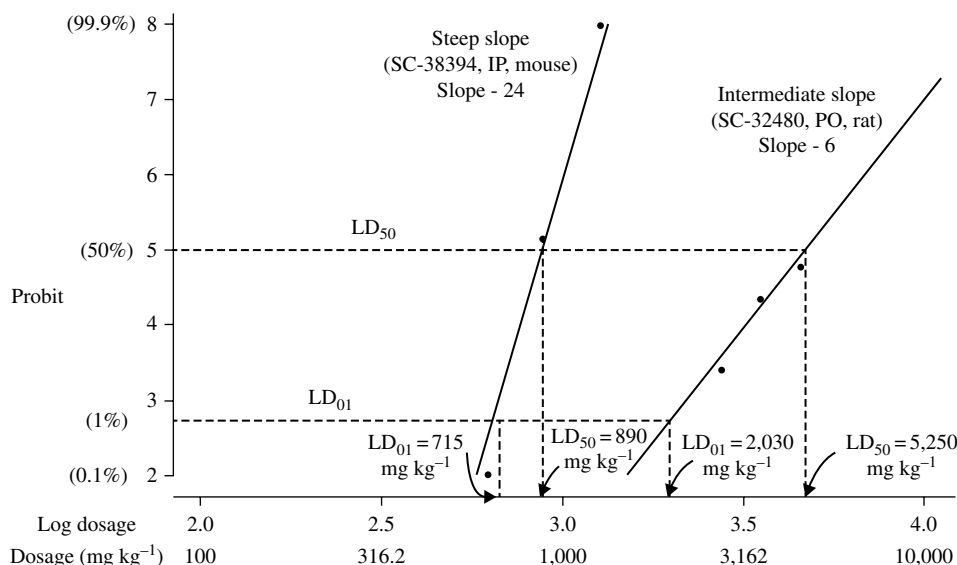


FIGURE 7.8 Examples of probit-log dosage response curves illustrating differences in slope curves and the relationship between the slope, LD_{50} , and LD_{01} .

A number of different safety factors and therapeutic indices have been proposed in the literature. Despite their similarity, some distinction should be made between these two. A therapeutic index applies only to drugs and is the ratio between a toxic dosage (TD, “toxic dose” or LD, “lethal dose”: the toxic end point does not always have to be death) and the pharmacologically effective dosage (ED, “effective dose”) in the same species. A safety index can be calculated for all xenobiotics, not just drugs. A safety index is the ratio of likely human exposure (or dosage) and the dosage that causes death or other forms of toxicity in the most sensitive experimental animal species. The most conservative (L)SI is obtained by dividing the maximum estimated human dosage or exposure by either the MLD or the maximum nonlethal dose (MNLD) (the highest dose tested which does not result in the death of an animal).

7.2.2.1 Minimum Lethal Dose Protocols Stating that the MLD is preferable to the LD_{50} for safety considerations is one thing; trying to determine what a specific MLD may be or could be is another. There are no commonly used experimental designs that have the MLD as an end point. Assuming a log-dose response, the MLD may become a function of group size. Theoretically, if enough animals are dosed, at least one animal could die at any reasonable dosage. There are, however, practical considerations that can and should be applied to determining an MLD. As a practical rule of thumb, we recommend that the estimated LD_{01} —the dose that would be expected to kill 1% of the experimental animals exposed—be used as an estimate of the MLD. If one already has sufficient data to describe a lethality curve, an LD_{01} can be calculated as easily as

the LD_{50} . This is often the case with acute toxicity data obtained to support regulatory submission.

How is the MLD calculated without a complete lethality curve? A modified pyramiding dosage design may be the most appropriate approach. With this design, groups of animals are treated with stepwise increases in dosage until death occurs or a limit dosage is attained. If one has no idea as to what the initial dosage should be or how to graduate the dosages, a dose-probing experiment can be conducted. If the dose-probing experiment produces no deaths, two to three more animals can be dosed at the limit dose to confirm the results; the lethality determination is now complete. If the probe experiment does produce death, then the additional dosages can be graduated between the lowest lethal and the highest nonlethal dosages. A typical progression may proceed as follows (Figure 7.9): on day 1 of the study, three probe animals are dosed at 10, 100, and 1000 $mg\ kg^{-1}$. The animal at 100 $mg\ kg^{-1}$ dies within a few hours of dosing. The two remaining animals are dosed at 300 $mg\ kg^{-1}$ on day 3. Neither dies. They are then dosed at 500 $mg\ kg^{-1}$ on day 5. One dies. Three additional animals should be dosed on day 7 or 8 at a dosage in between (i.e., 400 $mg\ kg^{-1}$ is a good estimate of the maximum nonlethal dose, or MNLD). While different by definition, there is usually not a great deal of distance between the MLD and the MNLD, as this example illustrates. In fact, even for a well-characterized lethality curve, the confidence limits for the LD_{01} will be quite broad and encompass both the MLD and MNLD.

Malmfors and Teiling (1983) have proposed a similar method for determining what they also termed the MNLD. Rather than initiating the study with probe animals, their design calls for three consecutive pyramiding-type studies

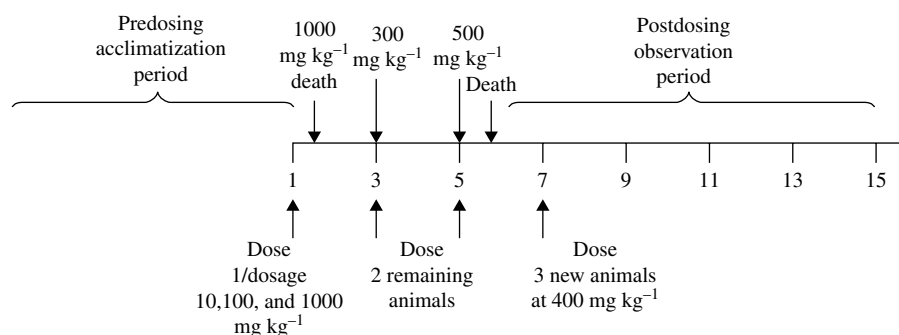


FIGURE 7.9 Example of minimum lethal dose (MLD) pyramiding dose design.

with the steps becoming increasingly smaller. For example, two animals will be sequentially dosed at 2, 200, and 2000 mg kg⁻¹. If death occurs at 2000 mg kg⁻¹, a new pair of animals is initiated at 200 mg kg⁻¹, and sequential dosages are increased by a factor of 1.8 until death occurs. Then another pair of animals is initiated at the highest nonlethal dosage, and successive dosages are increased by a factor of 1.15. The result of this exercise will be two dosages, one apparently nonlethal and the other lethal. Six animals are dosed at each dosage. If none die at the lower dosage and one dies at the higher dose, then the lower dose is considered to be the MNLD. At least 24 h between dosing rounds are recommended. While this method may have some utility, there are some disadvantages. First, the recommended limiting dosage of 6.5 g kg⁻¹ is too high. Second, 24 h between doses may be too short a period to allow for recovery. Third, even with only 24 h between doses, this is a time-consuming procedure—it may take up to 2 weeks to complete the dosing. Finally, it does not decrease the number of animals needed, since it may use 18–20 animals.

Dose probing is not generally used for nonrodents (rather, the pyramiding dose scheme is), and the initiating dosage is normally in the range of 1–5 times the projected human clinical dosage. The limit is generally in the area of 1 g kg⁻¹ or 100–200 times the human dosage, whichever is less. The normal study will include two animals of each sex treated with the test article. For simple lethality, there is seldom any need to include control animals. If the projected human dosage is 4 mg kg⁻¹, for example, the initial dosage in an MLD range finder in dogs will be 20 mg kg⁻¹, and succeeding dosages will increase stepwise at half-log intervals; thus, 20, 60, 200, and 600 mg kg⁻¹ doses are separated by at least 48 h. The MLD is simply reported as being between the highest observable nonlethal and the lowest lethal dosages, or at greater than the limit dosage—in this case, 600 mg kg⁻¹. Studies should not be done with nonrodents solely for determining lethality, because this would not be an appropriate use of time and animals. Generally, these studies should also include some combination of extensive physical examinations, such as ECGs and rectal temperatures,

careful observations of behavior and activity, and extensive clinical laboratory workups after each dose.

The pyramiding dose study is not without disadvantages. First, the small number of animals used can cause simple random variation resulting in misestimation of lethality. It is a well-accepted statistical maxim that the smaller the sample size, the greater the impact of any random variation (error or outlier) on the population characteristic. This may be especially true for a nonrodent species where experimental animals are drawn from an outbred population. Second, the pyramiding dose regimen can permit the development of tolerance. For example, pyramiding dosage studies were conducted to range-find dosages for a 2-week study on a 1,4-benzodiazepine. Lethality in dogs was observed at 600 mg kg⁻¹ in the pyramiding study. For the subsequent subchronic study, the top dose was set at 300 mg kg⁻¹; both dogs died of CNS depression of the first day of dosing.

7.3 ACUTE SYSTEMIC TOXICITY CHARACTERIZATION

Acute systemic toxicity studies are performed to more completely define the acute toxicity of a drug. They are more extensive and time consuming than range-finding tests or screens and are normally the type of study done to satisfy regulatory requirements or to provide a more thorough early characterization or prediction of toxicity (McClain, 1983). In pharmaceutical development, rarely would an acute test be sufficient to support registration, but it could support an exploratory IND first-in-human single-dose study (USFDA, 2006) or a single human dose study of an imaging agent, and it may be required as part of an overall package. These protocols may resemble range-finding tests, but they call for collection of more data. A list of the types of data that can be obtained in well-conducted acute toxicity tests is given in Table 7.3. Given that these studies usually include control groups, the classical or traditional design is the most common because it allows for the most straightforward statistical analyses. In addition, while the use of staggered dosing days

TABLE 7.3 Information, Including Lethality, which Can Be Gained in Acute Toxicity Testing

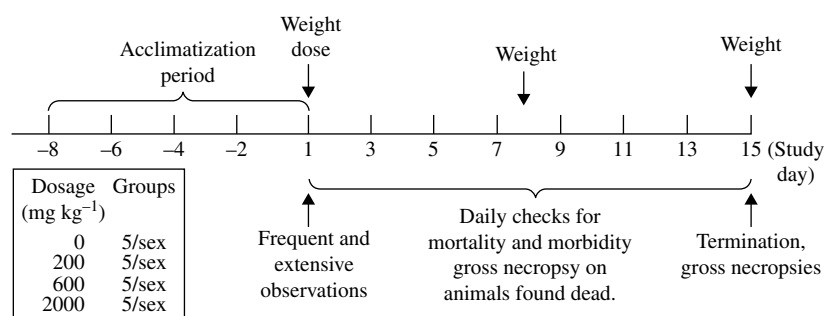
| |
|--|
| Lethality/mortality |
| LD ₅₀ with confidence limits |
| Shape and slope of lethality curves |
| Estimation of maximum nonlethal dose (MNLD) or minimum lethal dose (LD ₀₁) |
| Time to dose estimates |
| Clinical signs |
| Times of onset and recovery |
| Thresholds |
| Agonal versus nonagonal (i.e., do signs occur only in animals that die?) |
| Specific versus general responses |
| Separation of dose–response curves from lethality curves |
| Body weight changes |
| Actual loss versus decreased gain |
| Recovery |
| Accompanied by changes in feed consumption |
| Changes in animals that die versus those that survive |
| Target organ identification |
| Gross examinations |
| Histological examinations |
| Clinical chemical changes |
| Hematological changes |
| Specialized function tests |
| Immunocompetency |
| Neuromuscular screening |
| Behavioral screening |
| Pharmacokinetic considerations |
| Different routes of administration yielding differences in toxicity |
| Plasma levels of test article |
| Areas under the curves, volume of distribution, half-life |
| Metabolic pattern of test article |
| Distribution to key organs |
| Relationship between plasma levels and occurrence of clinical signs |

for different groups is still a fairly common practice, data analyses may be more sensitive if all animals are dosed on the same day, requiring that one have preliminary range-finder data that permit selection of appropriate dosages. Studies of more than one species and/or more than one route should be limited to those instances where they are required by statute.

In general, traditionally designed acute toxicity tests can be divided into three types that can be called the minimal acute toxicity test, the complete acute toxicity test, and the supplemented acute toxicity test. Of these, the minimal protocol is by far the most common and is discussed first. The other two represent increasing orders of complexity as additional parameters of measurement are added to the basic minimal study.

7.3.1 Minimal Acute Toxicity Test

An example of a typical minimal acute toxicity test protocol is shown in Figure 7.10. This study resembles a traditional lethality test in terms of the number of groups and the number of animals per group. Standard protocols consist of three or four groups of treated animals and one group of control animals, each group consisting of five animals per sex per dosage. Traditionally, the emphasis in these types of studies was on determining the LD₅₀, time to death, slope of the lethality curve, and the prominent clinical signs, as illustrated by the data reported by Jenner et al. (1964). More recent designs specify, in addition to lethality and clinical observations, that body weights be recorded during the study and gross necropsies performed at the end of the postdosing observation period. For an excellent example of a well-performed acute toxicity evaluation, the reader is referred to the paper by Peterson et al. (1987) of the acute toxicity of the alkaloids of *Lupinus angustifolius*, in which the LD₅₀s, time to death, clinical signs, body weight effect, and gross necropsy findings were all discussed. For pharmaceuticals, where acute toxicity data for more than one species are often required, these studies will be done as batteries on both rats

**FIGURE 7.10** Example of minimal acute toxicity protocol.

and mice. In addition, because many drugs will be given by more than one route to human patients, these batteries will include groups treated by two different routes. Thus, an acute study on a pharmaceutical agent will often result in eight “curves”—one per route per species per sex. For tests on nonrodent species, as required for pharmaceuticals, a different design is used (discussed later).

The animals should be acclimated to laboratory conditions for 7–14 days prior to dosing. For acute toxicity tests, this pretreatment period should be more than just a holding period. Animals should be checked daily for signs of ill health and/or abnormal behavior. Body weights may also be determined. These data should be used to exclude abnormal animals from the test. Such data also provide an additional basis for interpreting the data gathered during the postdosing period. Finally, these activities acclimate the animals to the frequent handling that is a necessary part of an acute toxicity test.

In selecting dosages for an acute systemic toxicity study, the same general guidelines apply as with lethality testing:

1. There is little to be gained from testing dosages that are so high that the physical rather than biological properties become prominent. Generally, little additional information is gained by pushing dosages past 2 g kg^{-1} . The usual regulatory limit for pharmaceuticals is now 1.5 g kg^{-1} .
2. The highest dosage should be no larger than 100–300 times the anticipated human dosage.
3. Widely spaced dosages are better than narrowly spaced dosages.

This latter point is particularly true in an acute toxicity test on a drug, because pharmacologically based clinical signs may occur at dosages considerably lower than those that cause death. Also, as discussed by Sperling (1976) and Gad (1982), the effects at high dosages may mask the effects that would be observed at low dosages. As human beings are more likely to be exposed to lower dosages than experimental animals, these low-dosage effects may be important parameters to define.

Historically, it has been stated in various regulatory communications that a well-conducted acute toxicity test should contain sufficient data to calculate an LD_{50} . This is no longer necessarily the case. Simpler, less resource-intensive range-finding protocols should be used for defining lethality. Because it is rare that an extensive acute protocol would be attempted without preliminary lethality data, the lethality objectives of acute systemic testing are not always critical. Ideally, the highest dosage should elicit marked toxicity (such as lethality), but it does not need to kill all of the animals to satisfy one's need to show due diligence in stressing the test system. If one already has sufficient preliminary data to suspect that the top dosage will be nonlethal

or otherwise innocuous, the test can be conducted as a limit test, consisting of one treated group and one control group.

7.3.1.1 Clinical Signs The nonlethal parameters of acute toxicity testing have been extensively reviewed by Sperling (1976) and Balazs (1970, 1976). Clinical observations or signs of toxicity are perhaps the most important aspect of a minimal acute toxicity test because they are the first indicators of drug- or chemical-related toxicity or morbidity, and they are necessary in the interpretation of other data collected. For example, body weight loss (or a reduction in body weight gain) would be expected if an animal had profound CNS depression lasting several hours.

With regard to clinical signs and observations, there are some basic definitions that should be kept in mind. Symptomatology is the overall manifestation of toxicity. Signs are overt and observable events (Brown, 1983). Symptoms are the subjective impressions of a human patient (e.g., headache) and cannot be described or reported by speechless animals (Balazs, 1970). Clinical signs can be reversible or irreversible. Reversible signs are those that dissipate as the chemical is cleared from the body or tolerance develops (Chan et al., 1982) and are generally not accompanied by permanent organ damage. Irreversible signs are those that do not dissipate and are generally accompanied by organ or tissue damage. Signs can also represent a normal biological or pharmacological response (Chan et al., 1982). For example, an antidepressant would be expected to cause decreased activity and some ataxia. These signs are generally reversible and can lead to secondary, nonspecific signs—nonspecific in that any number of agents or stimuli can evoke the same response and secondary in that they are probably not due (at least, one has no evidence to determine otherwise) to the direct action of the test article. Responses can also be abnormal, in that they are not due to a homeostatic process. The increases in serum urea and creatinine due to kidney damage, for example, are abnormal responses. These are often irreversible, but this is not always the case, depending on the repair capacity or functional reserves of the target organ. These abnormal responses may also be called primary effects because they reflect the direct action of a test article. Agonal signs are those occurring immediately prior to, or concomitantly with, death. They are obviously irreversible but not necessarily reflective of a specific effect of a test article. For example, regardless of the cause, labored breathing will occur in a moribund animal. It is, therefore, important to distinguish between signs that occur in animals that die and those that do not. It should also be kept in mind that agonal signs may mask (make it difficult or impossible) to observe other signs, including those clearly seen at lower doses.

In their simplest form, clinical observations are those done on an animal in its cage or, preferably, in an open plane, such as on the top of a counter or laboratory cart. These are

TABLE 7.4 Clinical Observation in Acute Toxicity Tests

| Organ System | Observation and Examination | Common Signs of Toxicity |
|--------------------------|--------------------------------|--|
| CNS and somatomotor | Behavior | Unusual aggressiveness, unusual vocalization, restlessness, sedation |
| | Movements | Twitch, tremor, ataxia, catatonia, paralysis, convulsion |
| | Reactivity to various stimuli | Irritability, passivity, anesthesia, hyperesthesia |
| | Cerebral and spinal reflexes | Sluggishness, absence of reflex |
| | Muscle tone | Rigidity, flaccidity |
| Autonomic nervous system | Pupil size | Miosis, mydriasis |
| Respiratory | Nostrils | Discharge (color vs. uncolored) |
| | Character and rate | Bradypnea, dyspnea, Cheyne–Stokes breathing, Kussmaul breathing |
| Cardiovascular | Palpation of cardiac region | Thrill, bradycardia, arrhythmia, stronger or weaker beat |
| Gastrointestinal | Events | Diarrhea, constipation |
| | Abdominal shape | Flatulence, contraction |
| | Feces consistency and color | Unformed, black or clay colored |
| Genitourinary | Vulva, mammary glands | Swelling |
| | Penis | Prolapse |
| | Perineal region | Soiled |
| | Color, turgor, integrity | Reddening, flaccid skinfold, eruptions, piloerection |
| Mucous membranes | Conjunctiva, mouth | Discharge, congestion, hemorrhage, cyanosis, jaundice |
| Eye | Eyelids | Ptoxis |
| | Eyeball | Exophthalmos, nystagmus |
| | Transparency | Opacities |
| Others | Rectal or paw skin temperature | Subnormal, increased |
| | Injection site | Swelling |
| | General condition | Abnormal posture, emaciation |

Source: Data from Balazs (1970).

considered passive observations. One can gain even more information by active examination of the animal, such as the animal's response to stimulation. Fowler and Rutty (1983) divide their clinical evaluation of toxicity into those signs scored by simple observations (e.g., ataxia), those scored by provocation (e.g., righting reflex), those scored in the hand (e.g., mydriasis), and those scored by monitoring (e.g., rectal temperature). Cage pans should always be examined for unusually large or small amounts of excreta or excreta of abnormal color or consistency. A list of typical observations is summarized in Table 7.4. A more extensive table has been prepared by Chan et al. (1982). Given the fact that the number of different signs displayed is not infinite and that some signs are simply easier to discern than others, most clinical signs are referable to the CNS (e.g., lack of activity), the GI tract (e.g., diarrhea), or the general autonomic nervous system (e.g., increased salivation or lacrimation). This is illustrated by an actual example set of data from acute toxicity studies summarized in Table 7.5.

Other signs can be detected by a well-trained observer but are, nonetheless, less common than those described previously. Respiratory distress can be diagnosed by examining the animal's breathing motions and listening for breathing noises. Cardiovascular signs are generally limited to pallor, cyanosis, delayed capillary refill time, and/or hypothermia. Changes in cardiac function can be difficult to detect in small animals and generally consist of "weak" or "slow" breathing.

TABLE 7.5 Summary of Clinical Observations from Actual Acute Toxicity Tests

| Drug (Route) | Indication | Acute Clinical Signs ^a |
|---------------|--------------------------|---|
| SC-37407 (PO) | Analgesic (opiate) | Reduced motor activity, mydriasis, reduced fecal output, hunched posture, convulsions (tonic), ataxia |
| SC-35135 (PO) | Arrhythmias | Reduced motor activity, lost righting reflex, tremors, dyspnea, ataxia, mydriasis |
| SC-32840 (PO) | Intravascular thrombosis | Reduced motor activity, ataxia, lost righting reflex, closed eyes, red/clear tears |
| SC-31828 (PO) | Arrhythmias | Reduced activity, dyspnea, ataxia, lost righting reflex, red/clear tears |
| SC-25469 (PO) | Analgesic (nonopiate) | Reduced motor activity, ataxia, lost righting reflex, dyspnea, convulsions (clonic) |

^a The five or six most frequent signs in descending order of occurrence.

Arrhythmias can be difficult to detect because the normal heart rate in a rodent is quite rapid. ECGs are difficult to meaningfully record from rodents on a routine basis. Therefore, the assessment of potential acute cardiovascular effect of a drug or chemical is usually restricted to a nonrodent

Acute observation record

(Days, other than study Day 1, on which no signs are observed are recorded on the long of animal observations)

| <u>Species</u> | <u>Sex</u> | <u>Route</u> | <u>Dose level</u> | <u>Animals coded*</u> | <u>Date dosed</u> | |
|------------------------|------------|--------------|-------------------|-----------------------|-------------------|---|
| Study day | | | | | | Page of notes: * <u>An code</u> <u>An ID</u> _____ _____ |
| Observations: | | | | | | |
| Time | | | | | | |
| Date | | | | | | |
| No signs observed | | | | | | |
| Reduced motor activity | | | | | | |
| Ataxia | | | | | | |
| Lost righting reflex | | | | | | Animal code for recording observations |
| Convulsions () | | | | | | |
| Mydriasia | | | | | | |
| | | | | | | |
| | | | | | | |
| | | | | | | |
| | | | | | | Read and understood _____ Date _____ |
| | | | | | | |
| | | | | | | |
| | | | | | | |
| Death | | | | | | |
| Observer | | | | | | |

FIGURE 7.11 Example of a form for recording clinical observations in acute systemic toxicity studies.

species, usually the dog. The DRF studies in nonrodents supply a place to get an initial evaluation of electrophysiological CV effects.

Given the subjective nature of recognizing clinical signs, careful steps must be taken to ensure uniformity (is the animal depressed or prostrated?) of observation so that the data can be analyzed in a meaningful fashion. There are three ways of achieving this. First, signs should be restricted to a predefined list of simple descriptive terms, such as those listed in Table 7.4 or in Appendix B. Second, if a computerized data acquisition system is unavailable, the use of standardized forms will add uniformity to the observation and recording processes. An example of such a form is shown in Figure 7.11. Third, technicians should be trained in studies (not intended for regulatory submission) using material of known toxicity, so that all personnel involved in such evaluations are using the same terminology to describe the same signs.

Animals should be observed continuously for several hours following dosing. Times of observation should be recorded as well as the actual observations. After the first day of the study, observations generally need only to consist of brief checks for remission or change of any signs and the development of new signs of morbidity. Data should be collected in such a way that the following could be concluded for each sign: (i) estimated times of onset and recovery, (ii) the range of threshold dosages, and (iii) whether signs are

directly related (primary) to the test article. An example of clinical signs provoked by a specific drug is given in Table 7.6. Incidences are broken down by dosage group and sex. These data illustrate the fact that mortality can censor (preclude) the occurrence of clinical signs. Note that reduced fecal output was a more frequent observation at the intermediate dosages because most of the animals died at the higher dosages.

Therapeutic ratios are traditionally calculated using the dose of the lowest observed adverse effect. A more sensitive therapeutic ratio could be calculated using the ED_{50} (effective dose) for the most prominent clinical sign. However, while it may be possible to describe a dosage–response curve (which may, in fact, have a different slope than the lethality curve) for a clinical sign and calculate the ED_{50} , in practice this is rarely done. It is more common for the approximate threshold dosages or no-observable-effect levels (NOELs) to be reported. A typical minimal acute toxicity study can be summarized as shown in Table 7.7.

7.3.2 Complete Acute Toxicity Testing

An example of the next-level test, the complete acute toxicity test, is given in Figure 7.12. As stated by Dayan (1983), the value of doing more than the minimal test will depend on the nature of subsequent testing. The complete protocol is

TABLE 7.6 Example of Clinical Observations Broken Down by Dosage Group and Sex in an Acute Toxicity Study of the Drug SC-37407^a

| Signs Observed | Dose Levels (mg kg ⁻¹) by Sex | | | | | | | | | |
|------------------------|---|-----|-----|-----|-----|-----|-----|-----|------|-----|
| | 0 | | 50 | | 160 | | 500 | | 1600 | |
| | M | F | M | F | M | F | M | F | M | F |
| Reduced motor activity | — | — | — | — | — | — | 5/5 | 5/5 | 4/5 | 4/5 |
| Mydriasis | — | — | — | — | 3/5 | 4/5 | 4/5 | 5/5 | 5/5 | 5/5 |
| Reduced fecal output | — | — | 5/5 | 5/5 | 3/5 | 5/5 | — | 1/5 | — | — |
| Hunched posture | — | — | — | — | — | 1/5 | 3/5 | 3/5 | — | — |
| Convulsions (tonic) | — | — | — | — | — | — | 5/5 | 1/5 | 5/5 | 3/5 |
| Ataxia | — | — | — | — | — | — | 5/5 | 4/5 | 2/5 | 1/5 |
| Tremors | — | — | — | — | — | — | 1/5 | 2/5 | 1/5 | — |
| Death | 0/5 | 0/5 | 0/5 | 0/5 | 0/5 | 0/5 | 5/5 | 4/5 | 5/5 | 5/5 |

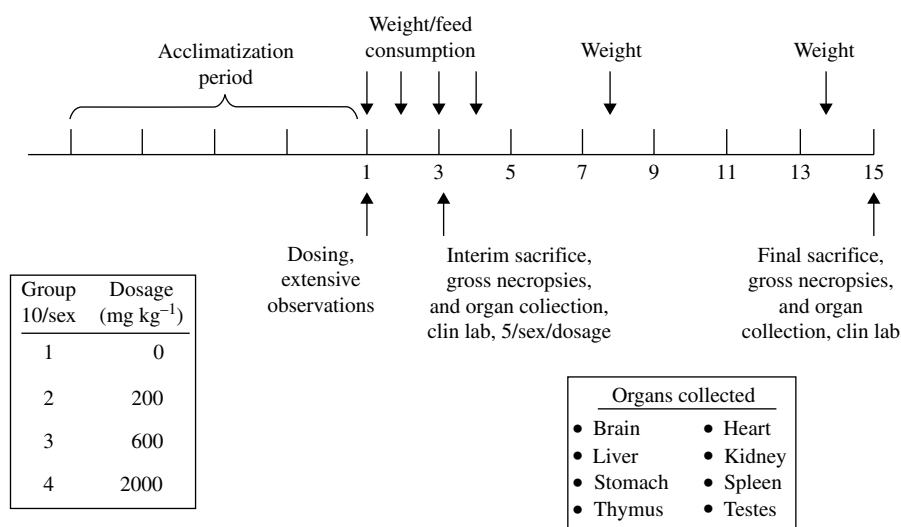
^a Signs observed in rats treated orally (no. exhibiting sign within 14 days after treatment/no. treated). A dash indicates the sign was not observed at that dose level.

TABLE 7.7 Minimal Acute Toxicity Study Summary of the Drug SC-34871

| Species (Route) | Dose (mg kg ⁻¹) | Dead/Dosed | LD ₅₀ (mg kg ⁻¹) | Signs Observed | Treatment to Death Intervals |
|-----------------|-----------------------------|------------|---|--|------------------------------|
| Rat (PO) | 2400 | 0/10 | >2400 ^a | None | None |
| Rat (IV) | 16 | 0/10 | ~67 | Reduced motor activity at 50 mg kg ⁻¹ ; convulsions, dyspnea, lost righting reflex at 160 mg kg ⁻¹ | 0–2 h |
| | 50 | 2/10 | | | |
| | 160 | 10/10 | | | |
| Mouse (PO) | 500 | 0/10 | >2400 | None | None |
| | 1600 | 0/10 | | | |
| | 2400 | 0/10 | | | |
| Mouse (IV) | 50 | 1/10 | 120 (75–200) ^b | Reduced motor activity, ataxia at 160 mg kg ⁻¹ ; tremors, convulsions, dyspnea at 500 mg kg ⁻¹ | 0–2 h |
| | 160 | 6/10 | | | |
| | 500 | 10/10 | | | |

^a Limit dosage.

^b Fiducial limits.

**FIGURE 7.12** Example of complete acute toxicity protocol.

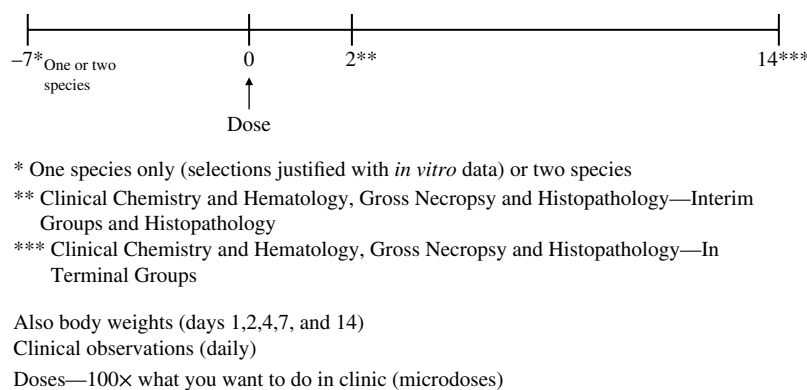


FIGURE 7.13 Exploratory IND enabling acute study. *Source:* Gad et al. (1984). Reproduced with permission of Taylor & Francis.

designed to provide for a more in-depth search for target organs than the minimal protocol. This type of study, which has been well described by Gad and coworkers (1984), is similar in design to a minimal acute toxicity study but includes feed consumption data, more frequent body weight determinations, and more detailed and frequent clinical sign assessment. Groups should consist of at least 10 animals per group; 5 per sex per dosage should then be sacrificed 24–48 h for more immediate examination of any pathological changes induced by the test article. Remaining animals will be sacrificed at the end of the 2-week period and examined for pathological changes. Blood will be collected at both sacrifices for clinical chemistry and/or hematology determinations. It should be noted that this design bears a striking resemblance to (is indistinguishable from) the design specified for an “expanded acute” study as required under the exploratory IND guidance (USFDA, 2006), as shown in Figure 7.13.

7.3.2.1 Body Weight Considerations Body weight and feed consumption are frequently determined parameters in toxicity testing. To an extent, the ability of an animal to gain or maintain weight may be considered a sensitive, but non-specific, indicator of health. While this is true in subchronic or chronic studies, its relevance in acute studies must be carefully considered. In most protocols, body weights are determined on day 1 (prior to dosing), day 7, and day 14, which are the days mandated by most regulatory guidelines. Despite being common, the design is not well founded: if an animal has not died within 7 days postdosing, it has probably recovered and its body weight may not be noticeably different from controls by day 14. A complete protocol addresses this problem by specifying more frequent body weight determinations (daily for the first 3–5 days of the observation period) so that not only can initial decreases (if they occur) be detected, but recovery can also be charted. Feed consumption measurements should be made at the same times, because it is difficult to determine the causes behind body weight changes in the absence of feed consumption data. Body weight loss accompanied by normal feed consumption

implies something very different than body weight loss (or lack of gain) accompanied by lack of feed consumption. In the absence of feed consumption data, however, changes in body weight should still be considered indicative of a change in an animal’s health status.

Yet another reason why body weight determinations are of questionable value in acute studies has to do with the statistical analysis of the data. Deaths may substantially alter group size and complicate analysis. The death of two of five animals causes a 40% decrease in group size and a substantial diminution of the power of any statistical test. In addition, the resulting data sets are censored: comparisons will often be between the control group, a dosage group where all the animals survive, and a high-dosage group where less than 50% of the animals survive to the end of the observation period. One has to question the utility of body weight changes if they occur at dosages that are acutely lethal. The data in Table 7.8 illustrate this point. Body weight changes tended to occur only at dosages that were acutely lethal. Additionally, one would suspect that the censoring of body weights in groups where death occurs is not random; that is, the animals that die are most likely those that are most sensitive, while those that survive are the most resistant or robust. This problem can be addressed by building exclusionary criteria into a protocol. For example, one could statistically analyze body weight data in groups that only had less than 50% mortality.

Minimal rather than complete protocols tend to be more common in the acute testing of pharmaceutical agents. Drugs will almost always be subjected to at least one subchronic study. Body weight and feed consumption determinations are a standard feature of such studies. Additionally, changes in body weight and feed consumption are more likely in a subchronic than an acute study because the animals are dosed continuously between body weight determinations.

7.3.2.2 Pathology Considerations One of the objectives of any well-conducted toxicity study is to identify target organs. There is some question, however, concerning the

TABLE 7.8 Examples of Body Weight Changes in Rats from Minimal Acute Toxicity Studies

| Drug (Route) | Dosage (mg kg ⁻¹) | BWT Change (g) ^a | Mortality |
|------------------|-------------------------------|-----------------------------|-----------|
| SC-32561 (PO) | 0 | 45 ± 4 | 0/10 |
| | 5000 | 39 ± 10 | 0/10 |
| | 0 | 43 ± 4 | 0/10 |
| | 500 | 43 ± 9 | 0/10 |
| | 890 | 44 ± 11 | 0/10 |
| | 1600 | 6 ± 14* | 2/10 |
| SC-36250 (PO) | 2800 | 24 ± 20* | 3/10 |
| | 0 | 38 ± 10 | 0/10 |
| | 5000 | 34 ± 10 | 0/10 |
| | 0 | 34 ± 6 | 0/10 |
| | 670 | 50 ± 8* | 2/10 |
| | 890 | 46 ± 8* | 3/10 |
| SC-36602 (IV) | 1200 | 45 ± 4 | 4/10 |
| | 1400 | 35 ^b | 9/10 |
| | 0 | 38 ± 9 | 0/10 |
| | 58 | 38 ± 3 | 0/10 |
| | 67 | 36 ± 7 | 2/10 |
| | 77 | 49 ± 5* | 3/10 |
| SC-36602 (PO) | 89 | 41 ± 7 | 7/10 |
| | 0 | 38 ± 5 | 0/10 |
| | 2100 | 41 ± 5 | 3/10 |
| | 2800 | 38 ± 5 | 7/10 |
| | 3700 | 26 ± 6 | 7/10 |

^a Mean ± standard deviation. Body weight (BWT) changes in grams for each group during the first week of the postdosing observation period.

^b Only one animal survived, so there is no standard deviation.

* Statistically different from control (0 dosage group), $p \leq 0.05$.

utility of extensive pathological assessments as part of an acute study. Gross necropsies are generally the minimum assessments requested by most regulatory bodies. Hence, minimal protocols will include necropsies on all animals found dead and those sacrificed following the postdosing observation period. An example of necropsy findings is given in Table 7.9. This table illustrates that gross necropsy observations on acute studies rarely predict the toxicity that will be seen when the chemical is given for longer periods of time. This is not surprising, because most drug-related histological lesions are the result of chronicity; that is, discernible lesions tend to result from the cumulative effect of dosages that are acutely well tolerated.

The data in Table 7.9 also demonstrate that substantial gross macroscopic findings are rare in minimal acute studies and seldom suggestive of a specific effect. There are several reasons for the lack of specificity. The first is the rather limited nature of gross observations, in that they are limited

TABLE 7.9 Examples of Gross Necropsy Findings from Acute Toxicity Studies

| Drug | Acute Gross Pathology | Subchronic Target Organs ^a |
|----------|--|--|
| SC-36602 | Distended stomach and intestine, bloody fluid in intestine, congested lung, pale liver | None |
| SC-38394 | None | Liver, testes, bone marrow, thymus, kidney |
| SC-32840 | None | Heart, stomach, kidney, bladder |
| SC-25469 | Peritonitis (IP route only) | None |
| SC-36250 | Peritonitis (IP route only) | Adrenal, liver, thyroid |
| SC-27166 | None | Liver |

^a Organs that showed any evidence of test-article-related changes in repeated-dose studies of 2 weeks or longer duration.

to broad descriptive terms (size, shape, color, etc.). The second is for animals found dead, in that it is difficult to separate the chemically associated effects from agonal and/or autolytic changes. Finally, it is difficult to come to a conclusion about the nature of a gross lesion without histological assessment.

If there are any identifiable gross lesions, they often differ between animals that die and those that survive to the end of the observation period. The reason for these differences is very simple. An animal that dies less than 24 h after chemical exposure probably has not had sufficient time to develop a well-defined lesion. As mentioned earlier, most deaths occur within 24 h. Animals that survive for the 2-week observation period have probably totally recovered and rarely have apparent lesions. Hence, the animals that provide the best chance to identify test-article-specific lesions are those that die in the region of 24–96 h postdosing. This is, in fact, one of the problems with acute pathology data—that is, comparing animals found dead with those sacrificed at a different time and comparing both to controls. As mentioned, a complete protocol, where groups of animals are sacrificed 24–96 h after dosing, at least partially solves this problem.

Many guidelines suggest microscopic confirmation of gross lesions “when necessary”; however, these are seldom done because of the autolytic nature of many of the tissues collected from animals found dead. Additionally, the practice of collecting and examining only gross lesions is difficult to justify because it does not permit in-depth comparisons. Pathological findings are most easily interpreted when the same organs are collected and examined from all animals on a test regardless of the circumstances of death.

Elsberry (1986) recommends that the GI tract, kidney, heart, brain, liver, and spleen be specifically examined routinely in acute studies. Given the timing issues discussed in the previous paragraph, the amount of effort may not be worth the result. In an attempt to address these problems, Gad and coworkers (1984) have developed a complete protocol that includes groups of satellite animals that are sacrificed 48 h after exposure and necropsied, and a standardized organ list is collected, weighed, and prepared for histological assessment. This list routinely includes the “first-line” organs: brain, thyroid, liver, kidneys, heart, and adrenals. The same organs are collected from all other animals, that is, those that die as a result of the toxicity as well as control animals. Additional tests can be included if one has a specific concern. For example, the structure of a test article may suggest that it has anticholinesterase potential. Therefore, one could include serum pseudocholinesterase determinations in the clinical laboratory package, as is frequently done for organophosphate and carbamate structures.

7.3.2.3 Supplemented Acute Studies An example of the third-level acute toxicity test, a supplemented study, is given in Figure 7.14. Such tests are rarely performed but are of use when one wishes to obtain data other than descriptive toxicity data, for example, the addition of satellite groups of animals to be dosed with a radiolabeled compound to gain pharmacokinetic information which turns a “complete” study into a “supplemented” one. Another common practice is the addition of other examinations or measurements to gain more information about a potential target organ. An example of this would be recording EKGs in rats, which is too complicated and time consuming to do on a routine basis but should be considered if the heart is a potential target organ. One way of describing such a study is that it is a complete toxicity study carrying a specific screen “piggyback.”

An excellent example of a supplemented protocol is that described by Gad and colleagues (1984). A neuromuscular screen was developed (Gad, 1982) and incorporated into

their routine acute toxicity protocol for testing nonpharmaceuticals. Doing so allowed for the more systematic and quantifiable examination of effects of the CNS than reliance on simple clinical observations. The neuromuscular screen consists of a battery incorporating standard clinical observations plus some behavioral assessment techniques already described in the literature. These are summarized in Table 7.10. This screen has been further developed to become the now regulatorily required functional observational battery (FOB). An advantage of this screen is that it uses noninvasive techniques and, therefore, will require the use of no additional animals. If an animal is displaying signs of severe CNS depression 2 h postdosing, little useful data will be gathered by examining behavior. In testing a pharmaceutical it is probably better practice to apply the neuromuscular screen on days 2, 7, and 14 postdosing in an attempt to identify more subtle or lingering effects and to chart recovery from these effects. For drugs that produce no observable CNS effect following dosing, the neuromuscular screen can be done a few hours postdosing. The more extensive and detailed nature of the data generated by the neuromuscular screen permits more confidence in the conclusion that the test article had no effect on the CNS. As will be discussed in a later chapter, the FOB has been modified for use in rats, mice, dogs, primates, and minipigs. Any suspect target organ can be investigated in a similar fashion. Depending on the invasiveness of the supplementary techniques, satellite groups may or may not need to be added to the study. Care must be taken in this regard to prevent the study from becoming too cumbersome and too complicated to conduct. It may be better to address some questions as separate studies. For this reason, one should not attempt to address more than one supplemental question in any one study.

7.3.3 Acute Toxicity Testing with Nonrodent Species

The designs described thus far for acute toxicity testing generally assume that the test species being used is a rodent. Nonrodent species are also used for acute toxicity testing.

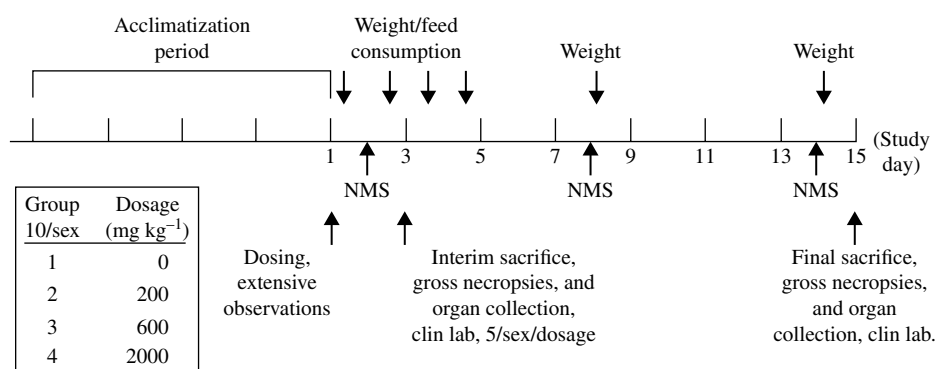


FIGURE 7.14 The design and conduct of a supplemented (or “heavy”) acute systemic toxicity study. The figure illustrates the approach to such a study when it is to serve as the definitive systemic toxicity study for some period of time.

TABLE 7.10 Neuromuscular Screen Observations

| Observation | Nature of Data Generated ^a | Correlates to which Neutral Component ^b |
|--------------------------|---------------------------------------|--|
| Locomotor activity | S/N | M/C |
| Righting reflex | S | C/M |
| Grip strength (forelimb) | N | M |
| Body temperature | N | C |
| Salivation | Q | P |
| Startle response | Q | S/C |
| Respiration | S | M/P/C |
| Urination | S | P/M |
| Mouth breathing | Q | S |
| Convulsions | S | C |
| Pineal response | Q | Reflex |
| Piloerection | Q | P/C |
| Diarrhea | S | GI tract/P/M |
| Pupil size | S | P/C |
| Pupil response | Q | P/C |
| Lacrimation | Q | S/P |
| Impaired gait | S | M/C |
| Stereotypy | Q | C |
| Toe pinch | S | S (surface pain; spinal reflex) |
| Tail pinch | S | S (deep pain) |
| Wire maneuver | S | C/M |
| Hind leg splay | N | P/M |
| Positional passivity | S | S/C |
| Tremors | S | M/C |
| Extensor thrust | S | C/M |
| Positive geotropism | Q | C |
| Limb rotation | S | M/C |

^a Data quantal (Q), scalar (S), or interval (N). Quantal data are characterized by being of an either/or variety, such as dead/alive or present/absent. Scalar data allow one to rank something as less than, equal to, or greater than other values, but one cannot exactly quantitate the difference between such rankings. Interval data is continuous data where one can assign (theoretically) an extremely accurate value to a characteristic that can be precisely related to other values in a quantitative fashion.

^b Peripheral (P), sensory (S), muscular (M), or central (C).

Many regulatory bodies require acute testing in at least one nonrodent species. The animals most often used are the dog, pig, or monkey. Veterinary products will also be tested in the target species. For example, a flea collar intended for cats must be tested in cats. While the rabbit is not technically a rodent, it is the species of choice for a variety of tests for assessing acute oral or intravenous toxicity and is considered a rodent for regulatory purposes. The section is written with the dog and monkey in mind. Clearly, there are some profound differences between these species and rodents with regard to handling, husbandry, and dosing. Here we focus on the design differences in toxicity testing in large species.

For financial, procurement, and ethical reasons, acute systemic toxicity tests on nonrodents are not performed using traditionally designed animal-intensive protocols.

The minimal acute study requires 30–50 animals. Complete and supplemented studies will usually require even more. At a cost of \$1500.00 per beagle dog, \$2000.00 per minipig, or \$5000.00+ per nonhuman primate, the animal costs alone are enough to make such studies with these species prohibitively expensive. Vivarium space and husbandry costs are also much higher with nonrodent species than with rodents. Nonrodents also require a much longer prestudy quarantine period than rodents: at least 6–8 weeks for dogs and pigs and 18–24 weeks for monkeys. Treatment during the quarantine period is more extensive than that given to rodents. The animals should be given frequent physical examinations including complete clinical laboratory panels and appropriate tests for common illnesses and parasites. Special care must be taken with monkeys not only because they can be vectors of human disease, but also because they can contract human diseases and a sick animal can compromise study outcome. All these factors dictate that these animals should be sparingly used. Hence, it is most common to study acute systemic toxicity in nonrodent animals using a pyramiding dosage design. The typical study will consist of two treated animals per sex and two control animals per sex for a total of eight animals. A typical protocol is shown in Figure 7.15.

The use of fewer but larger animals permits more extensive observations of each individual. Following each dose, animals can be given complete physical examinations that include palpations, behavioral checks, spinal reflex checks, pupillary light reflexes, respiration rate, ECG recording, and rectal temperature measurement. Blood samples can also be collected following each dose to determine standard clinical chemistry and hematology profiles. Hence, while fewer animals are used with the pyramiding dosage protocol, more information per animal is collected.

The small number of animals used in a pyramiding dosage study makes it difficult to do standard statistical comparisons. This difficulty can be overcome to a certain extent by taking advantage of two design aspects of the pyramiding protocol. First, pretreatment data can and should be obtained on all animals for all parameters examined or determined. In-study comparisons should be made both to pretreatment data and to concurrent control animals. Such comparisons can be made not only on the basis of absolute numbers but also on the magnitude of any changes from pretreatment values. Second, all animals should be measured repeatedly throughout the study. Hence, to reflect a true drug-related effect, the magnitude of change should increase following each dose (though one must be aware of the potential for the development of tolerance as induction of metabolism). This is, in fact, the only way one can make any dosage–response or threshold conclusions using the pyramiding protocol.

Seldom are drugs tested in nonrodent animals via routes other than the intended or likely routes of human exposure. Hence, the most common routes in these types of protocols are oral, intravenous, and respiratory. Rarely is a test article

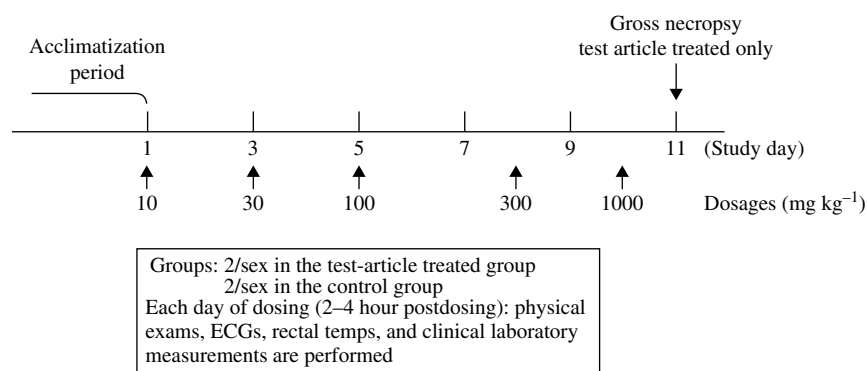


FIGURE 7.15 Example of pyramiding dose study for acute toxicity testing in a nonrodent species.

given to nonrodent species by the intraperitoneal route. Routes are discussed elsewhere in detail (Gad and Chengelis, 1999), but some discussion is appropriate here because of design considerations. Test articles are normally given orally by capsule to dogs and pigs and by gavage to monkeys. Nonrodents have to be restrained if dosed by gavage, making the process very labor-intensive. This is minimized by the small number of animals specified by the pyramiding protocol. In contrast, because of the differences in size, it is much easier to deliver a test article intravenously to nonrodents than to rodents. For topical studies, the rabbit is the nonrodent choice because it is easier to prevent a rabbit from grooming the delivery site, and considerably less material is required to deliver a comparable dose to a rabbit than a dog or pig. Acute dermal studies are not, however, usually done with a pyramiding study design but rather as a limit dose study.

The biggest problem with the pyramiding protocol is the development of tolerance. If no toxicity is observed, the chemical could be innocuous or animals could have developed tolerance during the study. The escalating dosage feature of the pyramiding protocol is an excellent vehicle for fostering the development of tolerance. One can check this by dosing additional naïve animals at the limit dosage to confirm, as it were, a negative result. Another problem, which is most peculiar to the dog, is emesis. Oral administration a large amount of almost any material will cause a dog to vomit. This is always somewhat of a surprise to toxicologists whose prior experience is primarily with rodents, which cannot vomit. One should pay close attention to dogs the first hour after capsule delivery. If the dog vomits up most of the dose, the actual dosage absorbed could be grossly overestimated. This can be a particular problem if one is using the results of a pyramiding dosage study to set the dosages for a repeat-dose study. Dogs can develop tolerance to the emetic effect of a set dosage. When this occurs, absorption and resulting blood concentrations of a test article can increase dramatically, resulting in more florid toxicity than expected on the basis of the pyramiding study. Another problem is that emesis can result in secondary electrolyte

changes—especially decreases in chloride—that may be mistaken for a direct effect of the test article. If emesis is a severe problem, one can study toxicity in a different nonrodent species or divide larger dosages into two or three divided dosages on the day of dosing.

As with traditionally designed rodent studies, the pathology component of pyramiding studies usually consists of gross necropsies followed by (when appropriate and necessary) histological assessment of gross lesions. Unfortunately, this study design does not permit the establishment of a dose–response relationship with regard to gross necropsy findings. In addition, the small number of animals makes definitive conclusions difficult. Usually, gross lesions are defined in absolute terms with few comparisons to control animals. Suspected target organs should be further investigated in subsequent subchronic studies or in rigorous and specific mechanistic studies. Because of the limited value of the pathology data generated by the pyramiding protocol, control animals should not be terminated but rather should be saved for reuse.

7.3.4 Factors that Can Affect Acute Tests

Many investigations into the sources of variability in acute toxicity testing have been conducted, and these have been reviewed by Elsberry (1986). The factors causing the greatest interstudy variation included lack of specifications for sex, strain, age, and weight range. When clearly defined, detailed protocols were used, interlaboratory variation was found to be minimal. Hence, it is equally important that the details of the protocol be well described and followed. It is not appropriate to draw dosage–response conclusion by comparing groups that differ substantially in age or that have been fed, fasted, or otherwise manipulated differently. Guidelines for standardization of acute toxicity testing were proposed by the interagency regulatory liaison group (Interagency Regulatory Liaison Group (IRLG) Office of Consumer Affairs, 1981; Elsberry, 1986). These do not differ markedly from those mandated by the Toxic Substance Control Act of 1986 (Gad and Chengelis, 1999).

7.3.4.1 Number, Size, and Sex of Dosage Groups The precision with which lethality and signs of toxicity are described will depend on the number of groups (hence, dosages) and the number of animals in each group. Between 1940 and 1980, the standard was to use from four to six dosages with 10 animals per dosage. The current emphasis is on limiting the number of animals used for acute testing, particularly with recognition of the limited value of “precise” lethality data (Gad and Chengelis, 1999). Retrospective analyses by DePass (1989) and Olson et al. (1990) have demonstrated that decreasing group size to two to three animals generally has little impact on overall study results. Hence, the number and size of dosage groups will depend, to an extent, on the methods of statistical analysis. The classic statistical methods for analyzing lethality data (or, indeed, any quantal dosage–response data) were published between 1930 and 1960 and have been extensively reviewed by Armitage and Allen (1959) and Morrison et al. (1968). These methods are mentioned here with regard to the demand they make on protocol design—specifically, the number of dosage groups, the spacing of the dosages, and the number of animals per dosage group. The probit and moving average methods are the most commonly used today. In general, all methods of calculation and computation are more precise if the designs are symmetrical (i.e., the dosages are evenly spaced and the group sizes are equal). The probit method, first developed by Bliss (1935, 1957) and later refined by Finney (1971, 1985), is considered to be the most precise, but it requires at least two groups of partial responses (i.e., mortality >0, but <100%). This may require dosing more than three groups until this criterion is met. It also deals ineffectively with groups that had either 0 or 100% mortality. (The most common correction for these groups is to substitute 0.1% for 0 and 99.7% for 100%.) The moving average method, first described by Thompson and Weil (1952), does not require partial responses, deals effectively with complete responses, and, therefore, can produce an acceptable estimate of an LD_{50} with as few as three groups of three to five animals each. The moving average method can also be used to design the experiment. Groups can be dosed in a sequential fashion as in a pyramiding study, with each step dictated by the moving average method. Once evidence of toxicity is observed, further dosing is discontinued. This method requires that the dosages be separated by a constant geometric factor (e.g., 2, 4, and 8 mg kg^{-1}) and that groups be of equal size. Weil (1952), and later Gad and Chengelis (1999) and Gad (2007) have published tables that allow for the easy calculation of the LD_{50} using $K=3$ (where K =the number of dosage groups minus 1). The LD_{50} for $K<3$ can be easily calculated without the aid of tables. In addition, methods for estimating the confidence limits of this calculated LD_{50} have also been published (Gad, 2007). Traditionally, the moving average method has not been extensively used because, while it yielded an estimate of the

LD_{50} , it did not give the slope of the (probit transformed) lethality curve. However, Weil (1983) also published a method for calculating a slope from the same data. Hence, an estimate of the LD_{50} and slope can be obtained from as few as three groups of three to five animals per group, provided that at least one group shows a response less than 50% and another shows a response greater than 50%.

The Litchfield and Wilcoxon (1949) plotting method was once commonly used. It is certainly a valid method, and it poses no more restrictions on study design than those imposed by the probit method. The Litchfield–Wilcoxon method has become a victim of technology as modern, handheld calculators and the ready availability of simple computer programs have made other methods more convenient to run. However, at least one software company has adopted the Litchfield–Wilcoxon method for its acute toxicity protocol package.

While much has been written about the influence of gender on acute lethality, most authors now agree that there are seldom any substantial differences in the LD_{50} due to sex (DePass et al., 1984; Gad and Chengelis, 1999). In those instances where there is a sex-related difference, females tend to be more sensitive than males (~85% of the time). If one is willing to accept this amount of uncertainty, only one sex needs to be tested. Alternatively, one could use only two to three animals per sex per dosage group. Schultz and Fuchs (1982) have demonstrated that, by pooling sexes, there are seldom any substantial differences in the LD_{50} calculations between groups consisting of five per sex versus three per sex. If there are no substantial differences between sexes (i.e., 70% mortality for males and 80% for females at a dosage), the results from pooling the sexes can provide a pooled LD_{50} . For most safety considerations, an LD_{50} derived on this basis will be acceptable and will result in the use of fewer animals.

7.3.5 Selection of Dosages

In setting dosages for acute studies, a few commonsense rules have to be applied. First, the intrinsic biological and chemical activity of the test article must be considered. Zbinden and Flury-Roversi (1981) have documented several cases where lethality was of no biological relevance. The oral lethality of tartaric acid, for example, is due to the caustic action of a high concentration of acid in the GI tract. In these instances, limit tests are more appropriate tests. Additionally, it is uncommon that a completely unknown chemical will be tested. Factors such as known pharmacological profile and chemical or physical characteristics including molecular weight, participant coefficient and the like, and the toxicity of related chemicals should be considered. For example, it is likely that a polymeric, poorly soluble molecule will not be bioavailable at an initial dosage of 100 mg kg^{-1} . A full understanding of all available data will permit one to pick dosages with more confidence and, thereby, save both time and resources.

TABLE 7.11 Sample Data Sets: LD₅₀ Calculations Using Fewer Dosages^a

| SC-27166 | | Theophylline | |
|----------------------------------|-----------|----------------------------------|-----------|
| Dosage (mg kg ⁻¹) | Mortality | Dosage (mg kg ⁻¹) | Mortality |
| 100 | 0/10 | 280 | 0/10 |
| 180 | 0/10 | 320 | 3/10 |
| 240 | 4/10 | 370 | 5/10 |
| 320 | 7/10 | 430 | 9/10 |
| 560 | 9/10 | 500 | 10/10 |
| 1000 | 10/10 | 670 | 10/10 |
| LD ₅₀ = 300 | | LD ₅₀ = 300 | |
| <i>Using every other dosage</i> | | | |
| 100 | 0/10 | 280 | 0/10 |
| 240 | 4/10 | 370 | 5/10 |
| 560 | 9/10 | 500 | 10/10 |
| LD ₅₀ = 290 | | LD ₅₀ = 290 | |

^a Adult male mice; drugs given by gavage.

Second, no protocol will yield high-quality data if all dosages given cause 100% lethality. Therefore, one is best advised to pick widely spaced, rather than closely spaced, dosages. In general, the best dosage regimen includes a dose that will definitely produce a high incidence of severe toxicity, another that will produce marginal toxicity, and one that will produce toxicity in between. If this pattern is obtained, adding more groups does not generally change the results. This point is illustrated by the data in Table 7.11. For two drugs, an LD₅₀ of 300 mg kg⁻¹ was obtained using six groups of 10 mice each. Essentially the same result was obtained if the second, fourth, and sixth groups were eliminated and not used in the calculations. Behrens (1929) noted this phenomenon almost 60 years ago.

Widely spaced dosages also decrease the likelihood of nonnormotonic data, where mortality does not necessarily increase with dosage (see Table 7.12). This can occur when the test chemical has a shallow dose-response curve and the group size is small (three to four animals). While it is possible to calculate an LD₅₀ from such data, the slope and confidence limits will be inaccurate. Nonnormotonic data can also occur if the lethality is indeed biphasic. If one suspects that this is occurring, additional dosages should be examined. For safety considerations, only the first part of the curve, the lowest LD₅₀, is of importance.

7.3.5.1 Timing The greatest precision in any lethality curve is obtained when the number of experimental variables is kept to a minimum. Hence, it is best if all the animals used for determining a specific curve are dosed on the same day and, if possible, at the same time of day, which limits age-related and diurnal effects. If a total of only 15 animals are being dosed, this is not a difficult task for a single well-trained technician. However, if the test substance is of unknown lethality, it is imprudent to deliver all doses on the

TABLE 7.12 Sample Data Sets: Homogeneous versus Heterogeneous Data

| Homogeneous ^a (Normotonic) | | Heterogeneous ^b (Nonnormotonic) | |
|--|-----------|---|-----------|
| Dosage (mg kg ⁻¹) | Mortality | Dosage (mg kg ⁻¹) | Mortality |
| 300 | 0/20 | 620 | 0/10 |
| 600 | 1/20 | 1600 | 2/10 |
| 800 | 10/20 | 2100 | 8/10 |
| 1000 | 17/20 | 2800 | 5/10 |
| | | 3700 | 8/10 |
| | | 5000 | 8/10 |

^a Data from study of SC-31828, using adult rats of both sexes.^b Data from study of SC-3894, using adult male rats.

same day. It is common practice for a single dosage group to be treated on the first day of an experiment and the dosages for the second and third groups to be adjusted pending the results of the first group. Generally, most acute deaths will occur within 24 h of dosing. Delayed deaths (those occurring more than 24 h after dosing) are relatively rare and generally restricted to the 72-h period following dosing (Gad et al., 1984; Bruce, 1985). Hence, waiting for 24 h between doses will generally yield sufficient data to allow the choice of the next dosage. For example, if all but one of the animals dosed in the first groups die, there is no doubt that the next dosage should be adjusted downward considerably, whether or not the final animal eventually dies. All the dosing for a single curve can be completed in 3 days. If a test article is being tested in traditional protocols (with two species, two routes, separate sexes), the two initial groups by a route can be treated on the first day of the dosing period and the second route initiated on the next day. Subsequent dosages can be adjusted on alternate days. Little real impact on the results will occur if there are 2–3 days between dosing sets. After that, however, the increasing age of the animals may result in a change in sensitivity. As reviewed by Balazs (1976), for example, the ratios of the LD₅₀s obtained in adult animals to the LD₅₀s obtained in neonates can vary from 0.002 to 160. One can use longer observation periods between dosing days if separate animal orders are timed for delivery to ensure that all animals dosed are closer in age. As a rule of thumb, the animals should not differ in age by more than 15%; hence, the younger the animals, the smaller the age window.

7.4 SCREENS

Screens are generally not safety studies in the regulatory sense. These are the studies done, as the name implies, to examine several chemicals in order either to select those with the most desirable properties for development or to eliminate those that have undesirable properties. There is nothing novel about screening; the process has been an integral part of

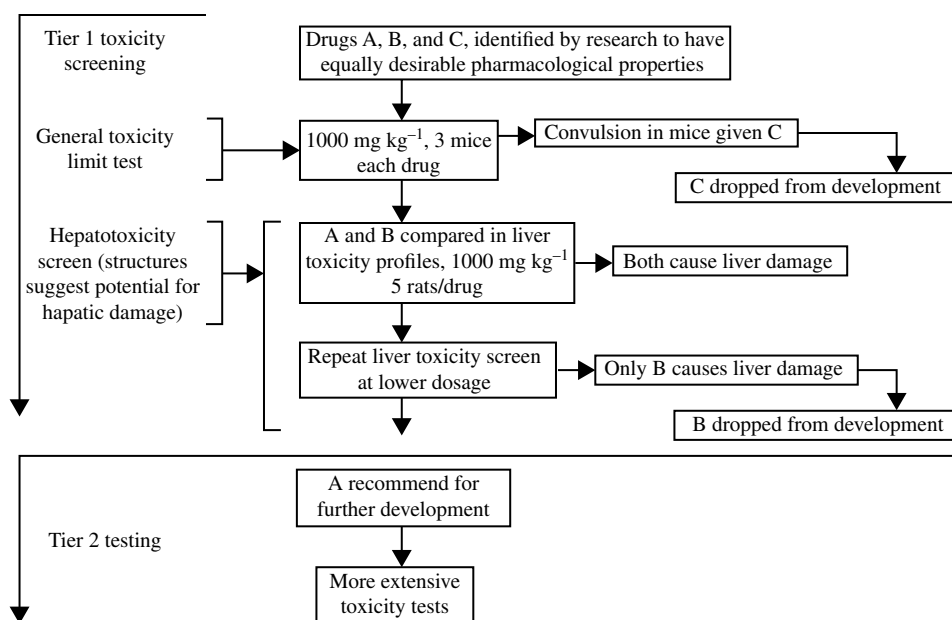


FIGURE 7.16 Example of use of screens in selecting drug candidates for development.

pharmaceutical research for decades (Irwin, 1962). In a pioneering paper, Smyth and Carpenter (1944) described a screening process for gathering preliminary toxicity data for a new chemical. In their discussion they clearly state the underlying rationale for toxicity screening:

Opinions upon the toxicity, hazards of manufacture, and fields for safe use must be expressed regarding many chemicals which will never be produced in quantity. Large expenditures of time and money upon securing these basic conclusions is not justified. Later, when a few of the new compounds are obviously going to be made commercially, more detailed studies can be undertaken.

Screens are designed for speed, simplicity, and minimal resource expenditure. They are designed to answer positive, single-sided questions. For example, the lack of an effect in an initial screen does not mean that toxicity will not be manifested with a different formulation or in a different species. It is for this reason that a screen should not, as stated by Zbinden et al. (1984), be seen as replacements for thorough safety testing. An acute toxicity screen can be the first leg in a decision tree or tier testing process for selecting a chemical or drug candidate for development. An example of this process is given in Figure 7.16.

7.4.1 General Toxicity Screens

There are two types of acute toxicity screens. In the general toxicity screen, animals (often, for economic reasons, mice) are exposed to two or three predefined dosages of chemical. No more than three mice per dosage are necessary and no control group is required. An example of this type of

protocol is shown in Figure 7.17. The animals are carefully observed for mortality and obvious signs of toxicity, such as convulsions, but no attempt should be made to quantify the severity of a response. There is seldom any need to have an observation period of more than 4–5 days. Because of the quantal nature of the data, interpretation is straightforward. There are four possible outcomes: (i) no death or signs of toxicity seen at dosages up to $X \text{ mg kg}^{-1}$, (ii) no deaths but evident signs of toxicity seen at $X \text{ mg kg}^{-1}$, (iii) deaths but no evident signs of toxicity at $X \text{ mg kg}^{-1}$, and (iv) deaths and evident signs of toxicity both occurred at $X \text{ mg kg}^{-1}$. General toxicity screens may also provide the preliminary information for picking the dosages for more definitive acute studies.

There are two ways to apply the data from toxicity screens to the development of a drug or chemical. On a relative basis, the drugs under consideration can be ranked according to screen results, and the one that appears to be the least toxic can be chosen for future development. Alternatively, decisions can be made on an absolute basis. All candidates that are positive below a certain dosage are dropped, and all those that are negative at or above that dosage will continue to the next tier of testing. If absolute criteria are used, the screen need be done only at the critical dosage. If only one dosage is examined, the test is a limit test. A limit test of this kind is the simplest form of toxicity screen, and depending on the nature of subsequent testing, it is highly recommended.

Fowler and his colleagues (1979) have described a rat toxicity screen (illustrated in Figure 7.18) that is more extensive and detailed than the one shown in Figure 7.17. It includes two rounds of dosing. In the first round, up to 12 rats are (singly) exposed to six different dosages by two different routes for the purpose of defining the MTD. In the second round of dosing, 16 rats are dosed at two-thirds

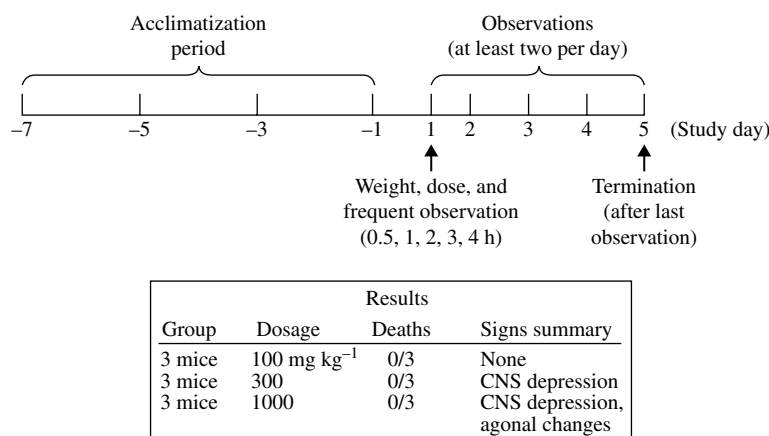


FIGURE 7.17 Example of general toxicity screen.

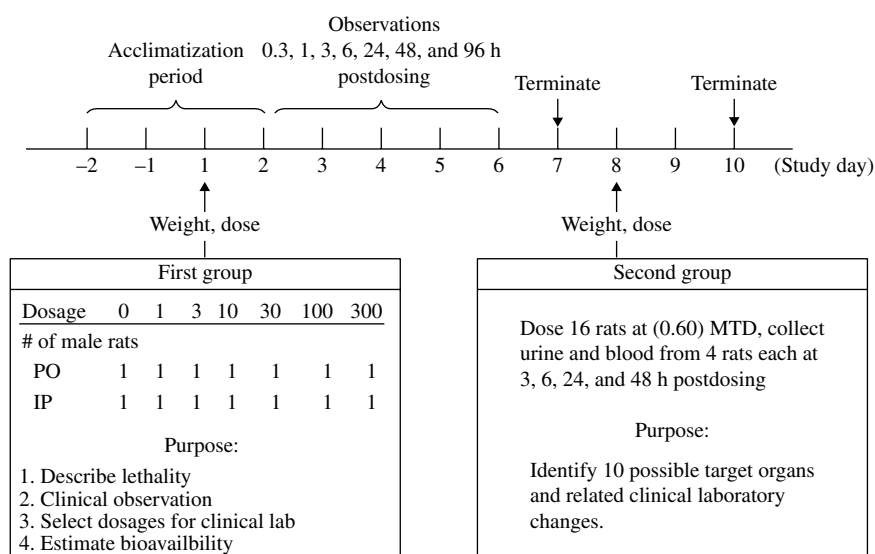


FIGURE 7.18 Example of rat toxicity screen for drugs.

(0.66) of the MTD and sacrificed on a serially timed basis for blood sample collections to determine test-article concentrations and for clinical laboratory tests. These features make this design too complicated, time consuming, and expensive to run as an initial screen. This design is better suited as a second-tier screen to provide a more extensive follow-up study for a more limited screen. Fowler et al. contend that their screen disclosed most toxicity uncovered by more conventional studies. This screen was most successful in defining acute CNS, liver, or kidney toxicity (Fowler et al., 1979). Lesions that require long-term exposure, such as those generally involving the eyes, may not be detected in this type of screen.

Up/down or pyramiding designs can be used for general toxicity screens, but this is not a common approach because of the time involved. In addition, if several chemicals are being compared, an up/down study where death occurs at

different dosages can be complicated to run. It is much easier to test several chemicals at the same time using a limit test design. Because only individual animals are dosed, these designs can be used when there is a very limited amount of test article available and/or there are few prior data on which to base an expected toxic dosage.

Hazelette and colleagues (1987) have described a rather novel pyramiding dosage screen that they term the rising dose tolerance (RDT) study (illustrated in Figure 7.19). The study, which uses a subacute rather than an acute dosing regimen, can also be used as a range-finding study design. The rats are exposed for 4 days to the initial dosage followed by 3 days of recovery before the next 4-day dosing period at the next highest dosage. This process is repeated for the three dosing cycles. Plasma and urine samples are collected for clinical chemistry and urinalysis, as well as test-article determinations. Necropsies and microscopic examinations

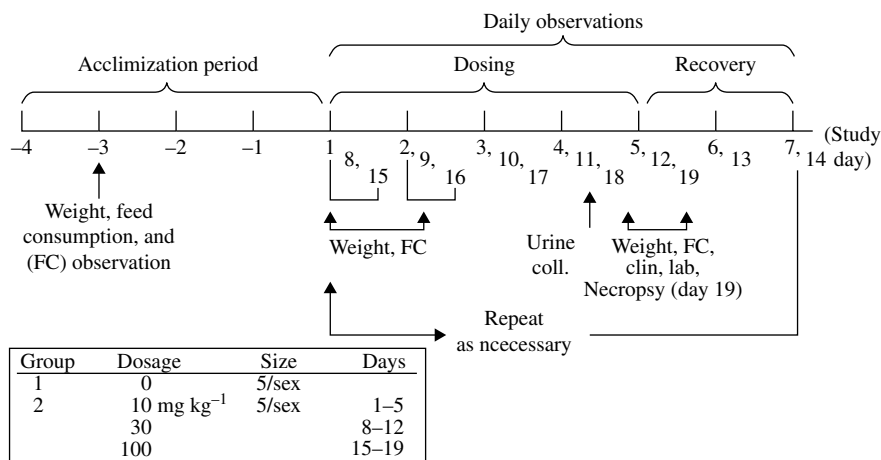


FIGURE 7.19 Example of rising dose tolerance test (no pharmacokinetic (PK) groups).

are performed. While this study design is novel, it appears to provide considerable acute data. It is also possible that this design could generate sufficient data to plan a pivotal sub-chronic study and, therefore, replace a traditional 2-week study, resulting in considerable savings of time and animals. This is not a simple study and, therefore, inappropriate as an initial screen, but it would appear to be appropriate for a second-tier test.

7.4.2 Specific Toxicity Screening

The second type of acute toxicity screen is the specific toxicity test. This type of test is done when one has a specific toxicological concern, for example, when prior structure–activity data suggest that a family of chemicals can be hepatotoxic. A screen to select the chemical with the least hepatotoxic potential is then in order. These tests are also done, as described by Zbinden (1984), to look for a specific toxicological effect that may be easily overlooked in a route safety study. Zbinden gives, as an example, screens that are designed to detect specific lesions to the hemostatic process. As pointed out by Irwin (1962) over three decades ago, such tests have their greatest power if more than one measure of specific target organ toxicity is used. Dayan (1983) refers to this technique as a matrix of measurements. In a liver toxicity screen, for example, liver weights (both absolute and relative), gross necropsy examinations, and a battery of serum enzyme assays should all be part of the protocol. As a general rule, because of the time and expense involved, screens should be designed to minimize the use of histopathological techniques. The dosages can be standardized or set on the basis of the results of a generalized toxicity screen. Specific toxicity screens can be the next-level test in the decision tree process for selecting a candidate for development (as illustrated in Figure 7.16).

The number of animals and the number of dosages are highly dependent on the type of data gathered. A few rules of

thumb should be followed: (i) keep it lean: each additional group, animal, or test article added to a protocol makes the study exponentially more difficult to conduct—simplicity is one of the most important features of a screen; (ii) the more parameters examined, the fewer the number of animals required; and (iii) if normal limits of a test parameter are relatively broad (e.g., serum glucose), more animals will be required than if the parameter is normally tightly controlled (e.g., prothrombin time). In general, 3 is the minimum and 10 is the maximum number of animals required per group. Further, if a single chemical is examined per study, no more than three groups will be required. If more than one chemical is included in the study, then a single dosage (limit) group per chemical is the best design.

Strictly speaking, an acute toxicity study is conducted to examine the effect of a single dose of a single compound. In designing specific toxicity screens, however, deviation from this principle is permissible if it increases screen sensitivity. For example, the sensitivity of mice to many indirect hepatotoxins will be enhanced by prior treatment with phenobarbital. Hence, the sensitivity of a hepatotoxicity screen will be enhanced if the mice are pretreated for 3 days with phenobarbital.

The screen should be validated for consistency of response with both positive and negative control articles. A positive control article is one that is known to reliably produce the toxic syndrome the screen is designed to detect. Concurrent control groups are not required with each replicate. Rather, control groups should be evaluated on some regular basis to ensure that screen performance is stable. Because a screen relies on a biological system, it is not a bad idea to test the control benchmarks, particularly the positive ones, on a routine period basis. Not only does that give one increased confidence in the screen, but it also provides a historical base against which to compare the results of new test articles. Zbinden and colleagues refer to the positive control as the reference compound, and they have discussed

some of the general criteria to be applied in the selection of these compounds (Zbinden et al., 1984). Any changes to the design should trigger revalidation. Any analytical methods should be subjected to precision, accuracy, sensitivity, and selectivity (PASS) validation.

Interpretation of specific toxicity screen data is not as straightforward as that of a general toxicity screen. This is because the data will often be continuous, following a Gaussian, or normal, distribution. This has two ramifications. First, for results around the threshold, it may be very difficult to differentiate between positive and negative responses. Second, for any one parameter, there is a real chance of false statistical significance (type I errors), especially if small numbers of animals are used. This occurrence is one of the reasons why specific toxicity screens should include the determination of more than one variable, since it is unlikely for multiple false positives to occur in the same group of animals. An undetected false positive could lead to the dropping of a promising candidate in error. False negatives, by contrast, may not be as critical (other than the time lost and the resources spent), because extensive subsequent tests should lead to the more complete description of the test article's toxic properties.

The problems described in the preceding paragraph assume that the screen will include a traditional ("negative," or vehicle) control group and that the data from the treated groups will be compared to those of the control group by standard methods. These problems will be minimized if no control group and, therefore, no traditional statistical comparisons are included. In addition, a decrease in the number of animals used simplifies the study. Data can be interpreted by comparison to a historical control data base as described by Zbinden (1984). The threshold, or test criterion X_c , is calculated according to the following formula:

$$X_c = m + (z)s$$

where m is the population mean, s is the standard deviation, and z is an arbitrary constant. This formula is essentially a method of converting continuous data to quantal data: it is used to determine if individual animals are over the test threshold, not if the group mean is over the threshold. Analysis of screening data by comparison to experience (i.e., historical control data) and an activity criterion are discussed in greater detail in Chapter 4 and by Gad (2007). The higher the z value, the lower the probability of a false positive but the lower the sensitivity of the screen. Again, including multiple parameters in the screen helps alleviate this problem. Zbinden has proposed a ranking procedure in which various levels of suspicion (LOS) or a level of certainty (LOC) is assigned to the result of a toxicity screen. This is simply a formalized fashion of stating that the more animals that respond, and the greater the severity of the response, the more certainty one has in drawing a conclusion. If relative

comparisons are being made, this system provides a framework for ranking test articles and selecting those to continue to the next tier of testing.

With regard to specific toxicity screening, behavioral toxicity screening is an area currently generating a great deal of interest. As reviewed by Hopper (1986), there are several reasons for this interest. First, the Toxic Substance Control Act of 1976 legislatively recognized behavioral measures as essential to examining chemicals for neurotoxic potential. Second, the structure and function of the CNS are not amenable to traditional methods of examination, in that profound behavioral changes can be induced in the absence of any detectable morphological lesions. This large and somewhat controversial subject is outside the scope of this chapter. Specific screening strategies are presented and critically discussed by Hopper (1986). Other recommended references to consult for different perspectives on acute toxicity testing are Rhodes (2000), Brown (1980), and Arnold et al. (1990).

7.5 PILOT AND DRF STUDIES

Between acute studies and the repeat-dose (14 or more typically 28 days) studies (described in Chapter 10), there are studies which are necessary to be able to set doses for the GLP studies required by regulation (ICH M3(R2) and its regional modifications) to be able to initiate initial clinical studies. These studies have evolved over time—until the mid-1990s, they were always miniature (usually 14-day) versions of the 30-day GLP repeat-dose studies that they were meant to enable the design and dose selection. Concerns for animal usage, costs, and pressures of time have changed this.

The DRF-style studies are not required by regulation but rather by sound scientific practice. As such, it should be noted that while the results must be included in the documents submitted to regulatory authorities, they do not have to be performed in compliance with GLPs. Their results are frequently used not just to select doses but also to justify dose selection, usually by demonstrating either that higher doses would be intolerable or (especially if TK has been performed) that there is a saturation of absorption effect making the use of higher doses meaningless, and to identify any needs for modification in the design of the studies. The studies should be conducted with dosing by the route and regimen (dose frequencies) intended for clinical studies. The formulations used should be those to be employed in the 14- or 28-day studies, but the actual drug substance (API) lot does not need to be the same (especially as the 28-day "pivotal" tox doses should use clinical lot material).

For nonrodent species (dog, pig, or primate), dose levels are set based on the maximum tolerated dose identified in the acute pyramiding dose study. The design of that study is presented in Table 7.13.

TABLE 7.13 Seven-Day Nonrodent Pilot Toxicology Study

| | | |
|---|-------|---------|
| Animal: 9 males, 9 females | | |
| Study design | Males | Females |
| Vehicle control | 3 | 3 |
| Low dose | 3 | 3 |
| High dose | 3 | 3 |
| Dosing: once daily all animal for 7 days | | |
| Observations: (mortality/moribundity) twice daily | | |
| Clinical examination: daily after dosing | | |
| Body weights: before start and on days 1, 2, 3, 5, and 7 | | |
| Food consumption: before start and on days 1, 2, 3, 5, and 7 | | |
| Physical examinations: conducted by a staff veterinarian on all animals prior to initiation of compound administration and at study termination | | |
| Electrocardiograms: all animals prior to initiation of compound administration and at study termination | | |
| Clinical pathology: hematology, clinical chemistry, and urinalysis evaluations on all animals pretest and at study termination | | |
| Necropsy: all animals on day after last dose | | |
| Organ weights: adrenals, brain, heart, kidneys, liver, lungs, ovaries with oviducts, pituitary, prostate, salivary glands, seminal vesicles, spleen, thyroid with parathyroid, thymus, testes, uterus | | |

TABLE 7.14 Seven-Day Rodent Pilot Toxicology Study

| | | |
|--|-------|---------|
| Study design: | | |
| Main study | Males | Females |
| Vehicle control | 5 | 5 |
| Low dose | 5 | 5 |
| Mid dose | 5 | 5 |
| High dose | 5 | 5 |
| Dosing: once daily all animal | | |
| Observations: (mortality/moribundity) twice daily | | |
| Detailed clinical observation: daily | | |
| Functional observational battery: after first dose | | |
| Body consumption: weekly | | |
| Clinical pathology: hematology, clinical chemistry, and urinalysis evaluations on all surviving main study animals at termination (see attachment) | | |
| Toxicokinetics: blood may be collected after first and last dose at six time points | | |
| Necropsy: all main study animals, toxicokinetics animals euthanized and discarded | | |
| Organ weights: adrenals, brain, heart, kidneys, liver, lungs, ovaries with oviducts, pituitary, prostate, salivary glands, seminal vesicles, spleen, thyroid with parathyroid, thymus, testes, uterus. | | |

The design for the rodent pilot study is shown in Table 7.14. Dose levels are set based on results seen in the acute study with the middose typically being the maximum tolerated dose seen in that study.

TABLE 7.15 Correlation of Clinical Signs and Target Organs

| Clinical Signs | Target of Toxicity | Potential Correlating Data |
|------------------------------|---------------------------------|----------------------------|
| Piloerection | Nonspecific | Body weight gain |
| Hypoactivity | | Food consumption |
| Ataxia | | FOB observations |
| Pustulation | | |
| Abnormal gait | | |
| Chromodacryorrhea | Nervous system | |
| Bulging eyes | | |
| Straub tail | | FOB observations |
| Tremors | | |
| Convulsions | | |
| Stereotypy | Kidney | |
| Anuria | | Relative kidney weights |
| Hematuria | | Gross necropsy |
| Polyuria | | Clinical chemistry finding |
| | | Urine chemistry |
| Pallor | Cardiovascular or Hematopoietic | FOB observations |
| Hypothermia | | Relative heart weight |
| | | Gross necropsy observation |
| | | Clinical chemistry |
| | | Hematology |
| Cyanosis | | FOB observations |
| Lacrimation | Autonomic nervous system | |
| Salivation | Nervous system/nonspecific | |
| Rhinorrhea | Gastrointestinal tract | |
| Aggression | | |
| Passivity | | |
| Resistance to handling | | |
| Vocalization | | |
| Hyperactivity | Gastrointestinal tract | Gross necropsy observation |
| Abnormal response to stimuli | | Stomach contents |
| Ptosis | | Change in body weight |
| Miosis | | Food consumption |
| Mydriasis | Respiratory tract | |
| Absent/few feces (droppings) | | |
| Diarrhea | | |
| Rapid breathing | | |
| Shallow breathing | | |
| Noisy breathing | Respiratory tract | Gross necropsy observation |
| Irregular breathing | | |
| Gasping | | |
| Emaciation | | |
| | Nonspecific | Food consumption |

Appendix B and Gad and Chengelis (1999) should be consulted for a lexicon of clinical signs. Chapter 6 presents a more complete discussion of the subject of this table.

As time and technology permit, as much information as possible should be collected in these studies. The hematology, clinical chemistry, and urinalysis measures are as described in Chapter 10. Toxicokinetic and organ weight data should be collected in these studies (Ballantyne, 2009; Denny and Stewart, 2013). Clinical signs should be obtained by careful observation of the animals over the course of the study, particularly shortly (usually 2 h) after the initial dose. Table 7.15 presents a correlation of such clinical signs and highly probable associated target organs.

In rodents, FOB observations are increasingly performed after the initial animal dosing to both add rigor to the collection of clinical signs (Gad, 1982) and to fulfill the safety pharmacology associated requirement. It should be noted that validated forms exist for the mouse, rat, dog, and primate.

Blood samples should be taken and analyzed for toxicokinetic analysis after the first and last dose administration in nonrodents (and potentially in rodents if no such *in vivo* data has previously been collected).

Though separate cardiovascular measurements are required and made in the cardiovascular safety pharmacology studies, these are generally performed at lower (more clinically relevant) doses.

Body weights should be measured prior to study start and throughout the course of the study and food consumption at least subjectively evaluated.

While a gross necropsy is performed and selected organ weights collected (and analyzed), it is not recommended (or required) in most cases to proceed to process tissues to slides and have histopathologic evaluations performed. In addition to cost, the 2 months or more of additional time before such data is available is rarely warranted. Rather, a careful evaluation of all the available in life data (everything up to and including organ weights and any gross necropsy observations) in an integrated manner serves to identify potential target organs, set GLP study dose levels, and modify designs for such studies.

REFERENCES

- Andrew DJ. (2009) Acute toxicity. In: *General and Applied Toxicology* (Ballantyne B, Marrs TC and Syversen T, Eds.), 3rd Ed. John Wiley & Sons, Inc., Hoboken, NJ, pp. 57–66.
- Armitage P, Allen J. (1959) Methods of estimating the LD₅₀ in quantal response data. *J. Hygiene*. 48:298–322.
- Arnold DL, Grice HC, Krewski DR. (1990) *Handbook of In Vivo Toxicity Testing*. Academic Press, San Diego, CA.
- Auletta C. (1998) Acute systemic toxicity testing. In: *Handbook of Product Safety Assessment* (Gad SC, Ed.), 2nd Ed. Marcel Dekker, New York.
- Balazs R. (1970) Measurement of acute toxicity. In: *Methods of Toxicology* (Paget G, Ed.). F.A. Davis Co., Philadelphia, PA, pp. 49–81.
- Balazs R. (1976) Assessment of the value of systemic toxicity studies in experimental animals. In: *Advances in Modern Toxicology, Vol. 1 Part 1: New Concepts in Safety Evaluation* (Mahlman M, Shapiro R and Blumenthal H, Eds.). Hemisphere Publishing, Washington, DC, pp. 141–153.
- Ballantyne B. (2009) Repeated exposure toxicology. In: *General and Applied Toxicology* (Ballantyne B, Marre TC and Syversen T, Eds.), 3rd Ed. John Wiley & Sons, Inc., Hoboken, NJ, pp. 73–88.
- Behrens B. (1929) Evaluation of digitalis leaves in frog experiments. *Arch. Exp. Pathol. Pharmacol.* 140:236–256.
- Berkson J. (1955) Estimate of the integrated normal curve by minimum normit chi-square with particular reference to bioassay. *J. Am. Stat. Assoc.* 50:529–549.
- Bliss C. (1935) The calculation of the dosage mortality curve. *Ann. Appl. Biol.* 22:134–167.
- Bliss C. (1957) Some principles of bioassay. *Am. Scientist.* 45:449–466.
- British Toxicology Society. (1984) A new approach to the classification of substances and preparations on the basis of their acute toxicity. A report by the British Toxicology Society working party on toxicology. *Hum. Toxicol.* 3:85–92.
- Brown VK. (1980) *Acute Toxicity*. John Wiley & Sons, New York.
- Brown VK. (1983) Acute toxicity testing. In: *Animals and Alternatives in Toxicity Testing* (Bals M, Riddell R and Worden A, Eds.). Academic Press, New York, pp. 1–13.
- Brownlee K, Hodges J, Rosenblatt M. (1953) The up-and-down method with small samples. *J. Am. Stat. Assoc.* 48:262–277.
- Bruce R. (1985) An up-and-down procedure for acute toxicity testing. *Fundam. Appl. Toxicol.* 5:151–157.
- Chan P, Hayes A. (2015) Acute toxicity and eye irritancy. In: *Principles and Methods of Toxicology* (Hayes A, Ed.), 6th Ed. CRC Press, Boca Raton, FL, pp. 1117–1171.
- Chan P, O'Hara G, Hayes A. (1982) Principles and methods for acute and subchronic toxicity. In: *Principles and Methods of Toxicology* (Hayes A, Ed.). Raven Press, New York, pp. 1–51.
- Dayan A. (1983) Complete program for acute toxicity testing—not only LD₅₀ determination. *Acta Pharmacol. Toxicol.* 52(Suppl. 2):31–51.
- Deichmann W, Gerarde H. (1969) *Toxicology of Drugs and Chemicals*. Academic Press, New York.
- Deichmann W, LeBlanc T. (1943) Determination of the approximate lethal dose with about six animals. *J. Ind. Hyg. Toxicol.* 25:415–417.
- Denny KH, Stewart CW. (2013) Acute, subacute, subchronic and chronic toxicity testing for preclinical drug development. In: *A Comprehensive Guide to Toxicology in Preclinical Drug Development* (Faqi AS, Ed.). Academic Press, San Diego, CA, pp. 87–105.
- DePass L. (1989) Alternative approaches in median lethality (LD₅₀) and acute toxicity testing. *Toxicol. Lett.* 49:159–170.
- DePass L, Myers R, Weaver E, Weil C. (1984) An assessment of the importance of number of dosage levels, number of animals per dosage level, sex, and method of LD₅₀ and slope calculations in acute toxicity studies. In: *Acute Toxicity Testing: Alternative Approaches* (Goldberg A, Ed.), Vol. 2. Mary Ann Liebert, Inc., New York, pp. 139–153.

- Dixon W. (1965) The up-and-down method for small samples. *Am. Statist. Assoc. J.* 60:967–978.
- Dixon WJ, Wood AM. (1948) A method for obtaining and analyzing sensitivity data. *J. Am. Stat. Assoc.* 43:109–126.
- Elsberry D. (1986) Screening approaches for acute and subacute toxicity studies. In: *Safety Evaluation of Drugs and Chemicals* (Lloyd W, Ed.). Hemisphere Publishing, Washington, DC, pp. 145–151.
- Finney DJ. (1971) *Probit Analysis*, 3rd Ed. Cambridge University Press, New York.
- Finney D. (1985) The median lethal dose and its estimation. *Arch. Toxicol.* 56:215–218.
- Fowler J, Rutty D. (1983) Methodological aspects of acute toxicity testing particularly LD₅₀ determinations: present use in development of new drugs. *Acta Pharmacol. Toxicol.* 52(Suppl. 2):20–30.
- Fowler J, Brown J, Bell H. (1979) The rat toxicity screen. *Pharmacol. Therapy.* 5:461–466.
- Gad SC. (1982) A neuromuscular screen for use in industrial toxicology. *J. Toxicol. Environ. Health.* 9:691–704.
- Gad SC. (2007) Statistics for toxicologists. In: *Principles and Methods in Toxicology* (Hayes A, Ed.), 5th Ed. Taylor & Francis, Philadelphia, PA.
- Gad SC, Chengelis CP. (1999) *Acute Toxicology Testing*, 2nd Ed. Academic Press, San Diego, CA.
- Gad SC, Smith A, Cramp A, Gavigan F, Derelanko M. (1984) Innovative designs and practices for acute systemic toxicity studies. *Drug Chem. Toxicol.* 7:423–434.
- Hazelette J, Thompson T, Mertz B, Vuolo-Schuessler L, Gree J, Tripp S, Robertson P, Triana V. (1987) Rising dose tolerance (RDT) study: a novel scenario for obtaining preclinical toxicology/drug metabolism data. *Toxicologist.* 7Abstract #846.
- van den Heuvel M, Clark D, Fielder R, Koundakjian P, Oliver G, Pelling D, Tomlinson N, Walker A. (1990) The international validation of a fixed-dose procedure as an alternative to the classical LD₅₀ test. *Food Chem. Toxicol.* 28:469–482.
- Hopper D. (1986) Behavioral measures in toxicology screening. In: *Safety Evaluation of Drugs and Chemicals* (Lloyd W, Ed.). Hemisphere Publishing, New York, pp. 305–321.
- Interagency Regulatory Liaison Group (IRLG), Office of Consumer Affairs. (1981) *Testing Standards and Guidelines Work Group (HFE-88)*. IRLG, Washington, DC.
- Irwin S. (1962) Drug screening and evaluative procedures. *Science.* 136:123–128.
- Jenner P, Hagan E, Taylor J, Cook E, Fitzhugh O. (1964) Food flavorings and compounds of related structure. I. Acute oral toxicity. *Food Cosmet. Toxicol.* 2:327–343.
- LeBeau J. (1983) The role of the LD₅₀ determination in drug safety evaluation. *Regul. Toxicol. Pharmacol.* 3:71–74.
- Litchfield J, Wilcoxon F. (1949) A simplified method of evaluating dose-effect experiments. *J. Pharmacol. Exp. Ther.* 96:99–113.
- Lörke D. (1983) A new approach to practical acute toxicity testing. *Arch. Toxicol.* 54:275–287.
- Malmfors T, Teiling A. (1983) LD₅₀—its value for the pharmaceutical industry in safety evaluation of drugs. *Acta Pharmacol. Toxicol.* 52(Suppl. 2):229–246.
- McClain R. (1983) Generating, interpreting and reporting information in toxicity studies. *Drug Inf. J.* 17:245–255.
- Morrison J, Quinton R, Reinert H. (1968) The purpose and value of LD₅₀ determinations. In: *Modern Trends in Toxicology* (Boyland E and Goulding R, Eds.), Vol. I. Appleton-Century-Crofts, London, pp. 1–17.
- Muller H, Kley H. (1982) Retrospective study on the reliability of an “approximate LD₅₀” determined with a small number of animals. *Arch. Toxicol.* 51:189–196.
- Office of Science Coordination, FDA. (1984) *Final Report on Acute Studies Workshop Sponsored by the U.S. Food and Drug Administration*. U.S. Government Printshop, Washington, DC.
- Olson H, Fabian R, Greener Y, Pack F, Zelinger D, Dean J. (1990) Reduction in animals used for acute toxicity testing based on retrospective analysis. *Toxicologist* 141 (abstract).
- Organization for Economic Co-operation and Development (OECD). (1991) *OECD Guidelines for Testing of Chemicals, No. 401: Acute Oral Toxicity*. OECD, Paris.
- Osterberg R. (1983) Today’s requirements in food, drug, and chemical control. *Acta Pharmacol. Toxicol.* 52(Suppl. 2): 201–228.
- Pendergast W. (1984) Biological drug regulation. In: *The Seventy-Fifth Anniversary Commemorative Volume of Food and Drug Law*. Edited and Published by the Food and Drug Law Institute, Washington, DC, pp. 293–305.
- Peterson D, Ellis Z, Harris D, Spadek Z. (1987) Acute toxicity of the major alkaloids of cultivated *Lupinus angustifolius* seeds to rats. *J. Appl. Toxicol.* 7:51–53.
- Piegorsch W. (1989) Quantification of toxic response and the development of the median effective dose (ED₅₀)—a historical perspective. *Toxicol. Ind. Health.* 5:55–62.
- Rhodes C. (2000) Principles of testing for acute toxic effects. In: (Ballantyne B (Marrs T and Syversen T, Eds.). *General and Applied Toxicology*. Macmillan References, London, pp. 33–54.
- Rowan A. (1981) The LD₅₀ test: a critique and suggestions for alternatives. *Pharm. Technol.* 5(April):64–94.
- Schiantarelli P, Cadel S. (1981) Piroxicam pharmacological activity and gastrointestinal damage by oral and rectal route: comparison with oral indomethacin and phenylbutazone. *Arzneim.-Forsch./Drug Res.* 31:87–92.
- Schultz E, Fuchs H. (1982) A new approach to minimizing the number of animals in acute toxicity testing and optimizing the information of the test results. *Arch. Toxicol.* 51:197–220.
- Smyth H, Carpenter C. (1944) The place of the range-finding test in the industrial toxicology laboratory. *J. Ind. Hyg. Toxicol.* 26:269–273.
- Sperling F. (1976) Nonlethal parameters as indices of acute toxicity: inadequacies of the acute LD₅₀. In: *Advances in Modern Toxicology, Vol. I, Part 1: New Concepts in Safety Evaluation* (Mehlman M, Shapiro R and Blumenthal H, Eds.). Hemisphere Publishing, Washington, DC, pp. 177–191.
- Thompson W, Weil C. (1952) On the construction of a table for moving average interpolation. *Biometrics.* 8:51–54.
- Trevarn J. (1927) The error of determination of toxicity. *Proc. Roy. Soc. B101*:483–514.

- USFDA. (1996) *Guidance for Industry: Single Dose Acute Toxicity Testing for Pharmaceuticals*. CDER, Rockville, MD.
- USFDA. (2006) *Guidance for Industry Investigations and Reviewers: Exploratory IND Studies*. U.S. Department of Health and Human Services, Food and Drug Administration, Center for Drug Evaluation and Research, Rockville, MD.
- Weil C. (1952) Table for convenient calculation of median effective dose (LD_{50} or ED_{50}) and instructions in their use. *Biometrics*. 8:249–263.
- Weil C. (1983) Economical LD_{50} and slope determination. *Drug Chem. Toxicol.* 6:595–603.
- Wilson DT, Hardisty JF, Hayes RH, Wilson NH. (2015) Short term, subchronic, and chronic toxicity studies. In: *Principles and Methods of Toxicology* (Hayes AW, Ed.), 6th Ed. CRC Press, Boca Raton, FL, pp. 1205–1249.
- Zbinden G. (1984) Acute toxicity testing, purpose. In: *Acute Toxicity Testing: Alternative Approaches* (Goldberg A, Ed.). Mary Ann Liebert, New York, pp. 5–22.
- Zbinden G, Flury-Roversi M. (1981) Significance of the LD_{50} test for the toxicological evaluation of chemical substances. *Arch. Toxicol.* 47:77–99.
- Zbinden G, Elsner J, Boelsterli U. (1984) Toxicological screening. *Regul. Toxicol. Pharmacol.* 4:275–286.

REPEAT-DOSE TOXICITY STUDIES

In the broadest sense, subchronic and chronic studies for pharmaceutical products can incorporate any of the routes used to administer a therapeutic agent, use any of a number of animal models, and conform to a broad range of experimental designs. They can be 2 weeks long (what used to be called “subacute” studies because they were conducted at dose levels below those employed for single-dose or acute studies) or last up to a year. Another name for these studies is repeat-dose studies (Gad, 2008; Ballantyne, 2009; Wilson et al., 2014)—that is, those studies whereby animals have a therapeutic agent administered to them on a regular and repeated basis by one or more routes over a period of 1 year or less. There is great flexibility and variability in the design of such studies. When the primary driver for testing was occupational exposure, test animal exposure was 5 days a week (simulating the work week). Such designs are almost gone, as environmental, consumer, and potential patient exposures (potentially occurring every day) now dictate 7 days a week dosing or exposure. Currently, the duration of such general repeat-dose toxicity studies ranges from 14 days (long enough for the drug levels in the body to reach steady state and for the adaptive immune response to begin to be active) to 9 months (in nonrodents).

This chapter seeks to provide a firm grasp of the objectives for repeat-dose studies, the regulatory requirements governing them, the key factors in their design and conduct, and the interpretation of their results. The key fundamental concepts in understanding mammalian toxicology are presented in Table 8.1.

8.1 OBJECTIVES

As with any scientific study or experiment (but especially for those in safety assessment), the essential first step is to define and understand the reason(s) for the conduct of the

study—that is, its objectives. There are three major (scientific) reasons for conducting subchronic and chronic studies, but a basic characteristic of all but a few subchronic studies needs to be understood. The subchronic study is (as are most other studies in whole animal toxicology) a broad screen. It is not focused on a specific end point; rather, it is a broad exploration of the cumulative biological effects of the administered agent over a range of doses. So broad an exploration, in fact, that it can be called a “shotgun” study.

The objectives of the typical subchronic and chronic studies fall into three categories. The first is to broadly define the toxicity (and, if one is wise, the pharmacology and hyperpharmacology) of repeated doses of a potential therapeutic agent in an animal model (Traina, 1983; Gad, 2014; Wilson et al., 2014). This definition is both qualitative (what are the target organs and the nature of the effects seen) and quantitative (at what dose levels or, more importantly, at what plasma and tissue levels are effects definitely seen and not seen).

The second objective (and the one that in the pharmaceutical industry testing facilities usually compel both timing and compromising of design and execution of studies) is to provide support for the initiation of and/or continued conduct of clinical trials in man (O’Grady and Linet, 1990; Smith, 1992). As such, subchronic studies should provide not only adequate clearance (therapeutic margin) of initial dose levels and duration of dosing but also guidance for any special measures to be made or precautions to be taken in initial clinical trials. Setting inadequate dose levels (either too low or too high) may lead to the failure of a study. A successful study must both define a safe, or “clean,” dose level (one that is as high as possible to allow as much flexibility as possible in the conduct of clinical studies) and demonstrate and/or characterize signs of toxicity at some

higher dose. The duration of dosing issue is driven by a compromise between meeting regulatorily established guidelines (as set out in Table 8.2) and the economic pressure to initiate clinical trials as soon as possible.

The third objective is one of looking forward to later studies. The subchronic study must provide sufficient information to allow a prudent setting of doses for later, longer studies (including, ultimately, carcinogenicity studies). At the same time, the subchronic study must also provide guidance for the other (than dose) design features of longer-term studies (such as what parameters to measure and when to measure them, how many animals to use, and how long to conduct the study).

The FDA and other ICH regulating agencies have specific expectations of these general toxicity studies (as

now codified in ICH M3(R2), 2009 and FDA, 2000). These include:

- a. At least three separate dose groups and concurrent control group.
- b. The high-dose group should if at all possible serve to identify one or more target organs (or organs systems) for toxicity at sublethal doses.
- c. The low-dose group should serve to identify a dose level at which there are no adverse effects (a NOAEL).
- d. The route and regimen employed should mirror those planned or used for clinical evaluation.
- e. Unless the drug is intended for use in just one sex (say for prostate or uterine cancer), equal numbers of males and females should be used (see Table 8.3). The data from the two genders is not pooled, and for all practical purposes, two substudies are thus performed—one for males, one for females.
- f. Additional animals (recovery) should be included in (at least) the high-dose and central groups to allow the evaluation of the reversibility of target organ effects.

The use of at least three (but not necessarily only three) separate treatment groups each receiving a different dose is intended to allow the evaluation of dose–response. Table 8.4

TABLE 8.1 Key Concepts in Understanding Toxicity

- Chronicity
- Absorption
- Distribution
- Reversibility
- Dose–response
- Time response
- Susceptibility
- Special populations
- Statistical versus biological significance
- Metabolism

TABLE 8.2 Recommended Duration of Repeated-Dose Toxicity Studies to Support the Conduct of Clinical Trials

| Maximum Duration of Clinical Trial | Recommended Minimum Duration of Repeated-Dose Toxicity Studies to Support Clinical Trials | |
|------------------------------------|---|--|
| | Rodents | Nonrodents |
| Up to 2 weeks | 2 weeks ^a | 2 weeks ^a |
| Between 2 weeks and 6 months | Same duration as clinical trial ^b | Same duration as clinical trial ^b |
| >6 months | 6 months ^{b,c} | 9 months ^{b,c,d} |

Source: ICH (2009).

^a In the United States, as an alternative to 2-week studies, extended single-dose toxicity studies can support single-dose human trials. Clinical studies of less than 14 days can be supported with toxicity studies of the same duration as the proposed clinical study.

^b In some circumstances clinical trials of longer duration than 3 months can be initiated, provided that the data are available from a 3-month rodent and a 3-month nonrodent study and that complete data from the chronic rodent and nonrodent study are made available, consistent with local clinical trial regulatory procedures, before extending dosing beyond 3 months in the clinical trial. For serious or life-threatening indications or on a case-by-case basis, this extension can be supported by complete chronic rodent data and in-life and necropsy data for the nonrodent study. Complete histopathology data from the nonrodent should be available within an additional 3 months.

^c There can be cases where a pediatric population is the primary population, and existing animal studies (toxicology or pharmacology) have identified potential developmental concerns for target organs. In these cases, long-term toxicity testing starting in juvenile animals can be appropriate in some circumstances.

^d In the EU, studies of 6 months duration in nonrodents are considered acceptable. However, where studies with a longer duration have been conducted, it is not appropriate to conduct an additional study of 6 months. The following are examples where nonrodent studies of up to 6 months duration can also be appropriate for Japan and the United States:

- When immunogenicity or intolerance confounds conduct of longer-term studies.
- Repeated short-term drug exposure even if clinical trial duration exceeds 6 months, such as intermittent treatment of migraine, erectile dysfunction, or herpes simplex.
- Drugs administered on a chronic basis to reduce the risk of recurrence of cancer.
- Drugs for indications for which life expectancy is short.

TABLE 8.3 Numbers of Animals for Chronic and Subchronic Study per Test Group

| Study Length | Mice or Rats per Sex | Dogs or Minipigs per Sex ^a | Primates per Sex ^a |
|-----------------------|-------------------------|--|----------------------------------|
| 2–4 weeks | 5–10 | 3–4 (+2/3) | 3 (+2/3) |
| 3 months | 20 | 6 (+3) | 5 (+3) |
| 6 months | 30 | 8 (+3) | 5 (+3) |
| 9 months or 1 year | N/A | 10 | 10 |

^a Numbers in parenthesis are for either animals added to high dose and control to allow assessment of reversibility.

TABLE 8.4 The Three Dimensions of Dose–Response

| As dose increases... |
|--|
| <ul style="list-style-type: none"> • Incidence of responders in an exposed population increases • Severity of response in effected individuals increases • Time to occurrence of response or of progressive stage of response decreases • An implication is that as duration of dosing increases, NOEL/NOAEL decreases |

presents some important aspects of such dose–response relationships.

These objectives are addressed by the usual subchronic study. Some subchronic studies, however, are unusual in being conceived, designed, and executed to address specific questions raised (or left unanswered) by previous preclinical or early clinical studies. Such a special purpose is addressed separately.

Chronic studies (those that last 6 or 9 months or a year) may also be conducted for the aforementioned purposes but are primarily done to fulfill registration requirements for drugs that are intended for continuous (or frequent intermittent) long-term (lifetime) use.

8.2 REGULATORY CONSIDERATIONS

Much of what is done (and how it is done) in repeat-dose studies is a response to a number of regulations. Three of these have very broad impact. These are the good laboratory practices (GLP) requirements, Animal Welfare Act requirements, and regulatory requirements that actually govern study design.

8.2.1 Good Laboratory Practices

Since 1978, the design and conduct of preclinical safety assessment studies for pharmaceuticals in the United States (and, indeed, internationally) have been governed and significantly influenced by development of GLPs. Strictly

speaking, these regulations cover qualifications of staff and facilities, training, record keeping, documentation, and actions required to ensure compliance with and the effectiveness of these steps. Though the initial regulations were from the US FDA (2014), they have always expanded in content and extended to cover studies performed overseas. Now, most other countries have adopted similar regulations. A discussion of these regulations is beyond the scope of the current chapter, but several aspects are central to this effort. Each technique or methodology to be employed in a study (such as animal identification, weighing and examination, blood collection, data recording, etc.) must be adequately described in a standard operating procedure (SOP) before the study begins. Those who are to perform such procedures must be trained in them beforehand. The actual design of the study, including start date and how it is to be ended and analyzed, plus the principal scientists involved (particularly the study director), must be specified in a protocol that is signed before the study commences. Any changes to these features must be documented in amendments once the study has begun. It is a good idea that the pathologist who is to later perform or oversee histopathology be designated before the start of the study and that the design be a team effort involving the best efforts of the toxicologist, pathologist, and (usually, for subchronic studies) drug metabolism scientist.

8.2.2 Animal Welfare Act

Gone are the days when the pharmaceutical scientist could conduct whatever procedures or studies that were desired using experimental animals. The Animal Welfare Act (APHIS, 1989, amended every 5 years since, most recently pending in 2015) (and its analogs in other countries) rightfully requires careful consideration of animal usage to ensure that research and testing uses as few animals as possible in as humane a manner as possible. As a start, all protocols must be reviewed by an Institutional Animal Care and Use Committee. Such review takes time but should not serve to hinder good science. When designing a study or developing a new procedure or technique, the following points should be kept in mind:

1. Will the number of animals used be sufficient to provide the required data yet not constitute excessive use? (It ultimately does not reduce animal use to utilize too few animals to begin with and then have to repeat the study.)
2. Are the procedures employed the least invasive and traumatic available? This practice is not only required by regulations but is also sound scientific practice, since any induced stress will produce a range of responses in test animals that can mask or confound the chemically induced effects.

8.2.3 Regulatory Requirements for Study Design

The first consideration in the construction of a study is a clear statement of its objectives, which are almost always headed by meeting regulatory requirements to support drug development and registration. Accordingly, the relevant regulatory requirements must be analyzed, which is complicated by the fact that new drugs are no longer developed for registration and sale in a single-market country. The expense is too great, and the potential for broad international sales is too appealing. While each major country has its own requirements as to study designs and studies required (with most of the smaller countries adhering to the regulations of one of the major players), harmonization has done much to smooth these differences (Gad, 2010). Meeting these regulatory requirements is particularly challenging for several reasons. First, the only official delineation of general requirements in the United States is dated (FDA, 1971), and recently special cases have arisen (anti-HIV agents, biotechnologically derived agents, therapeutic agents for neonates and the very elderly, etc.) that try the utility of these requirements. These needs have led to a stream of points to consider which seek to update requirements. Second, the term “guidelines” means different things in different countries (in the United States it means “requirements” and in Japan, “suggestions”).

Agents intended to treat or arrest the progress of rapidly spreading life-threatening diseases (such as AIDS) are subject to less stringent safety assessment requirements prior to initial clinical evaluations than are other drugs. However, even though approval (if clinical efficacy is established) for marketing can be granted with preclinical testing still under way, all applicable safety assessments (as with any other class of drugs) must still be completed (FDA, 1988).

Drugs intended for use in either the elderly or the very young have special additional requirements for safety evaluation, in recognition of the special characteristics and potential sensitivities of these populations. For the elderly, these requirements call for special consideration of renal and hepatic effects (Center for Drug Evaluation and Research (CDER), 1989). Likewise, drugs intended for the young require special studies to be performed in neonates and juvenile animals (usually of 2 or 4 weeks' duration in rats).

In the last 5–6 years, a number of potentially important drugs have been produced by recombinant DNA technology. These biomacromolecules, which are primarily endogenously occurring proteins, present a variety of special considerations and concerns, including the following:

- Because they are endogenously occurring molecules, assessing their pharmacokinetics and metabolism presents special problems.
- Is the externally commercially produced molecule biologically equivalent to the naturally occurring one?
- As proteins, are they immunogenic or do they provoke neutralizing antibodies that will limit their usefulness?

- Because they are available only in very small quantities, the use of traditional protocols (such as those that use ever-increasing doses until an adverse effect is achieved) is impractical.
- Agents with such specific activity in man may not be appropriately evaluated in rodents or other model species.

Each of these points must be addressed in any safety testing plan (Weissinger, 1989). The requirements set out in this chapter are designed to do this (for repeat-dose testing).

8.3 STUDY DESIGN AND CONDUCT

8.3.1 Animals

In all but a few rare cases, for pharmaceutical safety assessment, separate studies in at least two species are required. Regulations require that both species be mammalian, and one of these must be a nonrodent; practice and economics dictate that the other species will be a rodent. With extremely rare exception, the rodent species employed is the rat (though the mouse also sees significant use). There is considerably more variability in the nonrodent species, with a range of factors determining whether the dog (most common choice), a primate species (typically the rhesus or cynomolgus, though some others are used in particular cases), the pig (particularly in Europe), or some other animal (e.g., the ferret) is selected. The factors that should and do govern species selection are presented in detail in Gad (2015). The use of multiple species is a regulatory requirement arising from experience and the belief (going back to 1944, at least) that it will provide a better chance of detecting the full range of biological responses (adverse and otherwise) to the new molecular entity being evaluated. This belief has come under fire in recent years (Zbinden, 1993), and not followed for larger molecular entities after initial systemic toxicity studies, but is unlikely to be changed soon. Along the same lines, unless an agent is to be used by only one sex or the other of humans, equal numbers of both sexes of an animal species are utilized in the studies, with the sexes being treated as unrelated for purposes of statistical analysis. Also except in rare cases, the animals used are young, healthy adults in the logarithmic phase of their growth curve. (The FDA specifies that rodents be <6 weeks of age at the initiation of dosing.)

Numbers of animals to be used in each dose group of a study are presented in Table 8.3. Though the usual practice is to use three different dose groups and at least one equal-sized control group, this number is not fixed and should be viewed as a minimum (see Section 8.3.4). Use of more groups allows for a reduction in the risk of not clearly defining effects and establishing the highest possible safe

dose at a modest increase in cost. There must be as many control animals as are in the largest-size test group to optimize statistical power.

Animals are assigned to groups (test and control) by one or another form of statistical randomization. Prior to assignment, animals are evaluated for some period of time after being received in house (usually at least 1 week for rodents and 2 for nonrodents) to ensure that they are healthy and have no discernible abnormalities. The randomization is never pure; it is always “blocked” in some form or another (by initial body weight, at least) so that each group is not (statistically) significantly different from the others in terms of the “blocked” parameters.

Proper facilities and care for test animals are not only a matter of regulatory compliance (and a legal requirement) but also essential for a scientifically sound and valid study.

Husbandry requires clean cages of sufficient size and continuous availability of clean water and food (unless the protocol requires some restriction on their availability). Environmental conditions (temperature, humidity, and light–dark cycle) must be kept within specified limits. All of these must, in turn, be detailed in the protocols of studies. The limits for these conditions are set forth in relevant NIH and USDA publications.

8.3.2 Routes and Setting Doses

Route (how an agent is administered to a test animal) and dose (how much of and how frequently an agent is administered) are inseparable in safety assessment studies and really cannot be defined independently. The selection of both begins with an understanding of the intended use of the drug in humans. The ideal case is to have the test material administered by the same route, at the same frequency (once a day, three times a day, etc.), and for the same intervals (e.g., continuously, if the drug is an intravenously infused agent) as the drug’s eventual use in people. Practical considerations such as the limitations of animal models (i.e., there are some things you can’t get a rat to do), limitations on technical support¹ and the like, and regulatory requirements (discussed later as part of dose setting) frequently act or interact to preclude this straightforward approach.

Almost 30 routes exist for administration of drugs to patients (see Chapter 5), but only a handful of these are commonly used in preclinical safety studies (Gad, 1994). The most common deviation from what is to be done in clinical trials is the use of parenteral (injected) routes such as intravenous (IV) and subcutaneous (SC) deliveries. Such injections are loosely characterized as bolus (all at once or over a

very short period, such as 5 min) and infusion (over a protracted period of hours, days, or even months). The term *continuous infusion* implies a steady rate over a protracted period, requiring some form of setup such as an implanted venous catheter or infusion port.

It is rare that the raw drug itself is suitable (in terms of stability, local tissue tolerance, and optimum systemic absorption and distribution) for direct use as a dosage form. Either it must be taken into a solution or suspension in a suitable carrier, a more complex formulation (a prototype of the commercial form) must be developed. Gad (2015) should be consulted for a more complete discussion of dose formulation for animals or humans. One formulation or more must be developed (preferably the same one for both animals and humans) based on the specific requirements of preclinical dosage formulation. For many therapeutic agents, limitations on volumes that can be administered and concentrations of active ingredient that can be achieved impact heavily on dose setting.

Setting of doses for initial repeat-dose studies, when little is known as to dose–response and pharmacokinetics, is one of the most difficult tasks in study design. The doses administered must include one that is devoid of any adverse effect (preferably of *any* effect) and yet still high enough to “clear” the projected clinical dose by the traditional (10× for rodents, 5× for nonrodents) or regulatory safety factors (the HED factor times 10 for most therapeutic areas). At the same time, if feasible, at least one of the doses should characterize the toxicity profile associated with the agent (for some biotechnologically derived agents, particularly those derived from endogenous human molecules, it may only be possible to demonstrate biological effects in appropriate disease models and impossible to demonstrate toxicity). Because of limitations on availability of protodrugs, it is generally undesirable to go too high to achieve this second (toxicity) objective.

Traditionally, studies include three or more dose groups to fulfill these two objectives. Based on earlier results (generally, single-dose or 2-week studies), doses are selected. It is, by the way, generally an excellent idea to observe the “decade rule” in extrapolation of results from shorter to longer studies; that is, do not try to project doses for more than an order-of-magnitude-longer study (thus the traditional progression from single-dose to 14-day to 90-day studies). Also, one should not allow the traditional use of three dose groups plus a control to limit designs. If there is a great deal of uncertainty, it is much cheaper in every way to use four or five dose groups in a single study than to have to repeat the entire study. Finally, remember that different doses may be appropriate for the different sexes.

It should also be kept in mind that formulating materials may have effects of their own, and “vehicle” control groups may be required in addition to a negative control group. Additionally, the limitations on volumes of dose administration should be kept in mind as presented in Table 8.5.

¹ Many antiviral agents, particularly some anti-HIV agents, have rather short plasma half-lives, which require frequent oral administration of the agent. Thirteen-week studies have been conducted with TID dosing of rats and monkeys, requiring around-the-clock shift work for technical staff of the laboratory.

TABLE 8.5 Guidance on Volumes of Administration

| Species | Routes and Volumes (mL kg ⁻¹ Administration Site) | | | | | |
|----------|--|---------|---------|-------------------------------------|----------|--------------|
| | Oral | SC | IP | IM | IV Bolus | IV (Slowing) |
| Mouse | 10 (50) | 10 (40) | 20 (80) | 0.5 ^a (0.1) ^a | 5 | (25) |
| Rat | 10 (40) | 5 (10) | 10 (20) | 0.1 ^a (0.2) ^a | 5 | (20) |
| Rabbit | 10 (15) | 1 (2) | 5 (20) | 0.25 (0.5) | 2 | (10) |
| Dog | 5 (15) | 1 (2) | 1 (20) | 0.25 (0.5) | 2.5 | (10) |
| Macaque | 5 (15) | 2 (5) | —(10) | 0.25 (0.05) | 2 | (—) |
| Marmoset | 10 (15) | 2 (5) | —(20) | 0.25 (0.5) | 2.5 | (10) |
| Minipig | 10 (15) | 1 (2) | 1 (20) | 0.25 (0.05) | 2.5 | (5) |

Source: Adapted from ECVAM (2000).

(—) data not available. For nonaqueous injectates consideration must be given to time of absorption before redosing. No more than two IM sites should be used per day. SC sites should be limited to two to three sites per day.

IM, intramuscular; IP, intraperitoneal; SC, subcutaneous.

^a Milliliters.

8.3.3 Parameters to Measure

As was stated earlier, repeat-dose general (systemic) toxicity studies are “shotgun” in nature; that is, they are designed to look at a very broad range of end points with the intention of screening as broadly as indications of toxicity. Meaningful findings are rarely limited to a single end point; rather, what typically emerges is a pattern of findings. This broad search for components of toxicity profile is not just a response to regulatory guidelines intended to identify potentially unsafe drugs. An understanding of all the indicators of biological effect can also frequently help one to understand the relevance of findings, to establish some as unrepresentative of a risk to humans, and even to identify new therapeutic uses of an agent.

Parameters of interest in the repeat-dose study can be considered as sets of measures, each with its own history, rationale, and requirements. Chapter 6 sought to present an overview of such parameters. It is critical to remember, however, that the strength of the study design as a scientific evaluation lies in the relationships and patterns of effects that are seen not in simply looking at each of these measures (or groups) as independent findings but rather as integrated profiles of biological effects.

8.3.3.1 Pharmacokinetics and Metabolism All regulatory repeat-dose general toxicity studies now incorporate (either in the study itself or in a parallel study) evaluation of the basic pharmacokinetics of a compound. This is discussed in detail in Chapter 17.

Pharmaceutical subchronic toxicity studies are always accompanied by a parallel determination of the pharmacokinetics of the material of interest administered by the same route as that used in the safety study. This parallel determination consists of measuring plasma levels of the administered agent and its major metabolites either in animals that are part of the main study or in a separate set of animals (in parallel with the main study) that are dosed and evaluated to

determine just these end points. The purpose of these determinations is both to allow a better interpretation of the findings of the study and to encourage the most accurate possible extrapolation to humans. The first data of interest are the absorption, distribution, and elimination of the test material, but a number of other types of information can also be collected (Yacobi et al., 1989; Tse and Jaffe, 1991). For nonparenteral routes it is essential to demonstrate that systemic absorption and distribution of the test material did occur; otherwise, it is open to question whether the potential safety of the agent in man has been adequately addressed (not to mention the implication for potential human therapeutic efficacy). A complication, however, is that there are limits as to how much blood may be collected from specific species at one time (see Table 8.6) particularly as samples must be drawn to allow evaluation of clinical chemistry, clinical pathology, and pharmacokinetics. This is even worse in the case of biologics, where adequate samples must also be drawn to allow the evaluation of antibody levels. However, low volume sampling approaches are becoming more popular (Chapman et al., 2014).

8.3.4 Study Designs

The traditional design for a repeat-dose toxicity study is very straightforward. The appropriate numbers of animals of each sex are assigned to each of the designated dose and control groups. Unfortunately, this basic design is taken by many to be dogma, even when it does not suit the purposes of the investigator. There are many possible variations to study design, but four basic factors should be considered: controls, the use of interval and satellite groups, balanced and unbalanced designs, and staggered starts.

Classically, a single control group of the same size as each of the dose groups is incorporated into each study. Some studies incorporate two control groups (each with the same size as the experimental groups) to guard against

TABLE 8.6 Total Blood Volumes and Recommended Maximum Blood Sample Volumes for Species of Given Bodyweight

| Species | Blood Volume (mL) | 7.5% (mL) | 10% (mL) | 15% (mL) | 20% (mL) |
|-----------------------------|-------------------|-----------|----------|----------|----------|
| Mouse (25 g) | 1.8 | 0.1 | 0.2 | 0.3 | 0.4 |
| Rat (250 g) | 16 | 1.2 | 1.6 | 2.4 | 3.2 |
| Rabbit (4 kg) | 224 | 17 | 22 | 34 | 45 |
| Dog (10 kg) | 850 | 64 | 85 | 127 | 170 |
| Macaque (rhesus) (5 kg) | 280 | 21 | 28 | 42 | 56 |
| Macaque (cynomolgus) (5 kg) | 325 | 24 | 32 | 49 | 65 |
| Marmoset (350 kg) | 25 | 2.0 | 2.5 | 3.5 | 5 |
| Minipig (15 kg) | 975 | 73 | 98 | 146 | 195 |

having a statistically significant effect due to one control group being abnormal for one or more parameters (a much more likely event when laboratory animals were less genetically homogeneous than they are now). The belief is that a “significant” finding that differs from one (but not both) of the concurrent control groups, and does not differ from historical control data, can be considered as not biologically significant. This is, however, an indefensible approach. Historical controls have value, but it is the concurrent control group(s) in a study that is of concern.

Interval or satellite groups have been discussed at two earlier points in this chapter. They allow measurement of termination parameters at intervals other than at termination of the study. They are also useful when the manipulation involved in making a measurement (such as the collection of an extensive blood sample), while not terminal, may compromise (relative to other animals) the subject animals. Another common use of such groups is to evaluate recovery from some observed effect at study termination.

Usually, each of the groups in a study is the same size, with each of the sexes being equally represented. The result is called a balanced design, with statistical power for detection of effects optimized for each of the treatment groups. If one knows little about the dose toxicity profile, this is an entirely sound and rational approach. However, there are situations when one may wish to utilize an unbalanced design—that is, to have one or more dose groups larger than the others. This is usually the case when either greater sensitivity is desired (typically in a low-dose group), or an unusual degree of attrition of test animals is expected (usually due to mortality in a high-dose group), or as a guard against a single animal’s idiopathic response being sufficient to cause “statistical significance.”

As it is the normal practice to have a balanced design, it is also traditional to initiate treatment of all animals at the same time. This may lead to problems at study termination, however. It is a very uncommon toxicology laboratory that can “bring a study down” on a single day. In fact, there are no labs that can collect blood and perform necropsies in a single day on even the 48–80 dogs involved in a study, much less the 160–400+ rats in the rodent version. Starting all animals on study the same day presents a number of less than

desirable options. The first is to terminate as many animals as can be done each day, continuing to dose (and therefore, further affect) the remaining test animals. Assuming that the animals are being terminated in a random, balanced manner, this means that the last animals terminated will have received from 3 to 10 additional days of treatment. At the least, this is likely to cause some variance inflation (and therefore both decrease the power of the study design and possibly confound interpretation). If the difference in the length of treatment of test animals is greater than 3% of the intended length of the study, one should consider alternative designs.

An alternative approach to study design that addresses this problem employs one of several forms of staggered starts. In these, distinct groups of animals have their dosing initiated at different times. The most meaningful form recognizes that the two sexes are in effect separate studies anyway (they are never compared statistically, with the treatment groups being compared only against the same-sex control group). Thus if the termination procedure for one sex takes 3–5 days, then one sex should be initiated on dosing 1 week and the other on the following week. This maximizes the benefits of common logistical support (such as dose formulation) and reduces the impact of differential length of dosing on study outcome.

A variation on this is to either stagger the start of different dose groups or of the satellite and main study portions of dose groups. The former is to be avoided (it will completely confound study outcome), while the latter makes sense in some cases (pharmacokinetics and special measures) but not others (recovery and interval sacrifice).

8.4 STUDY INTERPRETATION AND REPORTING

For a successful repeat-dose study, the bottom line is the clear demonstration of a no-effect level, characterization of a toxicity profile (providing guidance for any clinical studies), enough information on pharmacokinetics and metabolism to scale dosages to human applications, and at least a basic understanding of the mechanisms involved in any identified pathogenesis. The report that is produced as a result of the study should clearly communicate these points—along with

TABLE 8.7 Summary Integrative Assessment of Study Results

| Parameter | Dose Group | | | | | |
|--------------------|------------|--------|--------|--------|------|--------|
| | Low | | Middle | | High | |
| | Male | Female | Male | Female | Male | Female |
| Morbidity | | | | | | |
| Body weight | | | | | | |
| Clinical signs | | | | | | |
| Organ weights | | | | | | |
| Clinical chemistry | | | | | | |
| Hematology | | | | | | |
| Histopathology | | | | | | |

TABLE 8.8 Troubleshooting in General Toxicology

| | |
|---|---|
| Unexpected toxicity, compared with prior tests | Change in formulation or batch of test chemical. Poor predictability of dose range-finder studies due to factors such as differences in animal age, supplier, or husbandry |
| Variation in individual response | Metabolic polymorphism or other genetic factor and social factors in group housing, for example, nutrition status |
| Low systemic concentration or area under the curve (AUC) | Poor absorption or poor formulation; isotonicity is important in parenteral formulations. Extensive first-pass effect. Short half-life |
| Low toxicity | Low availability; inappropriate route of administration or dose selection |
| Interspecies differences | Different ADME; different mechanism of effect; species-specific mechanisms such as peroxisome proliferation; enterohepatic recirculation. Different expression of or affinity for pharmacological receptors |
| Different response in males and females | Especially in rodents; due to different activities of metabolism enzymes in liver particularly but also physiological differences such as α_2 -microglobulin excretion in males |

the study design and experimental procedures, summarized data, and their statistical analysis—and it should be GLP compliant and suitable for FDA submission format.

Interpretation of the results of a study should be truly scientific and integrative. It is elementary to have the report state only each statistically and biologically significant finding in an orderly manner and not just a recitation of all observations. The meaning and significance of each in relation to other findings, as well as the relevance to potential human effects, must be evaluated and addressed.

The author of the report should ensure that it is accurate and complete but also that it clearly tells a story and concludes with the relevant (to clinical development) findings. A useful approach is to construct a summary table (such as illustrated in

Table 8.7) which gives an overview by dose group, gender, and grouping of observation. The initial use of such a table should be as soon as the “in-life” data (all but the histopathology) from main groups in studies is available, as it can serve as both a tool for early understanding of findings and a guide to what examination may be added or modified in recovery group animals.

There are some common problems encountered in general toxicity studies. The most common of these and their usual causality are presented in Table 8.8.

The usual case is that over the course of drug development, we go from shorter (14 or 28 days) studies in progressive steps to longer studies (90 days/13 weeks) then chronic studies of 6 or 7 months. As we progress through this sequence, the results of earlier studies should modify the design of longer studies.

REFERENCES

- Animal and Plant Health Inspection Service (APHIS) (1989) 54 FR 36112, August 31, 1989, as amended at 78 FR 57227, September 18, 2013. United States Department of Agriculture (USDA), Animal Welfare Act, Washington, DC.
- Ballantyne B. (2009) Repeated exposure toxicity. In: *General and Applied Toxicology* (Ballantyne B, Marrs T and Syversen T, Eds.), 3rd Ed. John Wiley & Sons, Inc., Hoboken, NJ, pp. 73–88.
- CDER (Center for Drug Evaluation and Research). (November 1989) *Guideline for the Study of Drugs Likely to Be Used in the Elderly*. FDA, Washington, DC.
- Chapman K, Chivers S, Gliddon D, Mitchell D, Robinson S, Sangster T, Sparrow S, Spooner N, Wilson A. (2014) Overcoming the barriers to the uptake of nonclinical microsampling in regulatory safety studies. *Drug Discov. Today*. 19(5):528–532.
- ECVAM. (2000) A good practice guide to the administration of substances and removed of blood, including routes and volumes. *J. Appl. Toxicol.* 21:15–23.
- FDA. (1971) *FDA Introduction to Total Drug Quality*. US Government Printing Office, Washington, DC.
- FDA. (1988) Investigational new drug, antibiotic, and biological drug product regulations, procedures for drugs intended to treat life-threatening and severely debilitating illness. *Fed. Regist.* 53(204):41516–41524.

- FDA. (2000) *Toxicological Principles for the Safety of Food Ingredients, Redbook 2000*. Center for Food Safety and Applied Nutrition, FDA, Washington, DC.
- FDA. (2014) *Good Laboratory Practices for Nonclinical Laboratory Studies*. CFR 21 Part 58, April 2014. Department of Health and Human Services, Washington, DC.
- Gad SC. (1994) Routes in safety assessment studies. *J. Am. Coll. Toxicol.* 13:1–17.
- Gad SC. (2008) Repeat dose toxicity studies. In: *Preclinical Development Handbook: Toxicology* (Gad SC, Ed.). Wiley-Interscience, Hoboken, NJ.
- Gad SC. (2010) *International Regulatory Safety Evaluation of Pharmaceuticals and Medical Devices*. Springer, Berlin.
- Gad SC. (2014) Safety and regulatory requirements and challenge for CNS drug development. *Neurobiol. Dis.* 61(2):39–46.
- Gad SC. (2015) *Animal Models in Toxicology*, 3rd Ed. CRC Press, Boca Raton, FL.
- ICH. (2009) *M3(R2) nonclinical safety studies for the conduct of human clinical trials for pharmaceuticals*. International Conference on Harmonisation, Geneva, Switzerland.
- O'Grady J, Linet OI. (1990) *Early Phase Drug Evaluation in Man*. CRC Press, Boca Raton, FL.
- Smith CG. (1992) *The Process of New Drug Discovery and Development*. CRC Press, Boca Raton, FL.
- Traina VM. (1983) The role of toxicology in drug research and development. *Med. Res. Rev.* 3:45–72.
- Tse FLS, Jaffe JM. (1991) *Preclinical Drug Disposition*. Marcel Dekker, New York.
- Weissinger J. (1989) Nonclinical pharmacologic and toxicologic considerations for evaluating biologic products. *Regul. Toxicol. Pharmacol.* 10:255–263.
- Wilson DT, Hardisty JF, Hayes JR, Wilson NH. (2014) Short-term, subchronic and chronic toxicity studies. In: *Principles and Methods of Toxicology* (Hayes AW and Kruger CL, Eds.), 6th Ed. CRC Press, Boca Raton, FL, pp. 1205–1250.
- Yacobi A, Skelly JP, Batra VK. (1989) *Toxicokinetics and New Drug Development*. Pergamon Press, New York.
- Zbinden G. (1993) The concept of multispecies testing in industrial toxicology. *Regul. Toxicol. Pharmacol.* 17:85–94.

GENOTOXICITY

Genotoxicity encompasses all the potential means by which the genetic material of higher organisms may be damaged, with resulting serious consequences to both the organism they occur in and potentially their offspring. With our increasing knowledge of genomics, it needs to be clear that somatic cell genomic changes which are not passed to subsequent generations of cells are not considered genotoxic events. Most forms of genotoxicity are expressions of mutagenicity—the induction of deoxyribonucleic acid (DNA) damage and other genetic alterations, with changes in one or a few of DNA base pairs (gene mutations). Others are clastogenicity, with gross changes in chromosomal structure (i.e., chromosomal aberrations (CAs)) or in chromosome numbers. Clearly the potential of any pharmaceutical to cause such damage is a concern, though more than half of all currently marketed drugs have shown as genotoxic in one or more test systems (Snyder and Green, 2001).

It has been known for several hundred years that exposure to particular chemicals or complex mixtures can lead to cancer in later life (Doll, 1977), and it has been postulated more recently that chemicals can also induce heritable changes in man, leading to diseases in the next generation (ICEMC, 1983). There has been accumulating evidence that such changes can arise following damage to DNA and resulting mutations (see, e.g., Slater et al., 1971; Bridges, 1976). Therefore, it has become necessary to determine whether widely used drugs or potentially useful new drugs possess the ability to damage DNA (Ishidate, 1988a, b). In pharmaceutical development, such information may be used to discard a new candidate drug from further work, to control or eliminate human exposure for a mutagenic industrial compound, or, for a drug, to proceed with development if benefits clearly outweigh risks. Data concerning the genotoxicity of a new drug have become part of the safety

package, though the timing of the performance of the tests may vary. They are needed for decision making and to reduce risks that might otherwise be unforeseen.

The newest revision of the International Conference on Harmonization (ICH) guidance on this subject, ICH S2(R1), sets forth clear guidance in genotoxicity testing requirements in the form of two options (as summarized in Table 9.1). This newest revision, which also serves to merge the former S2A and S2B guidelines, is just (2011) into consideration. This revision fundamentally changed many aspects of how candidate drugs are evaluated for potential genotoxicity. The FDA follows the operative ICH guidances (ICH, 2011, 2012) and ISO (2012, 2014). The Japanese MAFF guidance was last updated in 2000 (MAFF, 2000).

9.1 ICH TEST PROFILE


The ICH recommends a rather different profile of genotoxicity tests for drugs than is used for environmental chemicals. They want to see at least a single relevant *in vivo* test conducted, with subsequent testing performed to allow evaluation of the relevance of test results to potential patient risk (Gad et al., 2015). While actual study design guidelines are described in the OECD test guidelines (OECD, 1997a, b, c, 2010, 2014a, b), the ICH S9 guideline details which studies should be performed.


9.2 DNA STRUCTURE

With the exception of certain viruses, the blueprint for all organisms is contained in a code saved in the form of DNA, a giant macromolecule whose structure allows a vast amount

TABLE 9.1 Genotoxicity Tests Recommended by ICH

| ICH S2A/B | ICH S2(R1) | |
|---|--|---|
| Standard battery | Option 1 | Option 2 |
| Ames microbial mutagenesis and repeat | Ames (one complete assay) | Ames (one complete assay) |
| <i>In vitro</i> mammalian cell assay: | <i>In vitro</i> mammalian cell assay: (1 nM top concentration) | No <i>in vitro</i> mammalian cell assay |
| Chromosome aberrations | Chromosome aberrations | |
| Or | Or | |
| TK mutations in mouse lymphoma cells | TK mutations in mouse lymphoma cells | |
| | Or | |
| | Micronucleus assay | |
| <i>In vivo</i> micronucleus (mouse or rat bone marrow, mouse blood) or chromosomal aberration assay (mouse or rat marrow) | <i>In vivo</i> micronucleus or chromosome aberration assay (mouse/rat bone marrow or blood) ^a Before phase I if integrated Before phase II if acute study | <i>In vivo</i> Gene-Tox with two tissues ^a : • Micronuclei • 2nd <i>in vivo</i> end point (comet, DNA breaks, adducts, mutation) |

 Before phase I.

 Before phase II.

^aCan be integrated into repeat dose studies. Micronuclei analysis can be performed by flow cytometry.

of information to be stored accurately. We have all arisen from a single cell, the fertilized ovum containing two sets of DNA (packaged with protein to form chromatin), one set from our mother, resident in the nucleus of the unfertilized ovum, and the second set from our father via the successful sperm. Every cell in the adult has arisen from this one cell and (with the exception of the germ cell and specialized liver cells) contains one copy of these original chromosome sets.

The genetic code is composed of four “letters”—two pyrimidine nitrogenous bases, thymine and cytosine, and two purine bases, guanine and adenine—which can be regarded functionally as arranged in codons (or triplets). Each codon consists of a combination of three letters; therefore, 4³ (64) different codons are possible. Sixty-one codons code for specific amino acids (three produce stop signals), and as only 20 different amino acids are used to make proteins, one amino acid can be specified by more than one codon.

The bases on one strand are connected together by a sugar (deoxyribose) phosphate backbone. DNA can exist in a single-stranded or double-stranded form. In the latter state, the two strands are held together by hydrogen bonds between the bases. Hydrogen bonds are weak electrostatic forces involving oxygen and nitrogen atoms. As a strict rule, one fundamental to mutagenesis, the adenine bases on one strand always hydrogen bond to the thymine bases on the sister strand. Similarly, guanine bases pair with cytosine bases. Adenine and thymine form two hydrogen bonds, and guanine and cytosine form three.

Double-stranded DNA has a unique property in that it is able to make identical copies of itself when supplied with precursors, relevant enzymes, and cofactors. In simplified terms, two strands begin to unwind and separate as the hydrogen bonds are broken. This produces single-stranded regions. Complementary deoxyribonucleotide triphosphates then pair with the exposed bases under the control of a DNA polymerase enzyme.

A structural gene is a linear sequence of codons which codes for a functional polypeptide, that is, a linear sequence of amino acids. Individual polypeptides may have a structural, enzymatic, or regulatory role in the cell. Although the primary structure of DNA is the same in prokaryotes and eukaryotes, there are differences between the genes of these two types of organism, in internal structure, numbers, and mechanism of replication. In bacteria, there is a single chromosome, normally a closed circle, which is not complexed with protein, and replication does not require specialized cellular structures. In plant and animal cells, there are many chromosomes, each present as two copies, as mentioned earlier, and the DNA is complexed with protein. Replication and cell division require the proteinaceous spindle apparatus. The DNA of eukaryotic cells contains repeated sequences of some genes. Also, eukaryotic genes, unlike prokaryotic genes, have noncoding DNA regions called introns between coding regions called exons. This property means that eukaryotic cells have to use an additional processing step at transcription.

9.2.1 Transcription

The relationship between the DNA in the nucleus and proteins in the cytoplasm is not direct. The information in the DNA molecule is transmitted to the protein-synthesizing machinery of the cell via another informational nucleic acid, called messenger RNA (mRNA), which is synthesized by an enzyme called RNA polymerase. Although similar to DNA, mRNAs are single stranded and possess the base uracil instead of thymine and the sugar ribose rather than deoxyribose. These molecules act as short-lived copies of the genes being expressed.

In eukaryotic cells, the initial mRNA copy contains homologs of both the intron and exon regions. The intron regions are then removed by enzymes located in the nucleus of the cell. Further enzymes splice the exon regions together to form the active mRNA molecules. In both groups of organisms, mature mRNA molecules then pass out of the nucleus into the cytoplasm.

9.2.2 Translation

The next process is similar in both eukaryotes and prokaryotes and involves the translation of mRNA molecules into polypeptides. This procedure involves many enzymes and two further types of RNA: transfer RNA (tRNA) and ribosomal RNA (rRNA). There is a specific tRNA for each of the amino acids. These molecules are involved in the transportation and coupling of amino acids in to the resulting polypeptide. Each tRNA molecule has two binding sites, one for the specific amino acid and the other containing a triplet of bases (the “anticodon”) which is complementary to the appropriate codon on the mRNA.

rRNA is complexed with protein to form a subcellular globular organelle called a ribosome. Ribosomes can be regarded as the “reading head” which allows the linear array of mRNA codons each to base-pair with an anticodon of an appropriate incoming tRNA/amino acid complex. The polypeptide chain forms as each tRNA/amino acid comes into register with the RNA codon and with specific sites on the ribosome. A peptide bond is formed between each amino acid as it passes through the reading head of the ribosome (Venitt and Parry, 1984).

9.2.3 Gene Regulation

Structural genes are regulated by a special set of codons, in particular “promoter” sequences. The promoter sequence is the initial binding site for RNA polymerase before transcription begins. Different promoter sequences have different affinities for RNA polymerase. Some sets of structural genes with linked functions have a single promoter, and their coordinate expression is controlled by another regulatory gene called an operator. A group of such genes is called an operon.

The activity of the operator is further controlled by a protein called a repressor, since it stops the expression of the whole operon by binding to the operator sequence, preventing RNA polymerase from binding to the promoter. Repressors can be removed by relevant chemical signals or in a time-related fashion.

In the ways described earlier, only the genes required at a given moment are expressed. This not only helps to conserve the energy of the cell but also is critical for correct cellular differentiation, tissue pattern formation, and formation of the body plan.

9.2.4 DNA Repair

All living cells appear to possess several different major DNA repair processes (reviews: Walker, 1984; Rossman and Klein, 1988). Such processes are needed to protect cells from the lethal and mutating effects of heat-induced DNA hydrolysis, ultraviolet (UV) light, ionizing radiation, DNA reactive chemicals, free radicals, etc. In single-celled eukaryotes such as the yeast *Saccharomyces cerevisiae*, the number of genes known to be involved in DNA repair approaches 100 (Friedberg, 1988). The number in mammalian cells is expected to be at least equal to this and emphasizes the importance of correction of DNA damage.

9.2.4.1 Excision Repair Some groups of enzymes (light independent) are apparently organized to act cooperatively to recognize DNA lesions, remove them, and correctly replace the damaged sections of DNA. The most comprehensively studied of these is the excision repair pathway.

Briefly, the pathway can be described as follows:

1. *Preincision reactions* UvrA protein dimers are formed which bind to the DNA at a location distant from the damaged site. The UvrB protein then binds to the DNA–UvrA complex to produce an energy-requiring topological unwinding of the DNA via DNA gyrase. This area of unwinding is then translocated, again using ATP as an energy source, to the site of the damaged DNA.
2. *Incision reactions* The UvrC protein binds to the DNA–UvrA/B complex and incises DNA at two sites—seven bases to the 5′ end and three bases to the 3′ end of the damage.
3. *Excision reactions* UvrD protein and DNA polymerase I excise the damaged bases and then resynthesize the strand, using the sister strand as a template. The Uvr complex then breaks down, leaving a restored, but nicked, strand.
4. *Ligation reaction* The nick in the phosphate backbone is repaired by DNA ligase.

A similar excision repair mechanism exists in mammalian cells (see, e.g., Cleaver, 1983). In both cases, the process is regarded as error-free and does not lead to the generation of mutations. However, this pathway can become saturated with excessive numbers of damaged DNA sites, forcing the cell to fall back on other repair mechanisms.

9.2.5 Error-Prone Repair

Exposure of *Escherichia coli* to agents or conditions that either damage DNA or interfere with DNA replication results in the increased expression of the so-called “SOS” regulatory network (Walker, 1984). Included in this network is a group of at least 17 unlinked DNA damage-inducible (*din*) genes. The *din* gene functions are repressed in undamaged cells by the product of the *lexA* gene (Little and Mount, 1982) and are induced when the LexA protein is cleaved by a process that requires modified RecA protein (RecA*), which then acts as a selective protease (Little, 1984). The *din* genes code for a variety of functions, including filamentation, cessation of respiration, etc. Included are the *umuDC* gene products, which are required for so-called “error-prone” or mutagenic DNA repair (Kato and Shinoura, 1977). The precise biochemical mechanism by which this repair is achieved is still not fully understood. Bacterial polymerase molecules have complex activities, including the ability to “proofread” DNA—that is, to ensure that the base-pairing rules of double-stranded DNA are met. It is hypothesized that Umu proteins may suppress this proofreading activity, so that base mismatches are tolerated (Villani et al., 1978). Recent evidence suggests that DNA lesions are bypassed, and this bypass step required UmuDC proteins and RecA* protein (Bridges et al., 1987). The net result is that random base insertion occurs opposite the lesion which may result in mutation.

Analogues of the *umuDC* genes can be found in locations other than the bacterial chromosome—for example, plasmid pKM101 (Walker and Dobson, 1979), a derivative of the drug resistance plasmid R46 (Mortelmans and Strock, 1979), which carried *mucAB* genes (Shanabruch and Walker, 1980) (see pp. 879–880). Mutagenic repair, as controlled by *umuDC*, is not universal even among enterobacteria (Sedgwick and Goodwin, 1985). For instance, *Salmonella typhimurium* LT2 does not appear to express mutagenic repair (Walker, 1984). Thus, the usefulness of strains of this species is greatly enhanced by using derivatives containing plasmids with genes coding for error-prone repair (MacPhee, 1973; McCann et al., 1975a, b).

9.2.6 Mismatch Repair

Mismatched pairs that break the normal base-pairing rules can arise spontaneously due to DNA biosynthetic errors, events associated with genetic recombination and the deam-

ination of methylated cytosine (Modrich, 1987). With the latter, when cytosine deaminates to uracil, an endonuclease enzyme, *N*-uracil-DNA glycosylase (Lindahl, 1979), excises the uracil residue before it can pair with adenine at the next replication. However, 5-methylcytosine deaminates to form thymine and will not be excised by a glycosylase. As a result, thymine exists on one strand paired with guanine on the sister strand, that is, a mismatch. This will result in a spontaneous point mutation if left unrepaired. For this reason, methylated cytosines form spontaneous mutation “hot spots” (Miller, 1985). The cell is able to repair mismatches by being able to distinguish between the DNA strand that exists before replication and a newly synthesized strand.

The mechanism of strand-directed mismatch correction has been demonstrated in *E. coli* (see, e.g., Wagner and Meselson, 1976). In this organism, adenine methylation of d(G-A-T-C) sequences determines the strand on which repair occurs. Thus, parental DNA is fully methylated, while newly synthesized DNA is undermethylated, for a period sufficient for mismatch correction. By this means the organism preserves the presumed correct sequence—that is, that present on the original DNA strand—and removes the aberrant base on the newly synthesized strand. Adenine methylation is achieved in *E. coli* by the *dam* methylase, which is dependent on *S*-adenosylmethionine. Mutants (*dam*) lacking this methylase are hypermutable, as would be expected by this model (Marinus and Morris, 1974).

9.2.7 The Adaptive Repair Pathway

The mutagenic and carcinogenic effects of alkylating agents such as ethyl methanesulfonate are due to the generation of *O*⁶-alkylguanine residues in DNA, which result in point mutations. Bacterial and mammalian cells can repair a limited number of such lesions before DNA replication, thus preventing mutagenic and potentially lethal events taking place.

If *E. coli* are exposed to low concentrations of simple alkylating agents, a repair mechanism is induced that causes increased resistance to subsequent challenge with a high dose. This adaptation response was first described by Samson and Cairns (1977) and has recently been reviewed by Lindahl et al. (1988). The repair pathway is particularly well understood.

9.2.8 Plasmids

Plasmids are extrachromosomal genetic elements that are composed of circular double-stranded DNA. In bacteria some can mediate their own transfer from cell to cell by conjugation—that is, they contain a set of *tra* genes coding for tubelike structures, such as pili, through which a copy of plasmid DNA can pass during transfer.

Plasmids range in size from 1.5 to 200 million daltons. The number of copies per cell differs from plasmid to plasmid. Copy number relates to control of replication, and this correlates with size—that is, small plasmids tend to have large copy numbers per cell. This may relate to a lack of replication control genes (Mortelmans and Dousman, 1986).

9.2.9 Plasmids and DNA Repair

Many plasmids are known to possess three properties: (i) increased resistance to the bactericidal effects of UV and chemical mutagens, (ii) increased spontaneous mutagenesis, and (iii) increased susceptibility to UV and chemically induced mutagenesis. Some plasmids possess all three properties; others may possess just one, for example, increased susceptibility to mutagenesis (review: Mortelmans and Dousman, 1986). Often the profile of activity depends on the DNA repair status of the host cell (Pinney, 1980). Plasmid pKM101 carries DNA repair genes and has been widely used in strains used in bacterial mutagenicity tests.

9.2.10 Nature of Point Mutations

The word “mutation” can be applied to point mutations which are qualitative changes involving one or a few bases in base sequences within genes, as described in the following text, as well as to larger changes involving whole chromosomes (and thus many thousands of genes) and even to changes in whole chromosome sets (described later under cytogenetics).

Point mutations can occur when one base is substituted for another (base substitution). Substitution of another purine for a purine base or of another pyrimidine for pyrimidine is called a transition, while substitutions of purine for pyrimidine or pyrimidine for purine are called transversions. Both types of base substitution have been identified within mutated genes. These changes lead to a codon change which can cause the “wrong” amino acid to be inserted into the relevant polypeptide and are known as missense mutations. Such polypeptides may have dramatically altered properties if the new amino acid is close to the active center of an enzyme or affects the three-dimensional makeup of an enzyme or a structural protein. These changes, in turn, can lead to change or reduction in function, which can be detected as a change in phenotype of the affected cells.

A base substitution can also result in the formation of a new inappropriate terminator (or nonsense) codon and are thus known as nonsense mutations. The polypeptide formed from such mutated genes will be shorter than normal and is most likely to be inactive. Owing to the redundancy of the genetic code, about a quarter of all possible base substitutions will not result in an amino acid replacement and will be silent mutations.

Bases can be deleted or added to a gene. As each gene is of a precisely defined length, these changes, if they involve a number of bases that is not a multiple of three, result in a change in the “reading frame” of the DNA sequence and are thus known as frameshift mutations. Such mutations tend to have a dramatic effect on the polypeptide of the affected gene, as most amino acids will differ from the point of the insertion or deletion of bases onwards. Very often a new terminator codon is produced, so, again, short inactive polypeptides will result.

Both types of mutation result in an altered polypeptide, which, in turn, can have a marked effect on the phenotype of the affected cell. Much use of phenotypic changes is made in mutagenicity tests.

Base substitutions and frameshift changes occur spontaneously and can be induced by radiations and chemical mutagens. It is apparent that the molecular mechanisms resulting in these changes are different in each case, but the potential hazards associated with mutagens capable of inducing the different types of mutation are equivalent.

9.2.11 Suppressor Mutations

In some instances a mutation within one gene can be corrected by a second mutational event at a separate site on the chromosome. As a result, the first defect is suppressed and the second mutation is known as a suppressor mutation. Most suppressor mutations have been found to affect genes encoding for tRNAs. Usually the mutation causes a change in the sequence of the anticodon of the tRNA. Thus, if a new terminator or nonsense codon is formed as the first mutation, this can be suppressed by a second mutation, forming a tRNA species that now has an anticodon complementary to a termination codon. Thus, the new tRNA species will supply an amino acid at the terminator site on the mRNA and allow translation to proceed. Surprisingly most suppressors of this type do not adversely affect cell growth, which implies that the cell can tolerate translation proceeding through termination signals, producing abnormal polypeptides. An alternative explanation is that the particular DNA sequences surrounding normal terminator codons result in a reduced efficiency of suppressor tRNAs (Bossi, 1985).

Frameshift suppression is also possible. This can be achieved by a second mutation in a tRNA gene such that the anticodon of a tRNA molecule consists of four bases rather than three—for example, an extra C residue in the CCC anticodon sequence of a glycine tRNA gene. This change will allow correction of a +1 frameshift involving the GGG codon for glycine (Bossi, 1985).

9.2.12 Adduct Formation

The earlier discussion of adaptive repair made reference to the fact that some unrepaired alkylated bases are lethal, owing to interference with DNA replication, while others

such as *O*⁶-methylguanine lead to mutation if unrepaired. These differences indicate that not all DNA adducts (i.e., DNA bases with additional chemical groups, not associated with normal DNA physiology) are equivalent. In fact, some adducts appear not to interfere with normal DNA functions or are rapidly repaired, others are mutagenic, and yet others are lethal. Chemicals that form electrophilic species readily form DNA adducts. These pieces of information are hard won, and the reader is recommended to read reviews of the pioneering work of Brooks and Lawley (review: Lawley, 1989) summarizing work identifying the importance of DNA adduct formation with polycyclic hydrocarbons and the importance of “minor” products of base alkylation such as *O*⁶-methylguanine and, in addition, the work of the Millers in linking attack of nucleophilic sites in DNA by electrophiles to mutagenesis and carcinogenesis (Miller and Miller, 1971).

If a DNA adduct involves the nitrogen or oxygen atoms involved in base pairing, and the adducted DNA is not repaired, base substitution can result. Adducts can be small, such as the simple addition of methyl or ethyl groups, or they can be very bulky, owing to reaction with multiringed structures. The most vulnerable base is guanine, which can form adducts at several of its atoms (e.g., *N*⁷, *C*⁸, *O*⁶, and exocyclic *N*²) (Venitt and Parry, 1984). Adducts can form links between adjacent bases on the same strand (intrastrand cross-links) and can form interstrand cross-links between each strand of double-stranded DNA.

The induction of frameshift mutation does not necessarily require covalent adduct formation. Some compounds that have a flat, planar structure, such as particulate polycyclic hydrocarbons, can intercalate between the DNA strands of the DNA duplex. The intercalated molecules interfere with DNA repair enzymes or replication and cause additions and deletions of base pairs. The precise mechanism is still unclear, although several mechanisms have been proposed. Hot spots for frameshift mutation often involve sections of DNA where there is a run of the same base—for example, the addition of a guanine to a run of six guanine residues. Such information led to a “slipped mispairing” model for frameshift mutation (Streisinger et al., 1966; Roth, 1974). In this scheme single-strand breaks allow one strand to slip and loop out one or more base pairs, the configuration being stabilized by complementary base pairing at the end of the single-stranded region. Subsequent resynthesis results ultimately in additions or deletions of base pairs (Miller, 1985).

9.2.13 Mutations Due to Insertion Sequences

The subject of mutations due to insertion sequences is reviewed in Cullum (1985). Studies of spontaneous mutation in *E. coli* detected a special class of mutations that were strongly polar, reducing the expression of downstream genes (Jordan et al., 1967). These genes mapped as point mutations

and reverted like typical point mutations. However, unlike point mutations, mutagens did not increase their reversion frequency. Further studies showed that these mutations were due to extra pieces of DNA that can be inserted into various places in the genome. They are not just random pieces of DNA but are “insertion sequences” 0.7–1.5 kb long that can “jump” into other DNA sequences. They are related to transposons, which are insertion sequences carrying easily detected markers such as antibiotic resistance genes, and Mu phages (bacterial viruses).

9.2.14 The Link between Mutation and Cancer

The change in cells undergoing normal, controlled cell division and differentiation to cells that are transformed, dividing without check, and are undifferentiated or abnormally differentiated does not appear to occur as a single step—that is, transformation is multistage. Evidence for this comes from *in vitro* studies, animal models, and clinical observations—in particular, the long latent period between exposure to a carcinogen and the appearance of a tumor in the target tissue. There is much evidence for the sequence of events shown in Figure 9.1 (tumor initiation, promotion, malignant conversion, and progression). Such a scheme provides a useful working model but clearly does not apply to all “carcinogens” in all circumstances.

Study of Figure 9.1 shows that there are several points where genetic change appears to play a role. Such change may occur spontaneously, due to rare errors at cell division such as misreplication of DNA or spindle malfunction, or may be induced by exposure to viruses (e.g., acute transforming retroviruses), ionizing and nonionizing radiations absorbed by DNA (e.g., X-rays, UVC), or particular chemical species capable of covalently interacting with DNA (as discussed earlier) or with vital proteins, such as tubulin, that polymerize to form the cell division spindle apparatus.

9.2.15 Genotoxic versus Nongenotoxic Mechanisms of Carcinogenesis

The previous discussions of oncogene activation and human DNA repair deficiencies provide strong evidence for carcinogenesis via genotoxic mechanisms. However, it has been recognized for many years that cancers can arise without biologically significant direct or indirect interaction between a chemical and cellular DNA (see, e.g., Gatehouse et al., 1988). The distinction between nongenotoxic and genotoxic carcinogens has recently been brought into a sharper focus following the identification of a comparatively large number of “nongenotoxic” carcinogens by the US National Toxicology Program (Tennant et al., 1987). These include a wide range of chemicals acting via a variety of mechanisms, including augmentation of high

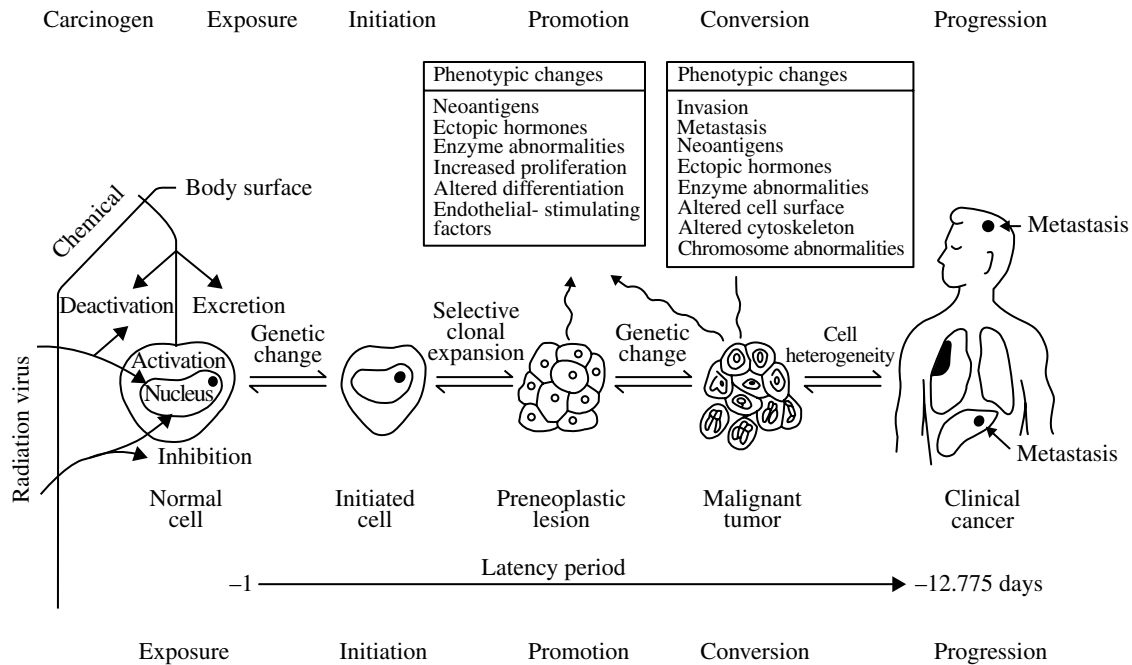


FIGURE 9.1 Schematic representation of events leading to neoplasia. Source: Adapted from Harris et al. (1987).

“spontaneous” tumor yields, disruption of normal hormonal homeostasis in hormone-responsive tissues, peroxisome proliferation, proliferation of urothelial cells following damage via induced kidney stones, etc. (Clayson, 1989). This author points out that a major effort is under way to determine whether many of these compounds can elicit similar effects in humans.

Ashby and Tennant (1988) and Ashby et al. (1989) stress the significance of their observations that 16 tissues are apparently sensitive to genotoxic carcinogens, while a further 13 tissues are sensitive to both genotoxic and nongenotoxic carcinogens (Table 9.2). Also, genotoxic carcinogens tend to induce tumors in several tissues of both males and females in both rats and mice. This contrasts with nongenotoxic carcinogens, which may induce tumors at high doses, in one tissue, of one sex, or of one species. Although it is most unlikely that all nongenotoxic carcinogens will prove to be irrelevant in terms of human risk, it appears from the earlier analysis that a proportion of carcinogens identified by the use of near-toxic levels in rodent bioassays are of dubious relevance to the induction of human cancer. For further discussion, see Butterworth and Slaga (1987).

9.2.16 Genetic Damage and Heritable Defects

Concern about the effects of radiations and chemicals on the human gene pool, and the resulting heritable malformations and syndromes, has steadily risen during this century. The recognition that changes in morphology would result from changes in the hereditary material due to mutations (from

TABLE 9.2 Tissues Sensitive to Genotoxic and/or Nongenotoxic Carcinogens

| Tissues Sensitive Primarily to Genotoxins | Tissues Sensitive to Both Genotoxins and Nongenotoxins |
|---|--|
| Stomach | Nose |
| Zymbal's gland | Mammary gland |
| Lung | Pituitary gland |
| Subcutaneous tissue | Integumentary system |
| Circulatory system | Kidney |
| Clitoral gland | Urinary bladder |
| Skin | Liver |
| Intestine/colon | Thyroid gland |
| Uterus | Hematopoietic system |
| Spleen | Adrenal gland |
| Tunica vaginalis | Pancreas |
| Bile duct | Seminal vesicle |
| Ovary | Urinary tract |
| Harderian gland | Lymphatic system |
| Preputial gland | |
| (Multiple organ sites) | |

the Latin word *mutare*, to change) was adopted by de Vries following observations on the evening primrose, *Oenothera* (deVries, 1901). Muller went on to demonstrate that X-rays could induce mutations in the germ cells of the fruit fly *Drosophila melanogaster* (Muller, 1927).

The human gene pool is known to carry many deleterious genes acquired from preceding generations which result in numerous genetic diseases. It is clear that these arise as a

result of DNA changes affecting particular chromosomes or genes. They can be grouped as follows:

1. Chromosome abnormalities, small changes in either number or structure
2. Autosomal dominant gene mutations, in which a change in only one copy of the pair of genes is sufficient for the condition to be expressed
3. Autosomal recessive gene mutations in which both copies of a gene must be mutated for the trait to become manifest
4. Sex-linked conditions, which may also be recessive or dominant, where the mutant gene is on an X chromosome and will be expressed at high frequency in males (XY) and at a much lower frequency in females (XX), if the gene acts in a recessive manner
5. Polygenic mutations, in which the condition results from the interaction of several genes and may include an environmental component

9.2.17 Reproductive Effects

If a potent genotoxin is able to cross the placental barrier, it is very likely to interfere with differentiation of the developing embryo and thus possess teratogenic potential. Indeed, many of the better studied teratogens are also mutagenic (Kalter, 1977). However, mutagens form only one class of teratogens, and a large proportion of teratogens are not mutagenic. Alternative mechanisms of teratogenesis include cell death, mitotic delay, retarded differentiation, vascular insufficiency, inhibited cell migration, etc. (Beckman and Brent, 1986).

It is known that more fetal wastage and many spontaneous abortions arise as a result of the presence of dominant lethal mutations in the developing embryo, many of which appear to be due to major chromosomal damage. In addition, impairment of male fertility may also be a consequence of exposure to mutagens.

9.3 CYTOGENETICS

There are various types of cytogenetic change which can be detected in chromosomes. These are structural CAs, numerical changes which could result in aneuploidy, and sister chromatid exchanges (SCEs). CA assays are used to detect the induction of chromosome breakage (clastogenesis) in somatic or germinal cells by direct observation of chromosomal damage during metaphase analysis or by indirect observation of micronuclei. Chromosomal damage detected in these assays is mostly lethal to the cell during the cell cycle following the induction of the damage. Its presence, however, indicates a potential to induce more subtle chromosomal damage which survives cell division to

produce heritable cytogenetic changes. Cytogenetic damage is usually accompanied by other genotoxic damage such as gene mutation.

9.3.1 Cytogenetic Damage and Its Consequences

Structural and numerical CAs in somatic cells may be involved in the etiology of neoplasia and in germ cells which can lead to perinatal mortality, dominant lethality, or congenital malformations in the offspring (Chandley, 1981), and some tumors (Anderson, 1990).

Chromosome defects arise at the level of the individual chromosome or at the level of the chromosomal set, so affecting chromosomal number.

9.3.2 Individual Chromosomal Damage

Damage to individual chromosomes consists of breakage of chromatids, which must result from a discontinuity of both strands of the DNA in a chromatid. How mutagens produce chromosome breakage is not totally understood, but DNA lesions which are not in themselves discontinuities will produce breakage of a chromosome as a consequence of their interference with the normal process of DNA replication. In haploid microorganisms and prokaryotes, chromosome breaks are usually lethal, but not in diploid eukaryotes. According to Bender et al. (1974), in these organisms chromosome breaks may reconstitute in the same order, probably as a result of an enzyme repair process, resulting in no apparent cytogenetic damage; they may remain unjoined as fragments, which could result in cell death at the next or following mitoses—if, for example, unrejoined fragments are introduced into the zygote via treated germ cells, the embryo may die at a very early stage from a dominant lethal mutation, or they may rejoin in a different order from the original one, producing chromosomal rearrangements. There are various types of chromosomal rearrangements:

Reciprocal translocations can result from the exchange of chromosomal segments between two chromosomes, and, depending on the position of the centromeres in the rearranged chromosomes, different configurations will result.

1. Asymmetrical exchanges arise when one of the rearranged chromosomes carries both centromeres and is known as dicentric while the other carries none and is acentric. The cell or zygote carrying this anomaly usually dies, death being caused by segregation difficulties of the dicentric or the loss of the acentric fragment at cell division. Such a translocation contributes to dominant lethality.
2. Symmetrical exchanges occur when each rearranged chromosome carries just one centromere. This allows the zygote to develop normally, but when such hetero-

zygous form germ cells at meiosis, about half of their gametes will be genetically unbalanced, since they have deficiencies and duplications of chromosomal material. The unbalanced gametes which survive produce unbalanced zygotes, which results in death shortly before and after birth or congenital malformations.

Centric fusions involve the joining together of two chromosomes, each of which has a centromere at or near one end, to produce a single metacentric or submetacentric chromosome. When such translocations are produced in a germ cell and result from breakage and rejoining in the short arms of the two chromosomes, as a consequence of loss of the derived acentric fragments, a genetic deficiency can result. Some Robertsonian translocations are able to survive, but others pose a risk. In heterozygotes the two arms of the translocation chromosome may pair with the two separate homologous chromosomes at meiosis but segregate in a disorderly manner. Some of the resultant germ cells lack copies (nullisomy) or carry two copies (disomy) of one or other of the two chromosomes involved, which results in monosomic or trisomic embryos. Monosomics die early, but trisomic embryos, which carry three copies of a chromosome, can survive to birth or beyond. If chromosome 21 is involved in the translocation, it can form a translocation trisomy and produce inherited Down syndrome (this differs from nondisjunctional Down syndrome trisomy).

Deletions and deficiencies are produced when two breaks arise close together in the same chromosome. The two ends of the chromosome join when the fragment between the breaks becomes detached. At the next cell division, the unattached piece of chromosome is likely to be lost. Large deletions may contribute to dominant lethality. Small deletions are difficult to distinguish from point mutations. Deletions may uncover preexisting recessive genes. If one gene that is essential for survival is uncovered, it can act as a lethal in a homozygote and as a partial dominant in a heterozygote.

Inversions occur when two breaks occur in the same chromosome. The portion between them is detached and becomes reinserted in the opposite way to its original position, that is, the gene order is reversed. This need not cause a genetic problem, but imbalanced gametes could result in congenital malformation or fetal death.

9.3.3 Chromosome Set Damage

Accuracy of chromosome replication and segregation of chromosomes to daughter cells requires accurate maintenance of the chromosome complement of a eukaryotic cell. Chromosome segregation in meiosis and mitosis is dependent upon the synthesis and functioning of the proteins of the spindle apparatus and upon the attachment and movement

of chromosomes on the spindle. The kinetochores attach the chromosomes to the spindle, and the centrioles are responsible for the polar orientation of the division apparatus. Sometimes such segregation events proceed incorrectly and homologous chromosomes separate, with deviations from the normal number (aneuploidy) into daughter cells or as a multiple of the complete karyotype (polyploidy). When both copies of a particular chromosome move into a daughter cell and the other cell receives none, the event is known as nondisjunction.

Aneuploidy in live births and abortions arises from aneuploid gametes during germ cell meiosis. Trisomy or monosomy of large chromosomes leads to early embryonic death. Trisomy of the smaller chromosomes allows survival but is detrimental to the health of an affected person—for example, Down syndrome (trisomy 21), Patau syndrome (trisomy 13), and Edward syndrome (trisomy 18). Sex chromosome trisomies (Klinefelter's and XXX syndromes) and the sex chromosome monosomy (XO), known as the Turner syndrome, are also compatible with survival.

Aneuploidy in somatic cells is involved in the formation of human tumors. Up to 10% of tumors are monosomic and trisomic for a specific chromosome as the single observable cytogenetic change. Most common among such tumors are trisomy 8, 9, 12, and 21 and monosomy for chromosomes 7, 22, and Y.

9.3.4 Test Systems

In vivo and *in vitro* techniques are available to test mutagenic properties to demonstrate presence or lack of ability of the test material to cause mutation or chromosomal damage or cause cancer, as summarized in Table 9.3. The material intended for intimate contact and long exposure should not have any genotoxic properties. The presence of unpolymerized materials and traces of monomers, oligomers, additives, or biodegradation products can cause mutations. Mutation can be a point mutation or chromosomal rearrangement caused by DNA damage. Therefore the material's ability to cause point mutation, chromosomal change, or evidence of DNA damage is tested. As we have seen, correlations exist between mutagenic and carcinogenic properties. Most carcinogens are mutagens, but not all mutagens are human carcinogens.

The Ames salmonella/microsome test is a principal sensitive mutagen screening test (McCann et al., 1975a; McCann and Ames, 1976). Compounds are tested on the mutants of *S. typhimurium* for reversion from a histidine requirement back to prototrophy. A positive result is seen by the growth of revertant bacteria (which do not require an external histidine source). A microsomal activation system should be included in this assay. The use of five different bacterial test strains is generally required.

TABLE 9.3 Fifteen Common Assays Described by OECD

| | In Vitro | In Vivo |
|---|----------|---------|
| <i>Assays for gene mutations</i> | | |
| <i>Salmonella typhimurium</i> reverse mutation assay (Ames test, bacteria) (OECD 471) | ✓ | |
| <i>Escherichia coli</i> reverse mutation assay (bacteria) (OECD 472) | ✓ | |
| Gene mutation in mammalian cells in culture (OECD 476) | ✓ | |
| <i>Drosophila</i> sex-linked recessive lethal assay (fruit fly) (OECD 477) | | ✓ |
| Gene mutation in <i>Saccharomyces cerevisiae</i> (yeast) (OECD 480) | ✓ | |
| Mouse spot test (OECD 484) | | ✓ |
| <i>Assays for chromosomal and genomic mutations</i> | | |
| <i>In vitro</i> cytogenetic assay (OECD 473) | ✓ | |
| <i>In vivo</i> cytogenetic assay (OECD 475) | | ✓ |
| Micronucleus test (OECD 474) | | ✓ |
| Dominant lethal assay (OECD 478) | | ✓ |
| Heritable translocation assay (OECD 485) | | ✓ |
| Mammalian germ cell cytogenetic assay (OECD 483) | | ✓ |
| <i>Assays for DNA effects</i> | | |
| DNA damage and repair: unscheduled DNA synthesis <i>in vitro</i> (OECD 482) | ✓ | |
| Mitotic recombination in <i>S. cerevisiae</i> (yeast) (OECD 481) | ✓ | |
| <i>In vitro</i> sister chromatid exchange assay (OECD 479) | ✓ | |

Two mammalian mutagenicity tests (one *in vitro*, one *in vivo*) are generally required to support the lack of mutagenic or carcinogenic potential. Some well-known tests are:

- The L5178Y mouse lymphoma assay (MLA) for mutants at the TK locus
- The induction of recessive lethals in *D. melanogaster*
- Metaphase analysis of cultured mammalian cells and of treated animals
- SCE assay
- Unscheduled DNA synthesis (UDS) assay
- Cell transformation assay
- Comet assay
- Gene mutation in cultured mammalian cells such as Chinese hamster V79 cell/hypoxanthine-guanine phosphoribosyltransferase (HGPRT) mutation system

ICH guidelines specifically require three genotoxicity assays for all devices (see Table 9.1). The assays should preferably evaluate DNA effects, gene mutations, and CAs, and two of the assays should preferably use mammalian cells. Guidance for providing tests for selection to meet these needs is the OECD guidelines, which include eight *in vitro* and seven *in vivo* assays.

9.3.5 *In Vitro* Test Systems

The principal tests can be broadly categorized into microbial and mammalian cell assays. In both cases the tests are carried out in the presence and absence of *in vitro* metabolic activation enzymes, usually derived from rodent liver.

9.3.5.1 *In Vitro* Metabolic Activation The target cells for *in vitro* mutagenicity tests often possess a limited (often overlooked) capacity for endogenous metabolism of xenobiotics. However, to simulate the complexity of metabolic events that occur in the whole animal, there is a critical need to supplement this activity.

Choice of Species A bewildering variety of exogenous systems have been used for one purpose or another in mutagenicity tests. The choice begins with plant or animal preparations. The attraction of plant systems has stemmed from a desire to avoid the use of animals, where possible, in toxicity testing. In addition, plant systems have particular relevance when certain chemicals are being tested, for example, herbicides.

If animal systems are chosen, preparations derived from fish (see, e.g., Kada, 1981) and birds (Parry et al., 1985) have been used. However, by far the most widely used and validated are those derived from rodents—in particular the rat. Hamsters may be preferred as a source of metabolizing enzymes when particular chemical classes are being screened—for example, aromatic amines, heterocyclic amines, *N*-nitrosamines, and azo dyes (Prival and Mitchell, 1982; Haworth et al., 1983).

Choice of Tissue The next choice is that of source tissue. Preparations derived from liver are the most useful, as this tissue is a rich source of mixed-function oxygenases capable of converting procarcinogens to genetically active electrophiles. However, many extrahepatic tissues (e.g., kidney, lung, etc.) are also known to possess important metabolic capacity which may be relevant to the production of mutagenic metabolites in the whole animal.

Cell-Free versus Cell-Based Systems Most use has been made of cell-free systems—in particular crude homogenates such as 9000g supernatant (S9 fraction) from rat liver. This fraction is composed of free endoplasmic reticulum, microsomes (membrane-bound packets of “membrane-associated” enzymes), soluble enzymes, and some cofactors. Hepatic S9 fractions do not necessarily completely reflect the metabolism of the whole organ, in that they mainly possess phase I metabolism (e.g., oxygenases) and are deficient in phase II systems (e.g., conjugation enzymes). The latter are often capable of efficient detoxification, while the former are regarded as “activating.” This can be a strength in that S9 fractions are used in screening tests as a surrogate for all tissues in animals, some of which may be exposed to reactive metabolites in the absence of efficient detoxification. Many carcinogens are organ specific in extrahepatic tissues, yet liver S9 fraction will reveal their mutagenicity. The deficiency of S9 fractions for detoxification can also be a weakness, in that detoxification may predominate in the whole animal, such that the potential carcinogenicity revealed *in vitro* is not realized *in vivo*.

Cell-free systems, when supplemented with relevant cofactors, are remarkably proficient, despite their crudity in generating reactive electrophiles from most procarcinogens. However, they provide at best a broad approximation of *in vivo* metabolism and can fail to produce sufficient quantity of a particular reactive metabolite to be detectable by the indicator cells, or they can produce inappropriate metabolites that do not play a role *in vivo* (see Gatehouse and Tweats (1987) for discussion).

Some of these problems can be overcome by the use of cell-based systems—in particular primary hepatocytes. Hepatocytes closely simulate the metabolic systems found in the intact liver and do not require additional cofactors for optimal enzyme activity. However, apart from greater technical difficulties in obtaining hepatocytes as opposed to S9 fraction, hepatocytes can effectively detoxify particular carcinogens and prevent their detection as mutagens. Despite these difficulties, hepatocytes have a role to play in mutagenicity screening, in both bacterial- and mammalian-based systems (Tweats and Gatehouse 1988).

Inducing Agents The final choice considered here is whether to use “uninduced” liver preparations or those derived from animals pretreated with an enzyme inducer to promote high levels of metabolic activity. If induced preparations are preferred, which inducer should be used?

It appears that uninduced preparations are of limited use in screening assays, as they are deficient in particular important activities such as cytochrome P450_{IA1} oxygenases. In addition, species and organ differences are most divergent with uninduced enzyme preparations (Brusick, 1987a).

The aforementioned differences disappear when induced microsomal preparations are used. A number of enzyme

inducers have been used, the most popular being Aroclor 1254, which is a mixture of polychlorinated biphenyls (as described by Ames et al., 1975). However, concern about the toxicity, carcinogenicity, and persistence of these compounds in the environment has led to the use of alternatives, such as a combination of phenobarbitone (phenobarbital) and β -naphthoflavone (5,6-benzoflavone). This combination results in the induction of a range of monooxygenases similar to that induced by Aroclor 1254 (see e.g., Ong et al., 1980). More selective inducers such as phenobarbitone (cytochrome P450_{IIA1}, P-450_{IIB1}) or 3-methylcholanthrene (cytochrome P450_{IA1}) have also been used.

In summary, genetic toxicity tests with both bacterial and mammalian cells are normally carried out with rat liver cell-free systems (S9 fraction) from animals pretreated with enzyme inducers. However, investigations should not slavishly follow this regimen: there may be sound scientifically based reasons for using preparations from different species or different organs or for using whole cells such as hepatocytes.

Standard Method of S9 Fraction Preparation The following method describes the production of hepatic S9 mix from rats induced with a combination of phenobarbitone and β -naphthoflavone and is an adaptation of the method described by Gatehouse and Delow (1979).

Male albino rats within the weight range 150–250g are treated with phenobarbitone sodium 16 mg mL⁻¹, 2.5 mL kg⁻¹ in sterile saline and β -naphthoflavone 20 mg mL⁻¹ in corn oil. A fine suspension of the latter is achieved by sonicating for 1 h. These solutions are dosed by intraperitoneal injection on days 1, 2, and 3.

Phenobarbitone sodium is normally administered between 0.5 and 2 h prior to β -naphthoflavone.

The animals are killed on day 4 by cervical dislocation and the livers removed as quickly as possible and placed on ice-cold KCl buffer (0.01 M Na₂HPO₄ + KCl 1.15%). The liver is cleaned, weighted, minced, and homogenized (in an Ultra-Turrax homogenizer) in the aforementioned buffer to give a 25% (w/v) liver homogenate. The homogenate is stored at 4°C until it can be centrifuged at 9000g for 15 min. The supernatant is decanted, mixed, and divided into 2 mL volumes in cryotubes. These are then snap-frozen in liquid nitrogen. Storage at -196°C for up to 3 months results in no appreciable loss of most P450 isoenzymes (Ashwood-Smith, 1980).

Quality control of S9 batches is usually monitored by ability to activate compounds known to require metabolism to generate mutagenic metabolites. This is a rather crude approach, and more accurate data can be obtained by measuring biochemical parameters—for example, protein, cytochrome P450 total activity (from crude S9) and related enzyme activities (from purified microsomes) such as 7-ethoxyresorufin-*O*-deethylase and 7-methoxycoumarin-*O*-demethylase to give an indication of S9 batch-to-batch variation and to set standards for rejecting suboptimal batches (Hubbard et al., 1985). For

further details on critical features affecting the use and limitations of S9 fraction, see Gatehouse and Tweats (1987).

S9 Mix The S9 fraction prepared as described earlier is used as a component in “S9 mix” along with buffers and various enzyme cofactors. The amount of S9 fraction in the S9 mix can be varied, but a “standard” level of 0.1 mL mL⁻¹ of S9 mix (or 10% S9) is often recommended for general screening.

No single concentration of S9 fraction in the S9 mix will detect all classes of genotoxic carcinogen with equal efficiency (Gatehouse et al., 1990). Some mutagens, including many polycyclic aromatic hydrocarbons, are activated to mutagens by higher than normal levels of S9 fraction in the S9 mix (see, e.g., Carver et al., 1985).

The mixed-function oxidases in the S9 fraction require NADPH, normally generated from the action of glucose-6-phosphate dehydrogenase acting on glucose-6-phosphate and reducing NADP, both of which are normally supplied as cofactors. As an alternative, isocitrate can be substituted for glucose-6-phosphate (to be used as a substrate by isocitrate dehydrogenase) (Linblad and Jackim, 1982). Additional cofactors may be added (e.g., flavin mononucleotide), when particular classes of compound such as azo dyes are being tested (Prival et al., 1984), or acetyl coenzyme A when aromatic amines such as benzidine are being tested (Kennelly et al., 1984). The composition of a “standard” S9 mix is given in Table 9.4.

9.3.6 Bacterial Mutation Tests

The study of mutation in bacteria (and bacterial viruses) has had a fundamental role in the science of genetics in the twentieth century. In particular, the unraveling of biochemical

anabolic and catabolic pathways, the identification of DNA as the hereditary material, the fine structure of the gene, the nature of gene regulation, etc. have all been aided by bacterial mutants.

As an offshoot of studies of genes concerned with the biosynthesis of amino acids, a range of *E. coli* (see, e.g., Yanofsky, 1971) and *S. typhimurium* strains (see e.g., Ames, 1971; Ames et al., 1973; Ames and McCann, 1981) with relatively well-defined mutations in known genes became available. Thus, bacteria already mutant at an easily detectable locus are treated with a range of doses of the test material to determine whether the compound can induce a second mutation that directly reverses or suppresses the original mutations. Thus, for amino acid auxotrophs, the original mutation has resulted in loss of ability to grow in the absence of the required amino acid. The second mutation restores prototrophy—that is, the affected cell is now able to grow in the absence of the relevant amino acid, if provided with inorganic salts and a carbon source. This simple concept, in fact, underlines the great strength of these assays, for it provides enormous selective power which can identify a small number of the chosen mutants from a population of millions of unmutated cells and cells mutated in other genes. The genetic target—that is, the mutated DNA bases in the gene in question (or bases in the relevant tRNA genes; see the discussion of suppressor mutations)—can thus be very small, just one or a few bases in length.

An alternative approach is to use bacteria to detect “forward mutations.” Genetic systems which detect forward mutations have an apparent advantage, in that a wide variety of genetic changes may lead to a forward mutation—for example, point mutation, deletions, insertions, etc. In addition, forward mutations in a number of different genes may lead to the same change in phenotype; thus, the genetic target is much larger than that seen in most reverse mutation assays. However, if a particular mutagen causes rare specific changes, these changes may be lost against the background of more common events (Gatehouse et al., 1990, 1994). Spontaneous mutation rates tend to be relatively high in forward mutation systems. Acquisition of resistance to a toxic chemical (e.g., an amino acid analog or antibiotic) is a frequently used genetic marker in these systems. For instance, the use of resistance to the antibiotic streptomycin preceded the reversion assays in common use today.

9.3.6.1 Reversion Tests: Background There are several excellent references describing the background and use of bacteria for reversion tests (Brusick, 1987a, b; Gatehouse et al., 1990). Three different protocols have been widely used: plate incorporation assays, treat and plate tests, and fluctuation tests. These methods are described in detail in the following sections. Fundamental to the operation of these tests is the genetic compositions of the tester strains selected for use.

TABLE 9.4 Composition of Standard S9 Mix

| Constituent | Final Concentration in Mix (mM) |
|--|------------------------------------|
| Glucose-6-phosphate | 5 |
| Nicotinamide adenine dinucleotide phosphate | 4 |
| MgCl ₂ ·6H ₂ O | } salt solution 8 |
| KCl | |
| Phosphate buffer (90.2 M) | 33 |
| Distilled water to make up to the required volume | 100 |
| S9 fraction added at 0.1 mL mL ⁻¹ of S9 mix | |

For assays using cultured mammalian cells, phosphate buffer and distilled water are replaced by tissue culture medium, as high concentrations of Na and K salts are toxic to such cells. The concentration of S9 fraction in the S9 mix varies, depending on the relevant assay (see individual sections). Once prepared, S9 mix should be used as soon as possible and should be stored on ice until required. S9 fraction, once thawed, should not be refrozen for future use.

TABLE 9.5 Genotype of Commonly Used Strains of *Salmonella typhimurium* LT2 and Their Reversion Events

| Strain | Genotype | Reversion Events |
|--------|---|---|
| TA1535 | hisG ₄₆ rfa \bigcirc gal chlD bio uvrB | Subset of base-pair substitution events |
| TA100 | hisG ₄₆ \bigcirc rfa gal chlD bio uvrB (pKM101) | Subset of base-pair substitution events |
| TA1537 | hisC ₃₀₇₆ \bigcirc rfa gal chlD bio uvrB | Frameshifts |
| TA1538 | hisD ₃₀₅₂ \bigcirc rfa gal chlD bio uvrB | Frameshifts |
| TA98 | hisD ₃₀₅₂ \bigcirc rfa gal chlD bio uvrB pKM101) | Frameshifts |
| TA97 | hisD ₆₆₁₀ hisO ₁₂₄₂ rfa \bigcirc gal chlD bio uvrB (pKM101) | Frameshifts |
| TA102 | his \bigcirc (G) ₈₄₇₆ rfa galE (pAQ1) (pKM101) | All possible transitions and transversions, small deletions |

9.3.6.2 Genetic Makeup of Tester Strains The most widely used strains are those developed by Bruce Ames and colleagues which are mutant derivatives of the organism *S. typhimurium*. Each strain carries one of a number of mutations in the operon coding for histidine biosynthesis. In each case the mutation can be reverted either by base change or by frameshift mutations. The genotype of the commonly used strains is shown in Table 9.5.

9.3.6.3 The Use of the Plasmid pKM101 *S. typhimurium* LT2 strains do not appear to possess classical “error-prone” repair as found in *E. coli* strains and some other members of the enterobacteria (Walker, 1984; Sedgwick and Goodwin, 1985). This is due to a deficiency in *umuD* activity in these *Salmonella* strains (Herrera et al., 1988; Thomas and Sedgwick, 1989). One way to overcome this deficiency and to increase sensitivity to mutagens is to use strains containing a plasmid carrying analogs to the *umuDC* genes, such as are present in the pKM101 plasmid.

9.3.6.4 Ames Salmonella/Plate Incorporation Method

The following procedure is based on that described by Ames and colleagues (Kado et al., 1983; Maron and Ames, 1983), with additional modifications:

1. Each selected test strain is grown for 10h at 37°C in nutrient broth (Oxoid No. 2) or supplemented minimal media (Vogel–Bonner) on an orbital shaker. A timing device can be used to ensure that cultures are ready at the beginning of the working day.
2. 2.0mL aliquots of soft agar overlay medium are melted just prior to use and cooled to 50°C, and relevant supplements added—that is, L-histidine, final concentration 9.55 $\mu\text{g mL}^{-1}$, and D-biotin 12 $\mu\text{g mL}^{-1}$. (N.B.: If *E. coli* WP2 tester strains are used, the only supplement required is tryptophan 3.6 $\mu\text{g mL}^{-1}$.) The medium is kept semimolten by holding the tubes containing the medium in a hot aluminum dry block, held at 45°C. It is best to avoid water baths as microbial contamination can cause problems.
3. The following additions are made to each tube of top agar: the test article (or solvent control) in solution

(10–200 μL), the test strain (100 μL), and, where necessary, S9 mix (500 μL). The test is carried out in the presence and absence of S9 mix. The exact volume of test article or solvent may depend on toxicity or solubility, as described in the preceding section.

4. There should be at least three replicate plates per treatment with at least five test doses plus untreated controls. Duplicate plates are sufficient for the positive and sterility control treatments. The use of twice as many negative control plates as used in each treatment group will lead to more powerful tests from a statistical standpoint (Mahon et al., 1989).
5. Each tube of top agar is mixed and quickly poured onto dried prelabeled Vogel–Bonner basal agar plates.
6. The soft agar is allowed to set at room temperature, and the plates are inverted and incubated (within 1 h of pouring) at 37°C in the dark. Incubation is continued for 2–3 days.
7. Before scoring the plates for revertant colonies, the presence of a light background lawn of growth (due to limited growth of nonrevertant colonies before the trace of histidine or tryptophan is exhausted) should be confirmed for each concentration of test article by examination of the plate under low power of a light microscope. At concentrations that are toxic to the test strains, such a lawn will be depleted, and colonies may appear that are not true revertants but surviving, nonprototrophic cells. If necessary, the phenotype of any questionable colonies (pseudorevertants) should be checked by plating on histidine- or tryptophan-free medium.
8. Revertant colonies can be counted by hand or with an automatic colony counter. Such machines are relatively accurate in the range of colonies normally observed (although regular calibration against manual counts is a wise precaution). Where accurate quantitative counts of plates with large numbers of colonies are required, only manual counts will give accurate results.

9.3.7 Controls

9.3.7.1 Positive Controls Where possible, positive controls should be chosen that are structurally related to the test article. This increases the confidence in the results. In the absence of structurally related mutagens, the set of positive controls given in Table 9.6 can be used. The use of such controls validates each test run and helps to confirm the nature of each strain. Pagano and Zeiger (1985) have shown that it is possible to store stock solutions of most routinely used positive controls (sodium azide, 2-aminoanthracene, benzo[*a*]pyrene, 4-nitroquinoline oxide) at -20 to -80°C for several months, without loss of activity. This measure can help reduce potential exposure to laboratory personnel.

9.3.7.2 Untreated/Vehicle Controls Untreated controls omit the test article but are made up to volume with buffer. The vehicle control is made up to volume with the solvent used to dissolve the test substance. It is preferable to ensure that each of the treated plates contains the same volume of vehicle throughout.

As detailed by Gatehouse and Tweats (1987), the nature and concentration of solvent may have a marked effect on the test result. Dimethyl sulfoxide (DMSO) is often used as the solvent of choice for hydrophobic compounds. However, there may be unforeseen effects, such as an increase in mutagenicity of some compounds—for example, *p*-phenylenediamine (Burnett et al., 1982)—or a decrease in mutagenicity of others, such as simple aliphatic nitrosamines (Yahagi et al., 1977). It is essential to use fresh batches of the highest purity grade available and to prevent decomposition/oxidation on storage. The products after oxidation are both toxic and can induce base-pair substitutions in both bacterial and mammalian assays. Finally, DMSO and

other organic solvents can inhibit the oxidation of different premutagens by microsomal monooxygenases (Wolff, 1977a, b). To reduce the risk of artifactual results, it is essential to use the minimum amount of organic solvent (e.g., $<2\%$ w/w) compatible with adequate testing of the test chemical.

It is important to keep a careful check of the number of mutant colonies present on untreated or vehicle control plates. These numbers depend on the following factors:

1. The repair status of the cell—that is, excision repair-deficient strains tend to have more “spontaneous mutants” than repair-proficient cells.
2. The presence of mutator plasmids. Excision-deficient strains containing pKM101 have a higher spontaneous mutation rate at both base substitution and frameshift loci than excision-proficient strains.
3. The total number of cell divisions that take place in the supplemented top agar. This is controlled by the supply of nutrients—in particular histidine. Rat liver extracts may also supply trace amounts of limiting nutrients, resulting in a slight increase in the spontaneous yield of mutants in the presence of S9 mix.
4. The size of the initial inoculum. During growth of the starting culture, mutants will arise. Thus, if a larger starting inoculum is used, more of these “preexisting” mutants will be present per plate. In fact, the “plate mutants” arising as described in point (3) predominate.
5. The intrinsic mutability of the mutation in question. In practice the control mutation values tend to fall within a relatively precise range for each strain. Each laboratory should determine the normal range of revertant colonies per plate for each strain.

TABLE 9.6 Positive Controls for Use in Plate Incorporation Assays

| Species | Strain | Mutagen | Concentration ($\mu\text{g plate}^{-1}$) ^a |
|-------------------------------|-------------------|-----------------------------|---|
| (a) In the absence of S9 mix | | | |
| <i>Salmonella typhimurium</i> | TA1535 | Sodium azide | 1–5 |
| | TA100 | | |
| | TA1538 | Hycanthone methanesulfonate | 5–20 |
| | TA98 | | |
| | TA1537 | | |
| <i>Escherichia coli</i> | WP2 uvrA | ICR 191 | 1 |
| | | Nifuroxime | 5–15 |
| (b) In the presence of S9 mix | | | |
| <i>E. coli</i> | WP2 uvrA (pKM101) | | |
| <i>S. typhimurium</i> | TA1538 | 2-Aminoanthracene | 1–10 |
| | TA1535 | | |
| | TA100 | | |
| | TA90 | | |
| | TA1537 | Neutral red | 10–20 |

^a The concentration given earlier will give relatively small increases in revertant count above the spontaneous level. There is little point in using large concentrations of reference mutagens which invariably give huge increases in revertant counts. This would give little information on the day-to-day performance of the assay.

Deviations in background reversion counts from the normal range should be investigated. It is possible that cross-contamination, variations in media quality, etc. have occurred that may invalidate particular experiments.

Frequent checks should also be made on the sterility of S9 preparations, media, and test articles. These simple precautions can prevent loss of valuable time and resources.

9.3.7.3 Evaluation of Results At least two independent assays are carried out for each test article. The criterion for positive response is a reproducible and statistically significant result at any concentration for any strain. When positive results are obtained, the test is repeated, using the strain(s) and concentration range with which the initial positive results were observed. This range may be quite narrow for toxic agents.

Several statistical approaches have been applied to the results of plate incorporation assays (Mahon et al., 1989). These authors make a number of important suggestions to maximize the power of statistical analyses; those that relate to the method of analysis are reproduced as follows:

1. Unless it is obvious that the test agent has had no effect, the data should be plotted, to give a visual impression of the form of any dose-response and the pattern of variability.
2. Three methods of analysis—linear regression (Armitage, 1971; Steel and Torrie, 1996; Gad, 2005); a multiple comparison analysis, such as Dunnett's method (Dunnett, 1955); and a nonparametric analysis, such as Kruskal-Wallis (Gad, 2005)—can all be recommended. Each has its strengths and weaknesses, and other methods are not excluded.
3. Linear regression assumes that variance across doses is constant and that the dose-response is linear. If the variance is not approximately constant, then a transformation may be applied or a weighted analysis may be carried out. If the dose scale tends to a plateau, then the dose scale may be transformed. If counts decline markedly at high doses, then linear regression is inappropriate.
4. Dunnett's method, perhaps with a transformation, is recommended when counts decline markedly at one or two high doses. However, when the dose-response shows no such decline, other methods may be more powerful.
5. Kruskal-Wallis's nonparametric method avoids the complications of transformations of weighting and is about as powerful as any other method. However, it is inappropriate when the response declines markedly at high dose.

9.3.7.4 Preincubation Tests Some mutagens are poorly detected in the standard plate incorporation assay, particularly those that are metabolized to short-lived reactive electro-

philes—for example, short-chain aliphatic *N*-nitroso compounds (Bartsch et al., 1976). It is also possible that some metabolites may bind to components within the agar. Such compounds can be detected by using a preincubation method first described by Yahagi et al. (1975) in which the bacteria, test compound, and S9 mix are incubated together in a small volume at 37°C for a short period (30–60 min) before adding the soft agar and pouring as for the standard assay. In this variation of the test, during the preincubation step, the test compound, S9 mix, and bacteria are incubated in liquid at higher concentrations than in the standard test, and this may account for the increased sensitivity with relevant mutagens. In the standard method the soluble enzymes in the S9 mix, cofactors, and the test agent may diffuse into the bottom agar. This can interfere with the detection of some mutagens—a problem that is overcome in the preincubation method (Forster et al., 1980; Gatehouse and Wedd, 1984).

The test is carried out as follows:

1. The strains are cultured overnight, and the inocula and S9 mix are prepared as in the standard Ames test.
2. The soft agar overlays are prepared and maintained at 45°C prior to use.
3. To each of 3–5 tubes maintained at 37°C in a Dri-Block are added 0.5 mL of S9 mix, 0.1 mL of the tester strain (10–18 h culture), and a suitable volume of the test compound to yield the desired range of concentrations. The S9 mix is kept on ice prior to use.
4. The reaction mixtures are incubated for use to 1 h at 37°C.
5. 2.0 mL of soft agar is added to each tube. After mixing, the agar and reaction mixture are poured onto previously labeled, dried Vogel-Bonner plates.
6. Once the agar has set, the plates are incubated for 2–3 days before revertant colonies are scored.

The use of controls is as described for the plate incorporation assay. It is crucial to use the minimum amount of organic solvent in this assay, as the total volume of the incubation mixture is small relative to the solvent component.

This procedure can be modified to provide optimum conditions for particular chemical classes. For instance, preincubation times greater than 60 min plus aeration have been found necessary in the detection of allyl compounds (Neudecker and Henschler, 1985).

9.3.7.5 *E. coli* Tester Strains Ames and colleagues have made an impressive contribution to mutagenicity testing by the development of the *Salmonella*/microsome test and, in particular, its application in the study of environmental mutagens. In genetic terms, *Salmonella* strains are, in some ways, not the best choice (see, e.g., Venitt and Crofton-Sleigh, 1981). Unlike the *Salmonella* strains, *E. coli* B strains

such as the WP2 series developed by Bridges, Green, and colleagues (Bridges, 1972; Green and Muriel, 1976) inherently possess the *umuDC*⁺ genes involved in generating mutations; they are also partly rough and thus allow many large molecules to enter the cell.

In addition to being effective general strains for mutagen detection, studies by Wilcox et al. (1990) have shown that a combination of *E. coli* WP2 *trp E* (pKM101), which has a functioning excision repair system for the detection of cross-linking agents, and *E. coli* WP2 *trp E uvrA* (pKM101) can be used as alternatives to *Salmonella* TA102 for the detection of oxidative mutagens. The *E. coli* strains have the advantage of lower spontaneous mutation rate and are somewhat less difficult to use and maintain. The *Salmonella* strains are, however, more commonly employed.

9.3.7.6 Storage and Checking of Tester Strains Detailed instructions for maintenance and confirmation of the phenotypes of the various tester strains are given in Maron and Ames (1983) and Gatehouse et al. (1990). Permanent master cultures of tester strains should be stored in liquid nitrogen or in dry ice. Such cultures are prepared from fresh nutrient broth cultures, to which DMSO is added as a cryopreservative. These cultures are checked for the various characteristics before storage as described in the following text. Cultures for use in individual experiments should be set up by inoculation from the master culture or from a plate made directly from the master culture, not by passage from a previously used culture. Passage in this way will inevitably increase the number of preexisting mutants, leading to unacceptably high spontaneous mutation rates (Gatehouse et al., 1990).

The following characteristics of the tester strains should be confirmed at monthly intervals or if the internal controls of a particular experiment fail to meet the required limits:

- Amino acid requirement
- Sensitivity to the lethal effects of the high molecular weight dye crystal violet for those strains carrying the *rfaE* mutation.
- Increased sensitivity to UV irradiation for those strains carrying the *uvrA* or *uvrB* mutations.
- Resistance to ampicillin for strains carrying pKM101 and resistance to tetracycline for strains carrying pAQ1.
- Sensitivity to diagnostic mutagens. This can be measured very satisfactorily by testing pairs of strains—one giving a strongly positive response and the partner a weak response.

The importance of these checks together with careful experiment-to-experiment controls of spontaneous mutation rates and response to reference mutation rates and response to reference mutagens cannot be overstressed; failure to apply them can result in much wasted effort.

9.3.8 Plate Incorporation Assay

9.3.8.1 Protocol for Dose Ranging and Selection Before carrying out the main tests, it is necessary to carry out a preliminary toxicity dose-ranging test. This should be carried out following the same basic protocol as the mutation test, except that instead of scoring the number of mutants on, for example, minimal media plates with limiting amounts of a required amino acid, the number of survivors is scored on fully supplemented minimal media. A typical protocol is outlined in the following:

1. Prepare a stock solution of the test compound at a concentration of 50 mg mL⁻¹ in an appropriate solvent. It may be necessary to prepare a lower concentration of stock solution, depending on the solubility of the test compound.
2. Make dilutions of the stock solution.
3. To 2.0 mL aliquots of soft agar overlay medium (0.6% agar and 0.5% sodium chloride in distilled water) containing a trace of histidine and excess biotin and maintained at 45°C in a dry block, add 100 µL of a solution of the test article. Use only one plate per dilution.
4. Mix and pour onto dried Vogel–Bonner minimal medium plates as in an Ames test, including an untreated control and a solvent control, if necessary. The final concentrations of test compound will be 5000, 1500, 500, 150, and 50 µg plate⁻¹.
5. Repeat step (3), using 0.5 mL of 8% S9 mix per 2.0 mL aliquot of soft agar in addition to the test compound and tester strain. The S9 mix is kept on ice during the experiment.
6. Incubate the plates for 2 days at 37°C and examine the background lawn of growth with a microscope (8 eye-piece lens, 10 objective lens). The lowest concentration giving a depleted background lawn is regarded as a toxic dose.

This test will also demonstrate excess growth, which may indicate the presence of histidine or tryptophan or their precursors in the test material, which could make testing for mutagenicity impracticable by this method.

When setting the maximum test concentration, it is important to test into the mg plate⁻¹ range where possible (Gatehouse et al., 1990), as some mutagens are only detectable when tested at high concentrations. However, for non-toxic, soluble mutagens an upper limit of 5 mg plate⁻¹ is recommended (DeSerres and Shelby, 1979). For less soluble compounds at least one dose exhibiting precipitation should be included.

Complex samples/mixtures can be assayed for component mutagens by using their longer chromatography plates as a basis for Ames assay testing (Bjørseth et al., 1982). Seifried et al. (2006) have compiled and published a summary of

comparative results from the Ames and mouse lymphoma mutagenicity test systems, with the Ames giving a better showing.

9.3.8.2 Forward Mutation Tests Forward mutation is an end point that may arise from various events, including base substitutions, frameshifts, DNA deletions, etc., as mentioned earlier.

9.3.9 Eukaryotic Mutation Tests

Prokaryotic systems, as described, have proved to be quick, versatile, and in many cases surprisingly accurate in identifying potential genetic hazards to man. However, there are intrinsic differences between eukaryotic and prokaryotic cells in the organization of the genome and the processing of the genetic information. Thus, there is a place for test systems based on mammalian cells for fundamental studies to understand the mutation process in higher cells and for the use of such tests for screening for genotoxic effects.

The early work of Muller showed the usefulness of the fruit fly *D. melanogaster* as a higher system for measuring germline mutations in a whole animal. The *Drosophila* sex-linked recessive lethal test has yielded much useful information and in the 1970s was a popular system for screening chemicals for mutation, but this test failed to perform well in international collaborative trials to study the utility of such tests to detect carcinogens and popularity waned. Another *Drosophila* test devised in the 1980s, the somatic mutation and recombination test (SMART), shows much promise and has revived the popularity of *Drosophila* for screening for genotoxic agent.

There are a number of test systems that use cultured mammalian cells, from both established and primary lines, that now have a large database of tested chemicals in the literature, that are relatively rapid, and that are feasible to use for genetic toxicity screening. These are discussed in the next section.

9.3.10 In Vitro Tests for the Detection of Mammalian Mutation

There have been a variety of *in vitro* mutation systems described in the literature, but only a small number have been defined adequately for quantitative studies (Cole et al., 1990). These are based on the detection of forward mutations in a similar manner to the systems described earlier for bacteria. A defined large number of cells are treated with the test agent and then, after a set interval, exposed to a selective toxic agent, so that only cells that have mutated can survive. As cultured mammalian cells are diploid (or near diploid), normally there are two copies of each gene. Recessive mutations can be missed if a normal copy is present on the homologous chromosome. As mutation frequencies for individual genes are normally very low, an impossibly large population

of cells would need to be screened to detect cells in which both copies are inactivated by mutation. This problem is overcome by measuring mutation in genes on the X chromosome in male cells where only one copy of the gene will be present or using heterozygous genes where two copies of a gene may be present, but one copy is already inactive through mutation or deletion.

Many genes are essential for the survival of the cell in culture, and thus mutations in such genes would be difficult to detect. However, use has been made of genes that are not essential for cell survival but allow the cell to salvage nucleotides from the surrounding medium. This saves the cell energy, as it does not have to make these compounds from simpler precursors by energy-expensive catabolism. These enzymes are located at the cell membrane. If the cell is supplied with toxic nucleotides, the "normal" unmutated cells will transport these into the cell and kill the cell. However, if the cells have lost the enzyme as a result of mutation (or chromosomal deletion, rearrangement, etc.), then they will not be able to "salvage" the exogenous toxic nucleotides and will survive. The surviving mutant cells can be detected by the formation of colonies on tissue culture plates or, in some cases, in the wells of microtiter plates.

One factor to take into account with these tests is that of expression time. Although a gene may be inactivated by mutation, the mRNA existing before the mutational event may decay only slowly, so that active enzyme may be present for some time after exposure to the mutagen. Thus, the cells have to be left for a period before challenging with the toxic nucleotide: this is the expression time and differs between systems.

9.3.10.1 Chinese Hamster Lines Chinese hamster cell lines have given much valuable data over the past 15 years, but their use for screening is limited by lack of sensitivity, as only a relatively small target cell population can be used, owing to metabolic cooperation (see Cole et al., 1990); however, they are still in use, so a brief description follows.

Chinese hamster CHO and V79 lines have high plating efficiencies and short generation times (<24h). These properties make the lines useful for mutagenicity experiments. Both cell lines have grossly rearranged chromosomal complements, which has an unknown effect on their responsiveness to mutagens (Tweats and Gatehouse, 1988). There is some evidence that Chinese hamster lines are undergoing genetic drift in different culture collections (Kirkland and Garner, 1987).

9.3.10.2 V79 System The Chinese hamster V79 line was established in 1958 (Ford and Yerganian, 1958). Publication of the use of the line for mutation studies (by measuring resistance to purine analogs due to mutation of the *Hgp^rt* locus) occurred 10 years later (Chu and Malling, 1968). The V79 line was derived from a male Chinese hamster; hence, V79 cells possess only a single X chromosome.

V79 cells grow as a cell sheet or monolayer on glass or plastic surfaces. If large numbers of cells are treated with a mutagen, when plated out, cells in close contact can link via intracellular bridges. These allow the transfer of cellular components between cells such as mRNA. Thus, if a cell carries a mutation in the *hprt* gene resulting in the inactivation of the relevant mRNA, it can receive viable mRNA or intact enzyme from adjacent nonmutated cells. Therefore, when the mutated cell is challenged with a toxic purine, it is lost, owing to the presence of active enzyme derived from the imported mRNA. This phenomenon is termed “metabolic cooperation” and severely limited the sensitivity of lines such as V79 for mutagen detection. This drawback can be overcome to an extent by carrying out the detection of mutant clones in semisolid agar (see, e.g., Oberly et al., 1987) or by using the “respreading technique” (Fox, 1981).

The preferred expression time for *Hprt* mutants is 6–8 days, although care needs to be taken when testing chemicals well into the toxic range, where the “expression time” needs to be extended to allow recovery.

9.3.10.3 Preliminary Cytotoxicity Testing An essential first step is to carry out a preliminary study to evaluate the toxicity of the test material to the indicator cells, under the conditions of the main mutagenicity test. When selecting dose levels, the solubility of the test compound, the resulting pH of the media, and the osmolality of the test solutions all need to be considered. The latter two parameters have been known to induce false-positive effects in *in vitro* mammalian tests (Brusick, 1986). The experimental procedure is carried out as follows:

1. Seek T75 plastic tissue culture flasks with a minimum of 2.5×10^6 cells in 120 mL of Eagle's medium containing 20 mM L-glutamine, 0.88 g L⁻¹ sodium bicarbonate, 20 mM HEPES, 50 µg mL⁻¹ streptomycin sulfate, 50 IU mL⁻¹ benzylpenicillin, and 7.5% of fetal bovine serum. The flasks are incubated for 18–24 h at 37°C in a CO₂ incubator to establish monolayer cultures.
2. Prepare treatment medium containing various concentrations of test compound—for example, 19.7 mL of Eagle's medium (without serum) plus 300 µL of stock concentration of compound in a preferred solvent (e.g., water, ethanol, DMSO, etc.). The final concentration of solvent other than water should not exceed 1% v/v. Normally a range of 0–5000 µg mL⁻¹ (final concentration) is covered. For a sparingly soluble compound, the highest concentration will be the lowest at which visible precipitation occurs. Similarly, if a compound has a marked effect on osmolality, concentrations should not be used that exceed 500 milliosmoles kilogram⁻¹ (mOsm kg⁻¹). In addition, a pH range of 6.5–7.5 should be maintained.

3. Each cell monolayer is rinsed with a minimum of 20 mL phosphate-buffered saline (PBS), and then 20 mL of treatment medium is carefully added. The flasks are incubated for 3 h at 37°C in a CO₂ incubator.
4. After treatment, carefully discard the medium from each flask and wash each monolayer twice with PBS. Care needs to be taken safely to dispose off contaminated solutions.
5. 10 mL of trypsin solution (0.025% trypsin in PBS) is added to each flask. Once the cells have rounded up, the trypsin is neutralized by the addition of 10 mL of complete medium. A cell suspension is obtained by vigorous pipetting to break up cell clumps.
6. The trypsinized cell suspension is counted and diluted in complete media before assessing for survival. For each treatment set up five petri dishes containing 200 cells per dish.
7. Incubate at 37°C in a CO₂ incubator for 7–10 days.
8. The medium is removed and the colonies are fixed and stained, using 5% Giemsa in buffered formalin. Once the colonies are stained, the Giemsa is removed and the colonies are counted.

The method can be repeated including 20% v/v S9 mix.

To calculate percentage survival, the following formula is used:

$$\frac{\text{Cell titre in treated culture}}{\text{Cell titre in control culture}} \times \frac{\text{mean no. of colonies on treated plates}}{\text{mean no. of colonies on control plates}} \times 100$$

The cloning efficiency (CE) of the control culture is calculated as follows:

$$\text{CE} = \frac{\text{mean no. of colonies per plate}}{\text{no. of cells per plate (i.e., 200)}} \times 100$$

In the absence of precipitation or effects on pH or osmolality, the maximum concentration of the main mutagenicity study is a concentration that reduces survival to approximately 20% of the control value.

Procedure for the Chinese Hamster V79/Hprt Assay The assay usually comprises three test concentrations, each in duplicate, and four vehicle control replicates. Suitable positive controls are ethyl methanesulfonate (–S9) and dimethylbenzanthracene (+S9). V79 cells with a low nominal passage number should be used from frozen stocks to help minimize genetic drift. The procedure described includes a reseeding step for mutation expression.

Steps 1–5 are the same as the cytotoxicity assay. As before, tests can be carried out in the presence and in the absence of S9 mix:

1. The trypsinized cultures are counted and a sample is assessed for survival as for the cytotoxicity assay. In addition, an appropriate number of cells are reseeded for estimation of mutation frequency at the day 8 expression time. The cells are transferred to roller bottles (usually 490 cm²) for this stage. The bottles are gassed with pure CO₂, the tops are tightened, and the bottles are incubated at 37°C on a roller machine (approximate speed 0.5–1.0 rev min⁻¹). Usually 10⁶ viable cells are reseeded in 50 mL of Eagle's medium containing serum, but more cells are required at the toxic dose levels.
2. The bottles are subcultured as necessary throughout the expression period to maintain subconfluency. This involves retrypsinization and determining the cell titer for each treatment. For each culture a fresh roller bottle is reseeded with a minimum of 10⁶ cells.
3. On day 8, each culture is again trypsinized, counted, and diluted so that a sample cell population can be assessed for cloning efficiency and a second sample can be assessed for the induction of 6TG-resistant cells.
4. The cell suspension is diluted in complete medium and 2 × 10⁵ cells added per petri dish (10 petri dishes per treatment). 6-Thioguanine is added to the medium at a final concentration of 10 µg mL⁻¹.
5. The petri dishes are incubated for 7–10 days and the medium is then removed. The colonies are fixed and stained as previously. The colonies (>50 cells per clone) are then counted.

Mutation frequency in each culture is calculated as

$$\frac{\text{Mean no. colonies on thioguanine plates}}{1000 \times \text{mean no. colonies on survival plates}}$$

9.3.10.4 Data Analysis A weighted analysis of variance is performed on the mutation frequencies, as the variation in the number of mutations per plate usually increases as the mean increases. Each dose of test compound is compared with the corresponding vehicle control by means of a one-sided Dunnett's test, and, in addition, the mutation frequencies are examined to see whether there is a linear relationship with dose (Arlett et al., 1989).

The criterion employed for a positive response in this assay is a reproducible statistically significant increase in mutation frequency (weighted mean for duplicate treated cultures) over the concurrent vehicle control value (weighted mean for four independent control cultures). Ideally, the response should show evidence of a dose–response relationship.

When a small isolated significant increase in mutation frequency is observed in only one of the two duplicate experiments, then a third test should be carried out. If the third test shows no significant effects, the initial increase is likely to be a chance result. In cases where an apparent treated-related increase is thought to be a result of unusually low variability or a low control frequency, comparison with the laboratory historical control frequency may be justified.

9.3.10.5 Chinese Hamster Ovary (CHO)/Hypoxanthine-Guanine Phosphoribosyltransferase (HGPRT) Assay

Chinese hamster ovary (CHO) cells have 21 or 22 chromosomes with one intact X chromosome and a large acrocentric marker chromosome (Natarajan and Obe, 1982). The use of these cells in mammalian mutation experiments was first reported by Hsie et al. (1975) and was refined into a quantitative assay for mutagenicity testing by O'Neill. The target gene is at the HGPRT locus on the X chromosome. The HGPRT enzyme is critical for purine salvaging and will incorporate toxic analogs into DNA. A forward mutation yields cells with an inactive gene product that allows cell survival and reproduction. The performance of this system has been reviewed by the US EPA Gene-Tox Program. The experimental procedure for this assay is similar to the V79/HgpRT system already described, and for more detailed descriptions the reader is referred to O'Neill et al. (1977) and Li et al. (1987). Statistical analysis of results requires some modified methodology (Snee and Irr, 1981).

9.3.10.6 Mouse Lymphoma L5178Y TK^{+/−} Assay Whereas the Chinese hamster cell systems are regarded as relatively insensitive, the mouse lymphoma L5178Y TK^{+/−} test is undoubtedly more sensitive. Unfortunately, there are persistent doubts regarding its specificity—that is, the ability to distinguish between mutagens and nonmutagens (see, e.g., Tennant et al., 1987). However, a great advantage is the ability of these cells to grow in suspension culture in which intracellular bridges do not occur. Thus, the problems of metabolic cooperation are avoided, which allows a large number of cells to be treated for optimum statistical analysis of results.

A candid historical overview of the development of the mouse lymphoma TK^{+/−} mutagenicity assay is given by its originator, Clive (1987). Initially methodologies were developed for producing the three TK genotypes (TK^{+/+} and TK^{−/−} homozygotes and the TK^{+/−} heterozygotes) (Clive et al., 1972, 1979, 1983, 1995). This first heterozygote was lost; however, it was recognized that subsequent heterozygotes produced distinctly bimodal distributions of mutant-colony sizes, owing to differences in growth rate. These were interpreted in terms of single-gene (large-colony mutants) and viable chromosomal mutations (small-colony mutants). A period of diversification of the MLA followed with controversy over the significance of small-colony mutants (Amacher et al., 1980; Honma et al., 1999a, b).

Following this, a series of cytogenetic studies confirmed the cytogenetic interpretation for small-colony mutants (see, e.g., Clive and Spector, 1975; Hozier et al., 1982). Molecular studies showed that most mutations resulting in small-colony mutants involve large-scale deletions (Evans et al., 1986). A current theory states that, for many compounds, deletion mutants are induced by binding of the compound to complexes between topoisomerase II and DNA (Clive, 1989). Topoisomerases are enzymes that control supercoiling via breakage and reunion of DNA strands; it is the latter step that is disrupted, which leads to chromosomal damage and deletions. Further molecular studies (Applegate et al., 1990) have shown that a wide variety of genetic events can result in the formation of TK^{+/-} genotype from the heterozygote, including recombinations and mitotic nondisjunction.

The TK^{+/-} line was originally isolated as a spontaneously arising revertant clone from a UV-induced TK^{-/-} clone. The parental TK^{+/+} cell and the heterozygote were then the only TK-competent mouse lymphoma cells that could be maintained in THMG medium (3 µg mL⁻¹ thymidine, 5 µg mL⁻¹ hypoxanthine, 0.1 µg mL⁻¹ methotrexate, and 7.5 µg mL⁻¹ glycine) (Clive, 1987). Thus, like most established lines, these cells are remote from wild-type cells. The karyotype of the TK^{+/-} -3.7.2C line has a modal chromosome number of 40 like the wild type but has a variety of chromosomal rearrangements and centromeric heteromorphisms (Blazak et al., 1986).

Two main protocols have been devised for carrying out mutation assays with mouse lymphoma L5178Y cells—that is, plating the cells in soft agar or a fluctuation test approach. It is the latter that is described in the following section, based on Cole et al. (1986). The reader is referred to Clive et al. (1987) for a full description of the soft agar method. Majeska and Holden (1994) have proposed a modification in test design to reduce test compound requirements.

Preliminary Cytotoxicity Assay The cells are maintained in RPMI 1640 medium containing 2.0 mM glutamine, 20 mM HEPES, 200 µg mL⁻¹ sodium pyruvate, 50 IU mL⁻¹ benzylpenicillin, 50 µg mL⁻¹ streptomycin sulfate, and 10% donor horse serum (heat inactivated for 30 min at 56°C). This medium is designated CM10. Conditioned medium is CM10 in which cells have grown exponentially for at least 1 day. Treatment medium contains 3% horse serum and 30% conditioned media (CM3). Medium without serum is known as incomplete medium (ICM). If treatment time exceeds 3 h, treatment is carried out in CM10. In 2006 it was recommended that treatment be for 24 h (Moore et al., 2000, 2006, 2007).

The method is as follows:

1. The cell titer of an exponentially growing culture of cells in CM10 is determined with a Coulter counter. The cell suspension is centrifuged at 70g for 5 min, and the supernatant is reduced such that 3 mL contains approximately 5×10^6 cells (3 h treatment) or 2×10^6 (treatment >3 h).

- 2a. For tests in the absence of S9 mix, treatment groups are prepared by mixing 3 mL of solution of test compound and 6.9 mL of ICM (3 h treatment) or 6.9 mL of CM10 (treatment >3 h).
- 2b. Tests in the presence of S9 mix are carried out in the same way, except that the treatment medium contains 10% v/v S9 mix at the expense of ICM—that is, 3 mL cell suspension, 5.9 mL ICM, 1 mL S9 mix, and 0.1 mL test compound solution/vehicle. The composition of the S9 mix is as described earlier. It is prepared immediately before required and kept on ice until it is added to the test system. For the vehicle controls, if an organic solvent is used, it should not exceed 1% v/v.
3. After the treatment period, cells are spun down at 70g for 5 min, and the supernatant is transferred for assessment of pH and osmolality. The cell pellet is washed twice in PBS and then resuspended in 10 mL CM10. (All contaminated material and waste should be disposed of safely.)
4. The cell titer of each culture is counted and a sample diluted in CM10 for assessment of posttreatment survival. For this two 96-well microtiter plates are charged with 200 µL of a diluted cell suspension, using a multichannel pipette such that each well contains on average one cell.
5. Plates are incubated for 7–8 days at 37°C and 5% CO₂ in 95 ± 3% relative humidity.
6. The plates are removed from the incubator, and 20 µL of MTT [3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyl-tetrazolium bromide] at 5 mg mL⁻¹ (in PBS) is added to each well with a multichannel pipette. The plates are left to stand for 1–4 h and are then scored for the presence of colonies with a Titertek mirror box, which allows direct viewing of the bottom surface of the plates.
7. Cytotoxicity can also be determined posttreatment as follows: T25 flasks are set up after treatment containing 0.75×10^5 cells mL⁻¹ in 5 mL CM10. Flasks are incubated with loose lids at 37°C with 5% CO₂ in 95 ± 3% relative humidity. Two days later the cell titer of each culture is determined with a Coulter counter.
8. Following this procedure, various calculations are carried out to aid selection of dose levels for the main mutation assay.
 - a. *Cloning efficiency* In microtiter assays calculations are based on the Poisson distribution

$$P(o) = \frac{\text{no. of wells without a colony}}{\text{total no. of cells}}$$

- b. *Relative survival* Relative survival (*S*) is calculated as follows:

$$S = \frac{\text{CE of treated group}}{\text{CE of control group}}$$

- c. *Growth* Growth in suspension (SG) is calculated as follows:

$$SG = \frac{\text{cell count after 3 days}}{0.75 \times 10^5}$$

Relative suspension growth (RSG) is calculated as follows:

$$RSG = \frac{\text{SG of treated group}}{\text{SG of control group}} \times 100\%$$

9.3.10.7 Selection of Dose Levels The highest test concentration is selected from one of the following options, whichever is lowest:

- A concentration which reduces survival to about 10–20% of the control value.
- A concentration which reduces RSG to 10–20% of the control value.
- The lowest concentration at which visible precipitation occurs.
- The highest concentration which does not increase the osmolality of the medium to greater than 400 mmol kg⁻¹ or 100 mmol above the value for the solvent control.
- The highest concentration that does not alter the pH of the treatment medium beyond the range 6.8–7.5.
- If none of these conditions are met, 5 mg mL⁻¹ should be used.

Lower test concentrations are selected as fractions of the highest concentration, usually including one dose which causes 20–70% survival and one dose which causes >70% survival.

9.3.10.8 Main Mutation Assay The assay normally comprises three test concentrations, a positive control and vehicle control. All treatment groups are set up in duplicate. The expression time is 2 days, unless there are indications that the test agent inhibits cell proliferation, where an additional or possibly alternative expression time should be employed.

Stock cultures are established from frozen ampoules of cells that have been treated with thymidine, hypoxanthine, methotrexate, and glycine for 24 h, which purge the culture of preexisting TK^{-/-} mutants. This cell stock is used for a maximum of 2 months.

Treatment is normally carried out in 50 mL centrifuge tubes on a roller machine. During the expression time the cells are grown in T75 plastic tissue culture flasks. For estimation of cloning efficiency and mutant induction, cells are plated out in 96-well microtiter plates. Flasks and microtiter plates are incubated at 37°C in a CO₂ incubator as in the cytotoxicity assays.

Cell titers are determined by diluting of the cell suspension in Isoton and counting an appropriate volume (usually 0.5 mL) with a Coulter counter. Two counts are made per suspension.

The experimental procedure is carried out as follows:

1. On the day of treatment, stock solutions for the positive control and the various concentrations of test compound (selected as per the previous selection) are prepared.
2. Treatment is carried out in 30% conditioned media. The serum concentration is 3% (3 h treatment) or 10% (treated >3 h).
3. Cell suspensions of exponentially growing cells are prepared as in the cytotoxicity assay, except that 6 mL of media required for each treatment culture contains 10⁷ cells (3 h treatment) or 3 × 10⁶ cells (>3 h treatment). The number of cells per treatment may be increased if marked cytotoxicity is expected, to allow enough cells to survive (e.g., if 20% survival or less is expected, 2 × 10⁷ cells may be treated).
4. For tests in the absence of S9 mix, 6 mL of cell suspension, 0.2 mL test compound/vehicle, and 13.8 mL ICM (3 h treatment) or 13.8 mL CM10 (treatment >7 h) are mixed in the presence of S9 mix, and 0.2 mL of test compound/vehicle is prepared.
5. After treatment the cells are centrifuged at 70g for 5 min, and supernatant is discarded and the cell pellet is resuspended in PBS (pH 7). This washing procedure is repeated twice, and finally the cell pellet is resuspended in CM10.
6. Each culture is counted so that a sample of cells can be assessed for posttreatment survival and the remaining cell population assessed for estimation of mutation frequency.
7. For survival estimation, cells are placed into 96-well microtiter trays at a cell density of one cell per well as per the cytotoxicity assay.
8. For mutation estimation, the cells are diluted to a cell density of 2 × 10⁵ cells mL⁻¹ with CM10 in tissue culture flasks, and the culture is incubated at 37°C in a CO₂ incubator. On day 1 each culture is counted and diluted with fresh medium to a cell density of 2 × 10⁵ cells mL⁻¹ in a maximum of 100 mL of medium.
9. On day 2 each culture is counted again and an aliquot of cells taken so that (i) a sample of the cell population can be assessed for cloning efficiency. Plates are incubated at 37°C in a CO₂ incubator for 7 days. (ii) A sample of the cell population can be assessed for the induction of TFT-resistant cells (mutants). For this 2 × 10³ cells are plated per well in 200 µL CM10 containing 4 µg mL⁻¹ TFT. TFT and TFT-containing cultures must not be exposed to bright light, as the material is light sensitive. The plates are incubated for 10–12 days at 37°C in a CO₂ incubator.

10. At the end of incubation, 20 μ L MTT is added to each well. The plates are left to develop for 1–4 h at 37°C and then scored for colony-bearing wells. Colonies are scored by eye and are classified as small or large.

The calculation for cloning efficiency is made as for the cytotoxicity assay.

Relative total growth (RTG) is a cytotoxicity parameter which considers growth in suspension during the expression time and the cloning efficiency of the end of the expression time as follows:

$$\text{Suspension growth (SG)} = \frac{24 \text{ h cell count}}{2 \times 10^4} \times \frac{48 \text{ h cell count}}{2 \times 10^5}$$

$$\text{RTG} = \frac{\text{SG treated culture}}{\text{SG control culture}} \times \frac{\text{CE of treated culture}}{\text{CE of control culture}}$$

Mutation frequency (MF) is calculated as follows:

$$\text{MF} = \frac{\text{in } P_0 \text{ for mutation plates}}{\text{no. of cells per well} \times \text{CE} / 100}$$

9.3.10.9 In Vivo Genotoxicity Tests for the Assessment of Primary DNA Lesions Primary DNA lesions are detected with so-called indicator tests. These tests do not directly measure consequences of DNA interaction (i.e., mutation) but do detect effects related to the process of mutagenesis, such as DNA damage, recombination, and repair. Results from indicator tests can provide additional useful information in the context of extended genotoxicity testing. However, primary DNA lesions may be repaired error-free and do not necessarily result in formation of mutations. The most commonly utilized assays in pharmaceutical development are the P-postlabeling assay and the comet assay. A comparison of different aspects of the methods described in the text is depicted in Table 9.7. Basic aspects regarding optimal study design for *in vivo* micronucleus assays are largely applicable to the design of supplemental *in vivo* assays. Specific or unique aspects on study protocols are described more extensively where appropriate.

9.3.10.10 The Comet Assay The *in vivo* comet assay (aka single-cell gel electrophoresis assay) is increasingly being used as a supplement genotoxicity test for drug candidates (Hartmann et al., 2001; Brendler-Schwaab et al., 2005) due to its simplicity, speed, and wide acceptance. The assay was first developed by Östling and Johansson (1984). There are general review articles on the comet assay (Tice et al., 2000; Speit and Hartmann, 2005), and a general guideline for test conductance has been published as a result of the International Workshop on Genotoxicity Test Procedures (IWGTP).

More specific recommendations with the goal of gaining more formal regulatory acceptance of the comet assay were published following the 4th International Comet Assay Workshop (Burlinson et al., 2007). An updated position paper on specific aspects of tests conditions and data interpretation was prepared following the IWGT in 2005 (Burlinson et al., 2007) (see Table 9.3).

9.3.10.11 Principle of Method The basic principle of the comet assay is the migration of DNA in an agarose matrix under electrophoretic conditions. When viewed through the microscope, a cell has the appearance of a comet, with a head (the nuclear region) and a tail containing DNA fragments or strands migrating in the direction of the anode. Among the various versions of the comet assay, the alkaline (pH of the unwinding and electrophoresis buffer ≥ 13) method enables detection of the broadest spectrum of DNA damage and is therefore urgently recommended (in the first instance) for regulatory purposes (Tice et al., 2000). The alkaline version detects DNA damage such as strand breaks, alkali-labile sites (ALS), and single-strand breaks associated with incomplete excision repair. Under certain conditions, the assay can also detect DNA–DNA and DNA–protein cross-linking, which (in the absence of other kinds of DNA lesions) appears as a relative decrease in DNA migration compared to concurrent controls. In contrast to other DNA alterations, cross-links may stabilize chromosomal DNA and inhibit DNA migration (Merk and Speit, 1998). Thus, reduced DNA migration in comparison to the negative control (which should show some degree of DNA migration) may indicate the induction of cross-links, which are relevant

TABLE 9.7 Alternative Test under ICH

| Aspect | Comet Assay | DNA Adducts | UDS Test (Liver) | Transgenic Gene Mutation |
|-------------------------------------|--------------|---------------|------------------|--------------------------|
| Test definition (accepted protocol) | Yes | No | Yes | Yes |
| Regulatory acceptance/use | Yes | Yes | Yes | Yes |
| Relevance of end point | Moderate | Moderate | Moderate | High |
| Technical demands | Low moderate | Moderate high | Moderate | High |
| Widespread use | Yes | No | Yes | No |
| Applicable to most tissues | Yes | Yes | No | Yes |
| Dependence of cell turnover | No | No | No | Yes |
| Cost | Low moderate | Moderate high | Low | High |

lesions with regard to mutagenesis and should be further investigated. Increased DNA migration indicated the induction of DNA strand breaks and/or ALS. Furthermore, enhanced activity of excision repair may result in increased DNA migration. DNA excision repair can influence comet assay effects in a complex way (Speit and Hartmann, 1995). While DNA repair generally reduces DNA migration by eliminating DNA lesions, ongoing excision repair may increase DNA migration to incision-related DNA strand breaks. Thus, the contribution of excision repair to the DNA effects seen in the comet assay depends on the types of induced primary DNA damage and the time point of analysis (Collins et al., 1993).

Test procedure aspects regarding test animals, test substance, use of concurrent negative and positive control animals, as well as dose selection for the design of a cytogenetic assay, as described in detail previously, are largely applicable to the design of an *in vivo* comet assay. In addition, more specific details can be found in an earlier publication. A single treatment or repeated treatments (generally at 24 h intervals) are equally acceptable. In both experimental designs, the study is acceptable as long as a positive effect has been demonstrated or, for a negative result, as long as an appropriate level of animal or tissue toxicity has been demonstrated or the limit dose with appropriate level tissue exposure has been used. For repeated treatment schedules, dosing must be continued until the day of sampling. On a daily basis, test substances may be administered as a split dose (i.e., two treatments separated by no more than a few hours), to facilitate administering a large volume of material. The test may be performed in two ways. If animals are treated with the test substances once, then tissue/organ samples are obtained at 2–6 and 16–26 h after dosing. The shorter sampling time is considered sufficient to detect rapidly absorbed as well as unstable or direct-acting compounds. In contrast, the late sampling time is intended to detect compounds that are more slowly absorbed, disturbed, and metabolized. When a positive response is identified at one sampling time, data from the other sample time need not be collected. Alternatively, if multiple treatments at 24 h intervals are used, tissue/organ samples need to be collected only once. The sampling time should be 2–6 h after the last administration of the test substance. Alternative sampling times may be used when justified on the basis of toxicokinetic data.

Selection of Tissues and Cell Preparation In principle, any tissue of the experimental animal, provided that a high-quality single-cell/nucleus suspension can be obtained, can be used for a comet assay. Selection of the tissue(s) to be evaluated should be based, wherever possible, on data from absorption, distribution, metabolism, and excretion studies and/or other toxicological information. A tissue should not be evaluated unless there is evidence of, or support for, exposure of the tissue to the test substance and/or its metabolite(s). In the absence of such information and unless

scientifically justified, two tissues should be examined. Recommended tissues are liver, which is the major organ for the metabolism of absorbed compounds, and a site of first contact tissue—for example, gastrointestinal for orally administered substances, respiratory tract for substances administered via inhalation, or skin for dermally applied substances. Which tissue is evaluated first is at the discretion of the investigator, and both tissues need not be evaluated if a positive response is obtained in the first tissue evaluated.

Single-cell suspension can be obtained from solid tissue by mincing briefly with a pair of fine scissors (Tice et al., 1991), by incubation with digestive enzymes such as collagenase trypsin (Brendler-Schwaab et al., 1994), or by pushing the tissue sample through a mesh membrane. Cell nuclei can also be obtained by homogenization. During mincing or homogenization, EDTA can be added to the processing solution to chelate calcium/magnesium and prevent endonuclease activation. In addition, radical scavengers (e.g., DMSO) can be added to prevent oxidant-induced DNA damage. Any cell dissociation method is acceptable as long as it can be demonstrated that the process is not associated with inappropriate background levels of DNA damage.

Cytotoxicity: A Potential Confounding Factor A general issue with DNA strand break assays such as the comet is the indirect mechanisms related to cytotoxicity which may lead to enhanced strand break formation. However, since DNA damage in the comet assay is assessed on the level of individual cells, dead or dying cells may be identified on microscopic slide by their specific image. Necrotic or apoptotic cells can result in comets with small or nonexistent head and large diffuse tails as observed *in vitro* upon treatment with cytotoxic and nongenotoxic articles. However, such microscopic images are not uniquely diagnostic for apoptosis or necrosis since they may also be detected after treatment with high doses of radiation or high concentrations of strong mutagens. For the *in vivo* comet assay, only limited data are available to establish whether cytotoxicity results in increased DNA migration in tissues of experimental animals. Despite necrosis or apoptosis in target organs of rodents such as kidneys, testes, liver, or duodenum, no elevated DNA migration was observed. However, enhanced DNA migration was seen in homogenized liver tissue of mice dosed with carbon tetrachloride when histopathological examination showed evidence of necrosis in the liver. Therefore, to avoid potential false-positive effects resulting from cytotoxicity, recommendations regarding a concurrent assessment of target organ toxicity have been made, including dye viability assays, histopathology, and a neutral diffusion assay (Tice et al., 2000; Hartmann et al., 2001).

Biological Significance of Lesions Detected DNA lesions leading to effects in the comet assay can be strand breaks which may be relevant to the formation of CAs or DNA modifications such as abasic sites (AP sites) with

relevance to the induction of gene mutations. However, primary lesions detected by the comet assay may also be correctly repaired without resulting in permanent genetic alterations. Neither the magnitude of DNA migration in the comet assay nor the shape of the comet can reveal the types of DNA damage causing the effect or other biological significance, that is, their mutagenic potential. Therefore, conclusions regarding the mutagenicity of a test compound cannot be made solely on the basis of comet assay effects. There are a few limitations of the comet assay with regard to its application and interpretation of test results. For example, short-lived primary DNA lesions such as single-strand breaks, which may undergo rapid DNA repair, could be missed when using inadequate sampling times. However, an appropriate study design including only early preparation time point (i.e., at 3–6 h) is considered sufficient to ensure that these lesions are captured, in particular at higher dose levels, where DNA repair may be significantly delayed or even overwhelmed. In any case, it should be kept in mind that a negative comet result can be considered as a strong indicator for the absence of a mutagenic potential.

Advantages The advantages of this assay for use in genotoxicity testing of drug candidates include its applicability to various tissues and/or special cell types, its sensitivity for detecting low levels of DNA damage, its requirement for small numbers of cells per sample, the general ease of test performance, the short time needed to complete a study, and its relatively low cost. The comet assay can be applied to any tissue in the given *in vivo* model, provided that a single-cell/nuclei suspension can be obtained. Therefore, the comet assay has potential advantages over other *in vivo* genotoxicity test methods, which are reliably applicable to rapidly proliferating cells only or have been validated preferentially in a single tissue only. The comet assay may detect a broader spectrum of primary DNA lesions, including single-strand breaks and oxidative base damage, which may not be detected in the UDS test because they are not repaired by nucleotide excision repair. The advantages of the comet assay over the alkaline elution test include the detection of DNA damage on a single-cell level and the requirement for only small numbers of cells per sample. In contrast, when using the alkaline elution assay, large quantities of cells are necessary for the determination of genotoxic effects, and, therefore, only a limited number of organs/tissues can be evaluated using this technique. In particular, this seems important for investigation of suspected tissue-specific genotoxic activity, which includes “site-of-contact” genotoxicity (cases of high local vs. low systemic exposure).

Limitations Experimental variability is an important issue and should be kept to a minimum to ensure reliable interpretation and comparability of the data obtained with other *in vivo* comet experiments. Experimental variability may result from shortcomings with regard to number of

doses tested, number of animals per dose, number of slides per animal, number of cells analyzed, lack of sufficient DNA migration in cells of concurrent controls, and deviation from minimum time for treatment of slides with alkaline buffer. Considering these discrepancies, the data of the comprehensive study, as well as other study reports not in agreement with the current recommendations, should be interpreted with caution. This point was highlighted recently in a position paper on the use and status of the *in vivo* comet assay in genotoxicity testing, which critically assessed published data produced under test conditions not fully in agreement with the minimal requirements for an acceptable test. For example, it was noted that positive comet assay data were published for compounds that have been assessed before to be neither genotoxic nor carcinogenic, such as food additives. Such isolated positive comet assay results should be critically evaluated in the light of current recommendations to exclude methodological shortcomings and potential artifacts. In cases where negative carcinogenicity data are already available and the *in vivo* comet assay result represents an isolated positive finding in the context of existing genotoxicity data, the biological significance of the effect seen in the comet assay should be assessed with caution.

Data Analysis Data from the fluctuation test described earlier are analyzed by an appropriate statistical method as described in Robinson et al. (1989). Data from plate assays are analyzed as described in Arlett et al. (1989) for treat and plate tests.

9.3.10.12 Status of Mammalian Mutation Tests At present the only practical assays for screening new chemical entities for mammalian mutation are the mammalian cell assays described earlier. The protocols are well defined, and mutant selection and counting procedures are simple and easily quantified. In general, the genetic end points are understood and relevant to deleterious genetic events in humans. For these reasons the assays are still regarded as valuable in safety evaluation (Li et al., 1991). It is, however, recognized that there are still unknown factors and molecular events that influence test results. This can be illustrated by the conclusions of the third UKEMS collaborative trial, which focused on tests with cultured mammalian cells. The following points were made:

- The number of cells to be cultured during expression imposes a severe limitation in the use of surface-attached cells.
- The importance of a careful determination of toxicity.
- That S9 levels may need to be varied.
- That the aromatic amine benzidine is mutagenic only at the TK locus in L5178Y TK^{+/−} cells. The most disturbing finding was that benzidine (detectable without metabolism by S9 mix) did not produce detectable

DNA adducts (as shown by ^{32}P -postlabeling) in L5178Y cells. Thus, in the mechanism for mutagenesis in L5178Y cells, benzidine remains to be elucidated (Arlett and Cole, 1990).

9.3.11 *In Vivo* Mammalian Mutation Tests

Mammalian mutation studies of chemicals in the whole animal have provided fundamental information on mutation parameters in germ cells such as dose–response, dose fractionation, sensitivity of various stages in gametogenesis, etc., just as is known for ionizing radiation (Russell, 1989). This has led to estimations of the possible impact chemical mutagens may have on heritable malformation, inborn errors of metabolism, etc. Today germ cell studies are still required when estimating the heritable damage a mutagen may inflict on exposed human populations.

The existing tests tend to be cumbersome and are not used for routine genetic toxicology screening, and thus only brief descriptions will follow. Reviews of existing data, particularly by Holden (Holden, 1982; Adler and Ashby, 1989; Glover et al., 1992), have indicated that most if not all germ cell mutagens also induce DNA damage in somatic cells, as detected by well-established assays such as the rodent micronucleus test. The converse is not true—that is, some mutagens/clastogens can induce somatic cell damage but do not induce germ cell changes, which probably reflects the special protection afforded to the germ cells, such as that provided by the blood–testis barrier. In other words, it appears that germ cell mutagens are a subset of somatic cell mutagens.

In vivo mammalian mutation tests are not restricted to germ cell tests. The mouse spot test described in the following text is, again, a test used first for studying radiation-induced mutation but has also been used for screening chemicals for *in vivo* mutagenic potential. This test has had several proponents, but compared with *in vivo* chromosomal assays is not widely used.

9.3.11.1 The Mouse Specific Locus Test The mouse somatic spot test is a type of specific locus test. The classical specific locus test was developed independently by Russell at Oak Ridge in the late 1940s (Russell, 1951, 1989) and Carter in Edinburgh (Carter et al., 1956). The test consists of treatment of parental mice homozygous for a wild-type set of marker loci. The targets for mutation are the germ cells in the gonads of the treated mice. These are mated with a tester stock that is homozygous recessive at the marker loci. The F_1 offspring that result are normally heterozygous at the marker loci and thus express the wild-type phenotype. In the event of a mutation from the wild-type allele at any of these loci, the F_1 offspring express the recessive phenotype.

The test marker strain (T) developed by Russell uses seven recessive loci, namely, *a* (nonagouti), *b* (brown), *c^{ch}*

(chinchilla), *d* (dilute), *p* (pink-eyed dilution), *s* (piebald) and *se* (short ear). As for the mouse spot test, these genes control coat pigmentation, intensity or pattern, and, for the *se* gene, the size of the external ear (Russell, 1984).

As the occurrence of mutation is rare even after mutagen treatment, the specific locus test is the ultimate study of mutation, requiring many thousands of offspring to be scored, plus significant resources of time, space, and animal husbandry. Because of these constraints it is often difficult to define a negative result, as insufficient animals are scored or all stages of spermatogenesis are not covered. Of the 25 compounds tested in the assay, as reviewed by Ehling et al. (1986), 17 were regarded as “inconclusive” and 8 positive. The scale studies can reach is illustrated by the test of ethylene oxide described by Russell et al. (1984), where exposures of 101 000 and 150 000 ppm were used over 16–23 weeks. A total of 71 387 offspring were examined. The spermatogonial stem cell mutation rate in the treated animals did not differ significantly from the historical control frequency!

With regard to the design of the test, mice are mated when 7–8 weeks old. By this age all germ cell stages are present. The test compound is normally administered by the IP route to maximize the likelihood of germ cell exposure. The preferred dose is just below the toxic level so long as fertility is not compromised. One lower dose should also be included.

In males spermatogonia are most at risk, but it is also desirable that later stages also be exposed. Thus, the mice are mated immediately after treatment to two to four females. This is continued each week for 7 weeks. Then the first group has completed its rearing of the first set of offspring and is remated. This cycle can be continued for the lifetime of the males. Tests can also be carried out by dosing females, when treatment is carried out for 3 weeks to cover all stages of oogenesis.

The offspring are examined immediately after birth for identification of malformations (dominant visibles) and then at weaning for the specific locus mutations. Presumptive mutant mice are checked by further crosses to confirm their status (Searle, 1984).

Comparison of mutation frequencies is made with the historical database. For definition of a positive result, the same principles are recommended as for the mouse spot test (Selby and Olson, 1981). A minimum size of 18 000 offspring per group is recommended by those authors for definition of a negative result.

9.4 IN VITRO CYTOGENETIC ASSAYS

The *in vitro* cytogenetic assay is a short-term mutagenicity test for detecting chromosomal damage in cultured mammalian cells (Fox, 1981).

Cultured cells have a limited ability metabolically to activate some potential clastogens. This can be overcome by adding an exogenous metabolic activation system such as S9 mix to the cells (Ames et al., 1975; Madle and Obe, 1980; Natarajan and Obe, 1982; Maron and Ames, 1983). It is important to accurately measure cytotoxicity in the assays, and a range of methods are available to measure them (Lorge et al., 2008).

Observations are made in metaphase cells arrested with a spindle inhibitor such as colchicine or colcemid to accumulate cells in a metaphase-like stage of mitosis (c-metaphase) before hypotonic treatment to enlarge cells and fixation with alcohol/acetic acid solution. Cells are then dispersed on to microscope slides and stained, and slides are randomized, coded, and analyzed for CAs with high-power light microscopy. Details of the procedure are given in Dean and Danford (1984) and Preston et al. (1981, 1987). The UKEMS guidelines (Scott et al., 1990) recommend that all tests be repeated regardless of the outcome of the first test and that, if a negative or equivocal result is obtained in the first test, the repeat should include an additional sampling time. In the earlier version of the guidelines (Scott et al., 1983), a single sampling at approximately 1.5 normal cycle times (–24 h for a 1.5 cell cycle) from the beginning of treatment was recommended, provided that a range of concentrations was used which induced marginal to substantial reductions in mitotic index (MI), usually an indicator of mitotic delay. However, Ishidate (1988a) reported a number of chemicals which gave negative responses with a fixation time of 24 h but which were positive at 48 h. This was when a Chinese hamster fibroblast line (CHO) with a doubling time of 15 h was used. It would appear, therefore, that there are chemicals which can induce extensive mitotic delay at clastogenic doses and may be clastogenic only when cells have passed through more than one cell cycle since treatment (Thust et al., 1980). A repeat test should include an additional sample at approximately 24 h later, but it may only be necessary to score cells from the highest dose at this later fixation time. When the first test gives a clearly positive result, the repeat test need only utilize the same fixation time. The use of other sampling times is in agreement with other guidelines (European Community EEC Directive—OECD, 1983; Japanese Ministry of Health and Welfare and Japanese Ministry of International Trade and Industry, 1987; American Society for Testing and Materials—Preston et al., 1987; Ishidate, 1988b; Japanese Guidelines—Japanese Ministry of Health and Welfare (JMHV), 1989).

9.4.1 Cell Types

Established cell lines, cell strains, or primary cell cultures may be used. The most often used are Chinese hamster cell lines and human peripheral blood lymphocytes. The merits of these two cell lines have been reported (Ishidate and

Harnois, 1987; Kirkland and Garner, 1987). The cell system must be validated and consistently sensitive to known clastogens. Rat lymphocytes have also been proposed for use (Sinha et al., 1989), with reports that they offer better specificity than hamster cells.

9.4.2 Chinese Hamster Cell Lines

CHO cells in which there has been an extensive rearrangement of chromosome material and the chromosome number may not be constant from cell to cell are frequently used. Polyploidy, endoreduplication, and high spontaneous CA frequencies can sometimes be found in these established cell lines, but careful cell culture techniques should minimize such effects (Hsie et al., 1981). Cells should be treated in exponential growth when cells are in all stages of the cell cycle. High concentration or low or high osmolarities can lead to false positives (Galloway et al., 1985, 1994).

9.4.3 Human Peripheral Blood Lymphocytes

Blood should be taken from healthy donors not known to be suffering from viral infections or receiving medication. Staff handling blood should be immunized against hepatitis B, and regular donors should be shown to be hepatitis B antigen negative. Donors and staff should be aware of AIDS implications, and blood and cultures should be handled at containment level 2 (Advisory Committee on Dangerous Pathogens, 1984).

Peripheral blood cultures are stimulated to divide by the addition of a T-cell mitogen such as phytohemagglutinin (PHA) to the culture medium. Mitotic activity is at a maximum at about 3 days but begins at about 40 h after PHA stimulation, and the chromosome constitution remains diploid during short-term culture (Evans and O'Riordan, 1975). Treatments should commence at about 44 h after culture initiation. This is when cells are actively proliferating and cells are in all stages of the cell cycle. They should be sampled about 20 h later. In a repeat study the second sample time should be about 92 h after culture initiation. Morimoto et al. (1983) report that the cycle time for lymphocytes averages about 12–14 h except for the first cycle. The incidence of spontaneous cytogenetic aberrations at 48 and 72 h is reported as markedly lower than in CHO cells (Sinha et al., 1984).

Female donors can give higher yields of chromosomal damage (Anderson et al., 1989).

9.4.4 Positive and Negative Controls

When the solvent is not the culture medium or water, the solvent, liver enzyme activation mixture, and solvent and untreated controls are used as negative controls.

Since cultured cells are normally treated in their usual growth medium, the solubility of the test material in the medium should be ascertained before testing. As pointed out earlier, extremes of pH can be clastogenic (Cifone et al., 1987), so the effect of the test material on pH should also be determined, but buffers can be utilized.

Various organic solvents are used, such as DMSO, dimethylformamide, ethanol, and acetone. The volume added must not be toxic to cells. Greater than 10% water (v/v) can be toxic because of nutrient dilution and osmolality changes.

A known clastogen should always be included as a positive control. When metabolic activation is used, a positive control chemical known to require metabolic activation should also be used to ensure that the system is functioning properly. Without metabolic activation, a direct-acting positive control chemical should be used. A structurally related positive control can also be used. Appropriate safety precautions must be taken in handling clastogens (IARC, 1979; MRC, 1981).

Positive control chemicals should be used to produce relatively low frequencies of aberrations so that the sensitivity of the assay for detecting weak clastogens can be established (Preston et al., 1987).

Aberration yields in negative and positive controls should be used to provide a historical database.

9.4.5 Treatment of Cells

When an exogenous activation system is employed, short treatments (about 2 h) are usually necessary because S9 mix is often cytotoxic when used for extended lengths of time. However, cells may be treated with chemicals either continuously up to harvest time or for a short time followed by washing and addition of fresh medium to allow cell-cycle progression. Continuous treatment avoids centrifugation steps required with washing of cells and optimizes the endogenous metabolic capacity of the lymphocytes.

When metabolic activation is used, S9 mix should not exceed 1–10% of the culture medium by volume. It has been shown that the S9 mix is clastogenic in CHO cells and mouse lymphoma cells (Cifone et al., 1987; Kirkland et al., 1989) but not in human lymphocytes, where blood components can inactivate active oxygen species which could cause chromosomal damage. When S9 mix from animals treated with other enzyme-inducing agents such as phenobarbitone/beta-naphthoflavone is used, clastogenesis may be minimized (Kirkland et al., 1989).

Prior to testing, it is necessary to determine the cytotoxicity of the test material, in order to select a suitable dose range for the chromosome assay both with and without metabolic activation. The range most commonly used determines the effect of the agent on the MI, that is, the percentage of cells in mitoses at the time of cell harvest. The highest dose should inhibit mitotic activity by approximately 50% (EEC Annex V) or 75% (UKEMS: Scott et al., 1990) or exhibit some

other indication of cytotoxicity. If the reduction in MI is too great, insufficient cells can be found for chromosome analysis. Cytotoxicity can also be assessed by making cell counts in the CA test when using cell lines. In the lymphocyte assay total white cell counts can be used in addition to MI. A dose which induces 50–75% toxicity in these assays should be accompanied by a suitable reduction in MI.

If the test material is not toxic, it is recommended that it be tested up to 5 mg mL⁻¹. The UKEMS recommends that chemicals be tested up to their maximum solubility in the treatment medium and not just their maximum solubility in stock solutions.

For highly soluble nontoxic agents, concentrations above 10 mM may produce substantial increases in the osmolality of the culture medium which could be clastogenic by causing ionic imbalance within the cells (Ishidate et al., 1984; Brusick, 1987a). At concentrations exceeding 10 mM, the osmolality of the treatment media should be measured, and if the increase exceeds 50 mmol kg⁻¹, clastogenicity resulting from high osmolality should be suspected and, according to the UKEMS, is unlikely to be of relevance to human risk. The UKEMS also does not recommend the testing of chemicals at concentrations exceeding their solubility limits as suspensions or precipitate.

A minimum of three doses of the test material should be used—the highest chosen as described earlier, the lowest on the borderline of toxicity and an intermediate one. Up to six doses can be managed satisfactorily, and this ensures the detection of any dose–response and that a toxic range is covered. MIs are as required for the preliminary study (at least 1000 cells per culture). It is also useful to score endoreduplication and polyploidy for historical data. Cells from only three doses need to be analyzed.

The range of doses used at the repeat fixation time can be those which induce a suitable degree of mitotic inhibition at the earlier fixation time, but if the highest dose reduces the MI to an unacceptably low level at the second sampling time, the next highest dose should be chosen for screening.

A complete assay requires the test material to be investigated at a minimum of three doses together with a positive (untreated), and solvent-only control can be omitted if tissue culture medium is used as a solvent. When two fixation times are used in repeat tests, the positive control is necessary at only one time, but the negative or solvent control is necessary at both times.

Duplicates of each test group and quadruplicates of solvent or negative controls should be set up. The sensitivity of the assay is improved with larger numbers scored in the negative controls (Richardson et al., 1989).

9.4.6 Scoring Procedures

Prior to scoring, slides should be coded, randomized, and then scored “blind.” Metaphase analysis should only be carried out by an experienced observer. Metaphase cells should

be sought under low-power magnification and those with well-spread, that is, nonoverlapping, clearly defined non-fuzzy chromosomes examined under high power with oil immersion. It is acceptable to analyze cells with total chromosome numbers or that have lost one or two chromosomes during processing. In human lymphocytes ($2n=46$) 44 or more centromeres and in CHO cells ($2n=22$; range 21–24) 20 or more centromeres can be scored. Chromosome numbers can be recorded for each cell, to give an indication of aneuploidy. Only cells with increases in numbers (above 46 in human lymphocytes and 24 in CHO cells) should be considered in this category, since decreases can occur through processing.

Recording microscope coordinates of cells is necessary and allows verification of abnormal cells. A photographic record is also useful of cells with aberrations. Two hundred cells (100 from each of two replicates) should be scored per treatment group. When ambiguous results are obtained, there may be further “blind” reading of these samples.

9.4.7 Data Recording

The classification and nomenclature of the International System for Human Cytogenetic Nomenclature (ISCN, 1985) as applied to acquired CAs is recommended. Score sheets giving the slide code, microscope scorer's name, date, cell number, number of chromosomes, and aberration types should be used. These should include chromatid and chromosome gaps, deletions, exchanges, and others. A space for the vernier reading for comments and a diagram of the aberration should be available.

From the score sheets, the frequencies of various aberrations should be calculated, and each aberration should be counted only once. To consider a break as one event and an exchange as two events is not acceptable, since unfounded assumptions are made about mechanisms involved (Revell, 1974).

9.4.8 Presentation of Results

The test material, test cells used, method of treatment, harvesting of cells, cytotoxicity assay, etc. should be clearly stated as well as the statistical methods used. Richardson et al. (1989) recommend that comparison be made between the frequencies in control cells and at each dose level using Fisher's exact test.

In cytogenetic assays the absence of a clear positive dose–response relationship at a particular time frequently arises. This is because a single common sampling time may be used for all doses of a test compound. CA yields can vary markedly with posttreatment sampling time of an asynchronous population, and increasing doses of clastogens can induce increasing degrees of mitotic delay (Scott et al., 1990). Additional fixation times should clarify the relationship between dose and aberration yield.

Gaps are by tradition excluded from quantification of CA yields. Some gaps have been shown to be real discontinuities in DNA (e.g., Heddle and Bodycote, 1970). Where CA yields are on the borderline of statistical significance above control values, the inclusion of gaps could be useful. Further details on this approach may be found in the UKEMS guidelines (Scott et al., 1990).

Since chromosome exchanges are relatively rare events, greater biological significance should be attached to their presence than to gaps and breaks.

Chemicals which are clastogenic *in vitro* at low doses are more likely to be clastogenic *in vivo* than those where clastogenicity is detected only at high concentrations (Ishidate et al., 1988). Negative results in well-conducted *in vitro* tests are a good indication of a lack of potential for *in vivo* clastogenesis, since almost all *in vivo* clastogens have given positive results *in vitro* when adequately tested (Thompson, 1986; Ishidate et al., 1988).

9.5 IN VIVO CYTOGENETIC ASSAYS

Damage induced in whole animals can be detected in *in vivo* chromosome assays in either somatic or germinal cells by examination of metaphases or the formation of micronuclei. The micronucleus test can also detect whole chromosome loss or aneuploidy in the absence of clastogenic activity and is considered comparable in sensitivity to chromosome analysis (Tsuchimoto and Matter, 1979).

Rats and mice are generally used for *in vivo* studies, with the mouse being employed for bone marrow micronucleus analysis and the rat for metaphase analysis, but both can be used for either. Mice are cheaper and easier to handle than rats, and only a qualitative difference in response has been found between the species (Albanese et al., 1988). Chinese hamsters are also widely used for metaphase analysis because of their low diploid chromosome number of 22. However, there are few other historical toxicological data for this species.

9.5.1 Somatic Cell Assays

9.5.1.1 Metaphase Analysis Metaphase analysis can be performed in any tissue with actively dividing cells, but bone marrow is the tissue most often examined. Cells are treated with a test compound and are arrested in metaphase by the administration of colcemid or colchicine at various sampling times after treatment. Preparations are examined for structural chromosomal damage. Because the bone marrow has a good blood supply, the cells should be exposed to the test compound or its metabolites in the peripheral blood supply, and the cells are sensitive to S-dependent and S-independent mutagens (Topham et al., 1983).

Peripheral blood cells can be stimulated to divide even though the target cell is relatively insensitive (Newton and

Lilly, 1986). It is necessary to stimulate them with a mitogen since the number of lymphocytes which are dividing at any one time is very low. Cells are in G_0 when exposure is taking place, so they may not be sensitive to cell-cycle stage-specific mutagens, and any damage might be repaired before sampling.

9.5.1.2 Micronuclei The assessment of micronuclei is considered simpler than the assessment of metaphase analysis (Heddle, 1973; Schmid, 1975; Heddle et al., 1983; Mavourin et al., 1990). This assay is most often carried out in bone marrow cells, where polychromatic erythrocytes are examined. Damage is induced in the immature erythroblast and results in a micronucleus outside the main nucleus, which is easily detected after staining as a chromatid-containing body. When the erythroblast matures, the micronucleus, whose formation results from chromosome loss during cell division or from chromosome breakage forming centric and acentric fragments, is not extruded with the nucleus. Micronuclei can also be detected in peripheral blood cells (MacGregor et al., 1980; Hayashi et al., 1994, 2007). In addition, they can be detected in the liver (Tates et al., 1980; Braithwaite and Ashby, 1988; Fellows and O'Donovan, 2007; Fenech, 2007; Corvi et al., 2008) after partial hepatectomy or stimulation with 4-acetylaminofluorene, or they can be detected in any proliferating cells. The assays can be performed after single or repeat exposure and not just in rodents but also primates (Hamada et al., 2001). Sex and strain of test animals (mouse) can influence test performance (The Collaborative Study Group for the Micronucleus Test, 1986, 1988). Interpretation requires careful consideration and control of test conditions—some can cause increases in micronuclei due to nongenotoxic mechanisms. The *in vivo* test system can detect genotoxic responses not seen *in vitro* (Tweats et al., 2007a, b).

9.5.2 Germ Cell Assays

The study of chromosomal damage is highly relevant to the assessment of heritable cytogenetic damage. Many compounds which cause somatic cell damage have not produced germ cell damage (Holden, 1982), and, so far, all germ mutagens have also produced somatic damage.

Germ cell data, however, are needed for genetic risk estimation, and testing can be performed in male or female germ cells. The former are most often used, owing to the systemic effects in females. Testing in the male is performed in mitotically proliferating premeiotic spermatogonia, but chromosomal errors in such cells can result in cell death or prevent the cell from passing through meiosis. Damage produced in postmeiotic cells, the spermatids, or sperm is more likely to be transmitted to the F_1 progeny (Albanese, 1987). In females it is during early fetal development of the ovary that oocyte stage is the most commonly tested stage in the

adult female. To test other stages during the first or second meiotic divisions demands the use of oocytes undergoing ovulation which occur naturally or are hormone stimulated. It is thus more difficult technically to test female germ cells.

9.5.3 Heritable Chromosome Assays

Damage may be analyzed in the heritable translocation test, which involves the examination in male F_1 animals if diakinesis metaphase I spermatocytes for multivalent association fall within the acceptable range for the laboratory for a substance to be considered positive or negative under the conditions of the study.

9.5.4 Germ Cell Cytogenetic Assays

Either mouse or rat can be used but the mouse is generally the preferred species. Normally such assays are not conducted for routine screening purposes.

Spermatogonial metaphases can be prepared by the air-drying technique of Evans et al. (1964) for the first and second meiotic metaphase (MI and MII) in the male mouse. This method is not so suitable for rat and hamster. The numbers of spermatogonial metaphases can be boosted if, prior to hypotonic treatment, the testicular tubules are dispersed in trypsin solution (0.25%). At least 1 month between treatment and sample should be allowed to pass in the mouse to allow treated cells to reach meiosis. Brook and Chandley (1986) established that 11 days and 4 h was required for spermatogonial cells to reach preleptotene and 8 days and 10 h to reach zygotene. It takes 4 h for cells to move from MI to MII but test compounds can alter this rate. A search for multivalent formation can be made at MI for the structural rearrangements induced in spermatogonia. Cawood and Breckon (1983) examined the synaptonemal complex at pachytene, using electron microscopy. Errors of segregation should be searched for at the first meiotic division in the male mouse, MII cells showing 19 (hypoploid) and 21 (hyperploid) chromosomes (Brook and Chandley, 1986). Hansmann and El-Nahass (1979), Brook (1982), and Brook and Chandley (1985) describe assays in the female mouse and procedures used for inducing ovulation by hormones and treatment of specific stages of meiosis.

9.6 SISTER CHROMATID EXCHANGE ASSAYS

SCEs are reciprocal exchanges between sister chromatids. They result in a change in morphology of the chromosome, but breakage and reunion are involved although the exact mechanism is unclear. They are thought to occur at homologous loci.

In 1958 Taylor demonstrated SCEs using autoradiographic techniques to detect the disposition of labeled DNA following incorporation of [3 H]-thymidine (Taylor, 1958).

5-Bromo-2'-deoxyuridine (BrdU) has now replaced [³H]-thymidine, and various staining methods have been used to show the differential incorporation of BrdU between sister chromatids: fluorescent Hoechst 33258 (Latt, 1973), combined fluorescent and Giemsa (Perry and Wolff, 1974), and Giemsa (Korenberg and Freedlender, 1974). The fluorescent plus Giemsa procedure is recommended in view of the fact that stained slides can be stored and microscope analysis is simpler.

So that SCEs can be seen at metaphase, cells must pass through S phase (Kato, 1973, 1974; Wolff and Perry, 1974). SCEs appear to occur at the replication point, since SCE induction is maximal at the beginning of DNA synthesis but drops to zero at the end of S phase (Latt and Loveday, 1978).

For SCE analysis *in vitro*, any cell type that is replicating or can be stimulated to divide is suitable. The incorporation of BrdU into cells *in vivo* allows the examination of a variety of tissues (Latt et al., 1980; Sinha et al., 1985, 1988). Edwards et al. (1993) suggest that it is necessary to standardize protocols measuring SCE since different responses can be obtained depending on the extent of simultaneous exposure of test compound and BrdU.

9.6.1 Relevance of SCE in Terms of Genotoxicity

SCEs do not appear to be related to other cytogenetic events, since potent clastogens such as bleomycin and ionizing radiation induce low levels of SCE (Perry and Evans, 1975). The mechanisms involved in CAs and SCE formation are dissimilar (e.g., Galloway and Wolff, 1979). There is no evidence that SCEs are in themselves lethal events, since there is little relationship to cytotoxicity (e.g., Bowden et al., 1979). It was suggested by Wolff (1977a, b) that they relate more to mutational events due to a compatibility with cell survival. However, there are examples of agents that induce significant SCE increases in the absence of mutation (Bradley et al., 1979) as well as they converse (Connell, 1979; Connell and Medcalf, 1982).

The SCE assay is particularly sensitive for alkylating agents and base analogs, agents causing single-strand breaks in DNA, and compounds acting through DNA binding (Latt et al., 1981). The most potent SCE inducers are S phase dependent. Painter (1980) reports that agents such as X-irradiation, which inhibits replicon initiation, are poor SCE inducers, whereas mitomycin C, which inhibits replication fork progression, is a potent SCE inducer.

9.6.2 Experimental Design

Established cell lines and primary cell cultures of rodents may be used. Detailed information on *in vitro* and *in vivo* assays may be obtained in reviews of SCE methods by Latt et al. (1977, 1981), Perry and Thomson (1984), and Perry et al. (1984). The *in vitro* methods will be briefly explored here.

Either monolayer or suspension cultures can be employed, or human lymphocytes. Human fibroblasts are less suitable because of their long cell-cycle duration.

The concentration of organic solvents for the test compound should not exceed 0.8% v/v, as higher concentrations could lead to slight elevations in the SCE level (Perry et al., 1984).

For monolayer cultures, the cultures are set up the day before BrdU treatment so that the cells will be in exponential growth before the addition of BrdU or the test compound. After BrdU addition the cells are allowed to undergo the equivalent of two cell cycles before cell harvest. A spindle inhibitor such as colchicine or colcemid is introduced for the final 1–2 h of culture to arrest cells in metaphase, after which the cells are harvested and chromosome preparations are made by routine cytogenetic techniques.

In the absence of metabolic activation, BrdU and the test agent can be added simultaneously and left for the duration of BrdU labeling. Shorter treatments should be used in the presence of metabolic activation or to avoid synergistic effects with BrdU, when cells can be pulse treated for, for example, 1 h before BrdU addition (see Edwards et al., 1993).

Peripheral blood cultures are established in medium containing BrdU and PHA. Colcemid is added 1–2 h before harvest and the cells are harvested between 60 and 70 h post-PHA stimulation. Cell harvest and slide preparations are conducted according to routine cytogenetic methods.

Heparinized blood samples may be stored at 4°C for up to 48 h without affecting the SCE response (Lambert et al., 1982). If the test agent is known to react with serum or red blood cells, the mononuclear lymphocytes may be isolated by use of a Ficoll/Hypaque gradient (Boyum, 1968).

If metabolic activation is not required, treatment is best conducted over the whole of the final 24 h of culture, or if metabolic activation is required, a pulse exposure may be employed to treat cultures at the first S phase at around 24–30 h or at 48 h for an asynchronous population.

Exposure of cells to fluorescent light during the culture period leads to photolysis of BrdU-containing DNA and a concomitant increase in SCE frequency (Wolff and Perry, 1974). Consequently, SCE cultures should be kept in the dark and manipulated under subdued light conditions such as yellow safe light. Furthermore, media used in SCE assays should be stored in the dark, since certain media components produce reactive SCE-inducing intermediates on exposure to fluorescent light (Monticone and Schneider, 1979).

Coded and randomized slides should be read. All experiments should be repeated at least once (Perry et al., 1984) with higher and lower concentrations of S9 mix if a negative response is achieved. Even for an apparently unambiguous positive response with a greater than twofold increase in SCEs over the background level at the highest dose and with at least two consecutive dose levels with an increased SCE response, a repeat study is necessary to show a consistent response.

The quality of differential staining will determine the ease and accuracy of SCE scoring, and, to eliminate variation, results from different observers should occasionally be compared. Furthermore, to avoid observer bias, scorers should have slides from different treatment groups equally distributed among them, as with all cytogenetic studies.

9.6.2.1 Issues in Assay Interpretation and Relevance The reason for consideration of a significant revision to S2 (such as is now under consideration) ties in issues and considerations as to perceived unacceptable error rates in test performance and difficulties in interpreting the relevance of findings.

Supplement *in vivo* genotoxicity studies are used (i) to follow up on positive finding in one or more tests of the standard genotoxicity battery, (ii) to elucidate a potential contribution of genotoxicity to the induction of preneoplastic and/or neoplastic changes detected in long-term tests in rodents, and (iii) to elucidate mechanisms of micronucleus formation to differentiate clastogenic from aneugenic effects since aneugenecity is well accepted to result from mechanisms of action for which thresholds exist, demonstrating that micronucleus formation is a result of chromosome loss should allow an acceptable level of human exposure to be defined. No matter the trigger for conducting supplemental *in vivo* genotoxicity testing, it is critical that the approach utilized—for example, the end point and target tissue assessed—is scientifically valuable, such that the results will aid in interpreting the relevance of the initial finding of concern. Ultimately, the goal of supplemental genotoxicity testing is to determine if a genotoxic risk is posed to patient under intended condition of treatment.

Follow-Up Testing of Drug Candidates Positive in the Standard Genotoxicity Test Battery It has been reported that approximately 30–50% of pharmaceuticals produce positive genotoxicity results *in vitro* (Kirkland and Muller, 2000). In contrast, results from bone marrow cytogenetic assays are frequently negative, even for those compounds that produce positive results *in vitro*. This discrepancy may result from a number of major differences that exist when testing in cultured cells versus intact animals. For example, differing metabolic pathways can exist *in vitro* and *in vivo*, metabolic inactivation can occur in the intact animal, parent compound or active metabolite may not reach the target cell *in vivo*, rapid detoxification and elimination may occur, or plasma levels *in vivo* may not be comparable to concentrations that generate positive responses in the *in vitro* assays, which is often accompanied by high levels of cytotoxicity. It is also worth nothing that positive results generated *in vitro* may be secondary to effects, such as cytotoxicity, which may never be achieved under *in vivo* exposure conditions (Kirkland et al., 2007). Data from *in vivo* experiments are therefore essential before definitive conclusions are drawn regarding the potential mutagenic hazard to humans from chemicals that produce positive results in one or more *in vitro* tests.

Follow-Up Testing of Tumorigenic Drug Candidates Negative in the Standard Genotoxicity Test Battery In carcinogenicity testing of pharmaceutical drug candidates of a tumorigenic response in rodents, the ICH guidance S2B currently stipulates that such tumorigenicity is not clearly based on a nongenotoxic mechanism. Typically, supplemental *in vivo* genotoxicity tests should be performed with cells of the respective tumor target organ to distinguish between genotoxic and nongenotoxic mechanism of tumor induction.

End Points Assessed in Supplemental Assays Commonly applied test systems are described that are used as supplemental genotoxicity assays. These assays differ with respect to the end points assessed:

1. Induction of primary DNA lesions, that is, measurement of exposure, uptake, and reactivity to DNA via the comet assay or P-postlabeling assay
2. Measurement of the repair of DNA lesion using the UDS test
3. Measurement of induction of genetic drug transgenic animal assays for point mutations or the mouse spot test

The comet assay is the most commonly applied of these.

The issue of how to assess the relevance of a finding of genotoxicity in a candidate drug is a complex one. If the judgment is that such a finding is not relevant to human risks, there are two approaches to assessing and defending such a finding: the weight of evidence (WOE) or the mode of action (MOA). The human biological relevance of a finding of genotoxicity is subject to some complex considerations (Müller and Kasper, 2000).

In the case of impurities in a marketed or candidate drug, the approach is more straightforward. One must reduce the level to or below the toxicological threshold of concern (TTC), which is such a level that no patient would receive more than $1.5 \mu\text{g day}^{-1}$ (EMEA, 2004).

REFERENCES

- Adler ID, Ashby J. (1989) The present lack of evidence for unique rodent germ-cell mutagens. *Mutat. Res.* 212:55–66.
- Advisory Committee on Dangerous Pathogens. (1984) Acquired immune deficiency syndrome (AIDS)—interim guidelines. In: *Circular HC(85)2*. Health Publications Unit, Haywood.
- Albanese R. (1987) Mammalian male germ cell cytogenetics. *Mutagenesis*. 2:79–85.
- Albanese R, Mirkova E, Gatehouse D, Ashby J. (1988) Species-specific response to the rodent carcinogens 1,2-dimethylhydrazine and 1,2-dibromochloropropane in rodent bone marrow micronucleus assays. *Mutagenesis*. 3:35–38.

- Amacher DE, Paillet SC, Turner GN, Ray VA, Salsburg DS. (1980) Point mutations at the thymidine kinase locus in L5178Y mouse lymphoma cells. 2. Test validation and interpretation. *Mutat. Res.* 72:447–474.
- Ames B. (1971) The detection of chemical mutagens with enteric bacteria. In: *Chemical Mutagens, Principles and Methods for Their Detection* (Hollaender A, Ed.), Vol. 1. Plenum Press, New York, pp. 267–282.
- Ames BN, McCann J. (1981) Validation of the Salmonella test: a reply to Rinkus and Legator. *Cancer Res.* 41:4192–4203.
- Ames BN, Durston WE, Yamasaki E, Lee FD. (1973) Carcinogens are mutagens: a simple test system combining liver homogenates for activation and bacteria for detection. *Proc. Natl. Acad. Sci. U.S.A.* 70:2281–2285.
- Ames BN, McCann J, Yamasaki E. (1975) Methods for detecting carcinogens and mutagens with the Salmonella/mammalian-microsome mutagenicity test. *Mutat. Res.* 31:237–364.
- Anderson D. (1990) Male mediated F1 abnormalities. *Mutat. Res.* 229:103–246.
- Anderson D, Jenkinson PC, Dewdney RS, Francis AT, Godbert P, Butterworth KR. (1989) Chromosome aberrations, mitogen induced blastogenesis and proliferative rate index in peripheral lymphocytes from 106 control individuals in the UK population. *Mutat. Res.* 204:407–420.
- Applegate ML, Moore MM, Broder CB, Burrell A, Juhn G, Kasweck KL, Lin PF, Wadhams A, Hozier JC. (1990) Molecular dissection of mutations at the heterozygous thymidine kinase locus in mouse lymphoma cells. *Proc. Natl. Acad. Sci. U.S.A.* 87:51–55.
- Arlett CF, Cole J. (1990) The third United Kingdom Environmental Mutagen Society collaborative trial: overview, a summary and assessment. *Mutagenesis*. 5(Suppl):85–88.
- Arlett CF, Smith DM, Clark GM, Green JHL, Cole J, McGregor DB, Asquith JC. (1989) Mammalian cell assays based upon colony formation. In: *UKEMS Subcommittee on Guidelines for Mutagenicity Testing. Report Part III: Statistical Evaluation of Mutagenicity Test Data* (Kirkland DJ, Ed.). Cambridge University Press, Cambridge, UK, pp. 66–101.
- Armitage P. (1971) *Statistical Methods in Medical Research*. John Wiley & Sons, New York.
- Ashby J, Tennant RW. (1988) Chemical structure, Salmonella mutagenicity and extent of carcinogenicity as indices of genotoxic carcinogens among 222 chemicals tested in rodents by the US NCI/NTP. *Mutat. Res.* 204:17–115.
- Ashby J, Tennant RW, Zeiger E, Stasiewicz S. (1989) Classification according to chemical structure, mutagenicity to Salmonella and level of carcinogenicity of a further 42 chemicals tested for carcinogenicity by the U.S. National Toxicology Program. *Mutat. Res.* 223:73–104.
- Ashwood-Smith MJ. (1980) Stability of frozen microsome preparations for use in the Ames Salmonella mutagenicity assay. *Mutat. Res.* 69:199–200.
- Bartsch H, Camus A-M, Malaveille C. (1976) Comparative mutagenicity of N-nitrosamines in a semi-solid and in a liquid incubation system in the presence of rat or human tissue fractions. *Mutat. Res.* 37:149–162.
- Beckman DA, Brent RL. (1986) Mechanism of known environmental teratogens: drugs and chemicals. *Clin. Perinatol.* 13:649–687.
- Bender MA, Griggs HG, Bedford JS. (1974) Mechanisms of chromosomal aberration production. III. Chemicals and ionizing radiation. *Mutat. Res.* 23:197–212.
- Bjørseth A, Eidså G, Gether J, Landmark L, Møller M. (1982) Detection of mutagens in complex samples by the Salmonella assay applied directly on thin-layer chromatography plates. *Science*. 215(4528):87–89.
- Blazak WF, Steward BE, Galperin I, Allen KL, Rudd CJ, Mitchell AD, Caspary WJ. (1986) Chromosome analysis of trifluorothymidine-resistant L5178Y mouse lymphoma cells colonies. *Environ. Mutagen.* 8:229–240.
- Bossi L. (1985) Information suppression. In: *Genetics of Bacteria* (Scaife J, Leach D and Galizzi A, Eds.). Academic Press, New York, pp. 49–64.
- Bowden GT, Hsu IC, Harris CC. (1979) The effect of caffeine on cytotoxicity, mutagenesis and sister chromatid exchanges in Chinese hamster cells treated with dihydrodiol epoxide derivatives of benzo(a)pyrene. *Mutat. Res.* 63:361–370.
- Boyum A. (1968) Separation of lymphocytes and erythrocytes by centrifugation. *Scand. J. Clin. Invest.* 21:77–85.
- Bradley MO, Hsu IC, Harris CC. (1979) Relationships between sister chromatid exchange and mutagenicity, toxicity and DNA damage. *Nature*. 282:318–320.
- Braithwaite I, Ashby J. (1988) A non-invasive micronucleus assay in rat liver. *Mutat. Res.* 203:23–32.
- Brendler-Schwaab SY, Schmezer P, Liegibel U, Weber S, Michalek K, Tompa A, Pool-Zobel BL. (1994) Cells of different tissues for in vitro and in vivo studies in toxicology: compilation of isolation methods. *Toxicol. In Vitro*. 8:1285–1302.
- Brendler-Schwaab SY, Hartmann A, Pfuhrer S, Speit G. (2005) The in vivo comet assay: use and status in genotoxicity testing. *Mutagenesis*. 20:245–254.
- Bridges BA. (1972) Simple bacterial systems for detecting mutagenic agents. *Lab. Pract.* 21:413–419.
- Bridges BA. (1976) Short-term screening tests for carcinogens. *Nature*. 261:195–200.
- Bridges BA, Woodgate R, Ruiz-Rubio M, Sharif F, Sedgwick SG, Hubschere U. (1987) Current understanding of UV-induced base pair substitution mutation in *E. coli* with particular reference to the DNA polymerase III complex. *Mutat. Res.* 181:219–226.
- Brook JD. (1982) The effect of 4CMB on germ cells of the mouse. *Mutat. Res.* 100:305–308.
- Brook JD, Chandley AC. (1985) Testing of 3 chemical compounds for aneuploidy induction in the female mouse. *Mutat. Res.* 157:215–220.
- Brook JD, Chandley AC. (1986) Testing for the chemical induction of aneuploidy in the male mouse. *Mutat. Res.* 164:117–125.
- Brusick D. (1986) Genotoxic effects in cultures mammalian cells produced by low pH treatment conditions and increased ion concentrations. *Environ. Mutagen.* 8:879–886.
- Brusick D. (1987a) Genotoxicity produced in cultures mammalian cell assays by treatment conditions. Special issue. *Mutat. Res.* 189:1–80.

- Brusick D. (1987b) *Principles of Genetic Toxicology*, 2nd Ed. Plenum Press, New York/London.
- Burlinson B, Tice RR, Speit G, Agurell E, Brendler-Schwaab SY, Collins AR, Escobar P, Honma M, Kumaravel TS, Nakajima M, Sasaki YF, Thybaud V, Uno Y, Vasquez M, Hartmann A. (2007) In vivo comet assay workgroup, part of the fourth international workgroup on genotoxicity testing: results of the in vivo comet assay. *Mutat. Res. Fundam. Mol. Mech. Mutagen.* 627(1):31–35.
- Burnett C, Fuchs C, Corbett J, Menkart J. (1982) The effect of dimethylsulphoxide on the mutagenicity of the hair-dye, p-phenylenediamine. *Mutat. Res.* 103:1–4.
- Butterworth BE, Slaga TJ. (1987) *Nongenotoxic Mechanisms in Carcinogenesis*, Banbury Report No. 25. Cold Spring Harbor Laboratory, New York.
- Carter TC, Lyon MF, Philips RJS. (1956) Induction of mutations in mice by chronic gamma irradiation; interim report. *Br. J. Radiol.* 29:106–108.
- Carver JH, Machado ML, MacGregor JA. (1985) Petroleum distillates suppress in vitro metabolic activation: higher (S9). required in the Salmonella/microsome mutagenicity assay. *Environ. Mutagen.* 7:369–380.
- Cawood AD, Breckon G. (1983) Synaptonemal complexes as indicators of induced structural change in chromosomes after irradiation of spermatogonia. *Mutat. Res.* 122:149–154.
- Chandley AC. (1981) The origin of chromosomal aberrations in man and their potential for survival and reproduction in the adult human population. *Ann. Genet.* 24:5–11.
- Chu EHY, Mallings HU. (1968) Mammalian cell genetics. II. Chemical induction of specific locus mutations in Chinese hamster cells in vitro. *Proc. Natl. Acad. Sci. U.S.A.* 61:1306–1312.
- Cifone MA, Myhr B, Eiche A, Bolisfodi G. (1987) Effect of pH shifts on the mutant frequency at the thymidine kinase locus in mouse lymphoma L5178Y TK+/- cells. *Mutat. Res.* 189:39–46.
- Clayson DB. (1989) ICPEMC publication No. 17: can a mechanistic rationale be provided for non-genotoxic carcinogens identified in rodent bioassays? *Mutat. Res.* 221:53–67.
- Cleaver JE. (1983) Xeroderma pigmentosum. In: *The Metabolic Basis of Inherited Disease* (Stanbury JB, Wyngaarden JB, Fredrickson DS, Goldstein JC and Brown MS, Eds.). McGraw-Hill, New York, pp. 1227–1248.
- Clive D. (1987) Historical overview of the mouse lymphoma TK+/- mutagenicity assay. In: *Mammalian Cell Mutagenesis*, Banbury Report 28 (Moore MM, Demarini DM, De Serres FJ and Tindall KR, Eds.). Cold Spring Harbor Laboratory, New York, pp. 25–36.
- Clive D. (1989) Mammalian cell genotoxicity: a major role for non-DNA targets? *Mutat. Res.* 223:327–328.
- Clive D, Spector JFS. (1975) Laboratory procedure for assessing specific locus mutations at the TK locus in cultured L5178Y mouse lymphoma cells. *Mutat. Res.* 31:17–29.
- Clive D, Flamm WG, Patterson JB. (1972) A mutational assay system using the thymidine kinase locus in mouse lymphoma cells. *Mutat. Res.* 16:77–87.
- Clive D, Johnson KO, Spector JFS, Batson AG, Brown MMM. (1979) Validation and characterization of the L5178Y TK+/- mouse lymphoma mutagen assay system. *Mutat. Res.* 59:61–108.
- Clive D, McCuen R, Spector JFS, Piper C, Mavourin KH. (1983) Specific gene mutation in L5178Y cells in culture: a report of the U.S. Environmental Protection Agency Gene-Tox Program. *Mutat. Res.* 115:225–251.
- Clive D, Caspary W, Kirkby PE, Krehl R, Moore M, Mayo J, Oberly TJ. (1987) Guide for performing the mouse lymphoma assay for mammalian cell mutagenicity. *Mutat. Res.* 189:145–146.
- Clive D, Bolcsfoldi G, Clements J, Cole J, Homna M, Majeska J, Moore M, Muller L, Myhr B, Oberly T, Oudelhkim M, Rudd C, Shimada H, Sofuni T, Thyband V, Wilcox P. (1995) Consensus agreement regarding protocol issues discussed during the mouse lymphoma workshop: Portland, Oregon, May 7, 1994. *Environ. Mol. Mutagen.* 25:165–168.
- Cole J, Muriel WJ, Bridges BA. (1986) The mutagenicity of sodium fluoride to L5178Y (wild type and TK+/- 3.7.2c) mouse lymphoma cells. *Mutagenesis.* 1:157–167.
- Cole J, Fox M, Garner RC, McGregor DB, Thacker J. (1990) Gene mutation assays in cultured mammalian cells. In: *UKEMS Subcommittee on Guidelines for Mutagenicity Testing. Report Part I Revised* (Kirkland DJ, Ed.). Cambridge University Press, Cambridge, UK, pp. 87–114.
- Collins AR, Duthie SJ, Dobson VL. (1993) Direct enzymic detection of endogenous oxidative base damage in human lymphocyte DNA. *Carcinogenesis.* 14:1733–1735.
- Connell JR. (1979) The relationship between sister chromatid exchange, chromosome aberration and gene mutation induction by several reactive polycyclic hydrocarbon metabolites in cultured mammalian cells. *Int. J. Cancer.* 24:485–489.
- Connell JR, Medcalf AS. (1982) The induction of SCE and chromosomal aberrations with relation to specific base methylation of DNA in Chinese hamster cells by N-methyl-N-nitrosourea and dimethyl sulphate. *Carcinogenesis.* 3:385–390.
- Corvi R, Albertini S, Hartung T, Hoffman S, Maricic D, Pfuhler S, van Benthem J, Vanparys P. (2008) ECVAM retrospective validation of the in vitro micronucleus test (MNT). *Mutagenesis.* 23:271–283.
- Cullum J. (1985) Insertion sequences. In: *Genetics of Bacteria* (Scarfe J, Leach D and Galizzi A, Eds.). Academic Press, New York, pp. 85–94.
- Dean BJ, Danford N. (1984) Assays for the detection of chemically-induced chromosome damage in cultures mammalian cells. In: *Mutagenicity Testing. A Practical Approach* (Venitt S and Parry JM, Eds.). IRL Press, Oxford, pp. 187–232.
- DeSerres FJ, Shelby MD. (1979) Recommendations on data production and analysis using the Salmonella/microsome mutagenicity assay. *Mutat. Res.* 64:159–165.
- Doll R. (1977) Strategy for detection of cancer hazards to man. *Nature.* 265:589–596.
- Dunnett CW. (1955) A multiple comparison procedure for comparing several treatments with a control. *J. Am. Stat. Assoc.* 50:1096–1121.
- Edwards AJ, Moon EY, Anderson D, McGregor DB. (1993) The effect of simultaneous exposure to bromodeoxyuridine and methyl methanesulphonate on sister chromatid exchange frequency in culture human lymphocytes and its mutation research. *Mutat. Res.* 247:117–125.

- Ehling UH, Chu EHY, DeCarli L, Evans HJ, Hayashi M, Lambert B, Neubert D, Thilly WG, Vainio H. (1986) Report 8. Assays for germ-cell mutations in mammals. In: *Long-term and Short-term Assays for Carcinogens. A Critical Appraisal*, IARC Scientific Publications, No. 83 (Montesano R, Bartsch H, Vainio H, Wilbourn J and Yamasaki H, Eds.). International Agency for Research on Cancer, Lyon, pp. 245–265.
- EMA (2006) *Guideline on the Limits of Genotoxic Impurities in Pharmaceuticals*. European Medicines Agency, London, UK.
- Evans HJ, O'Riordan ML. (1975) Human peripheral blood lymphocytes for the analysis of chromosome aberrations in mutagen tests. *Mutat. Res.* 31:135–148.
- Evans EP, Breckon G, Ford CE. (1964) An air-drying method for meiotic preparations from mammalian testes. *Cytogenetics*. 3:289–294.
- Evans HH, Mencl J, Horng MF, Ricanti M, Sanchez D, Hozier J. (1986) Lucus specificity in the mutability of L5178Y mouse lymphoma cells: the role of multilocus lesions. *Proc. Natl. Acad. Sci. U.S.A.* 83:4379–4385.
- Fellows M, O'Donovan MR. (2007) Cytotoxicity in cultured mammalian cells I a function of the method used to estimate it. *Mutagenesis*. 22(4):275–280.
- Fenech M. (2007) Cytokinesis-block micronucleus cytome assay. *Nat. Protoc.* 2(5):1084–1104.
- Ford DK, Yerganian G. (1958) Observations on the chromosomes of Chinese hamster cells in tissue culture. *J. Natl. Cancer Inst.* 21:393–425.
- Forster R, Green MHL, Priestley A. (1980) Optimal levels of S9 fraction in Ames and fluctuation tests: apparent importance of diffusion of metabolites from top agar. *Carcinogenesis*. 2:1081–1085.
- Fox M. (1981) Some quantitative aspects of the response of mammalian cells in vitro to induced mutagenesis. In: *Cancer Biology Reviews* (Marcheloni JJ and Hanna MG, Eds.), Vol. 3. Marcel Dekker, New York, pp. 23–62.
- Friedberg EC. (1988) DNA repair in the yeast *Saccharomyces cerevisiae*. *Microbiol. Rev.* 52:70–102.
- Gad SC. (2005) *Statistics and Experimental Design for Toxicologists and Pharmacologists*, 4th Ed. CRC Press, Boca Raton, FL.
- Gad SC, Spainhour CB, Gad-McDonald SE. (2015) Carcinogenicity and Genotoxicity. In: *Mammalian Toxicology* (Abou-Donia MB, Ed.). John Wiley & Sons, Inc., Hoboken, NJ, pp. 341–384.
- Galloway SM, Wolff S. (1979) The relation between chemically induced SCE's and chromatid breakage. *Mutat. Res.* 61:297–307.
- Galloway SM, Bean CL, Armstrong MA, Deasy D, Kraynak A, Bradley MO. (1985) False positive in vitro chromosome aberration tests with non-mutagens at high concentrations and osmolalities. *Environ. Mutagen.* 7(Suppl. 3):48–49.
- Galloway SM, Aardema MJ, Ishidate M Jr, Ivett JL, Kirkland DJ, Morita T, Mosesso P, Sofuni T. (1994) Report from working group in vitro tests for chromosomal aberrations. *Mutat. Res.* 312(3):241–261.
- Gatehouse DG, Delow GF. (1979) The development of a "microtitreR" fluctuation test for the detection of indirect mutagens and its use in the evaluation of mixed enzyme induction of the liver. *Mutat. Res.* 60:239–252.
- Gatehouse DG, Tweats DJ. (1987) Letter to the editor. *Mutagenesis*. 1:307–308.
- Gatehouse D, Wedd DJ. (1984) The differential mutagenicity of isoniazid in fluctuation assays and Salmonella plate tests. *Carcinogenesis*. 5:391–397.
- Gatehouse DG, Wedd DJ, Paes D, Delow G, Burlinson B, Pascoe S, Brice A, Stemp G, Tweats DJ. (1988) Investigations into the genotoxic potential of Ioxidine, a long-acting H2-receptor antagonist. *Mutagenesis*. 3:57–68.
- Gatehouse DG, Wilcox P, Forster R, Rowland IR, Callander RD. (1990) Bacterial mutation assays. In: *Basic Mutagenicity Tests: UKEMS Recommended Procedures* (Kirkland DN, Ed.). Cambridge University Press, Cambridge, UK, pp. 13–61.
- Gatehouse D, Haworth S, Cebula T, Gocke E, Kier L, Matsushima T, Melcion C, Nohmi T, Ohta T, Venitt S, Zeiger E. (1994) Recommendations for the performance of bacterial mutation assays. *Mutat. Res.* 312:217–233.
- Glover P, Kreul P, Poorman-Allen P, Clive D. (1992) Distinctive aspects of mammalian cell mutagenesis. II. Most mutations induced by most mutagens involve several kilobases of DNA. *Environ. Mol. Mutagen.* 19(Suppl. 20):20.
- Green MHL, Muriel WJ. (1976) Mutagen testing using TRP+ reversion in *E. coli*. *Mutat. Res.* 38:3–32.
- Hamada S, Sutou S, Morita T, Wakata A, Asanami S, Hosoya S, Ozawa S, Kondo K, Nakajima M, Shimada H, Osawa K, Kondo Y, Asano N, Sato S, Tamura H, Yajima N, Marshall R, Moore C, Blakey DH, Schechtman LM, Weaver JL, Torous DK, Proudlock R, Ito S, Namiki C, Hayashi M. (2001) Evaluation of the rodent micronucleus assay by a 28-day treatment protocol: summary of the 13th collaborative study by the collaborative study group for the micronucleus test (CSGMT)/Environmental Mutagen Society of Japan (JEMS)-Mammalian Mutagenicity Study Group (MMS). *Environ. Mol. Mutagen.* 37(2):93–110.
- Hansmann I, El-Nahass E. (1979) Incidence of non-disjunction in mouse oocytes. *Cytogenet. Cell Genet.* 24:115–121.
- Harris CC, Weston A, Willey JC, Trivers GE, Mann DL. (1987) Biochemical and molecular epidemiology of human cancer: indicators of carcinogen exposure, DNA damage and genetic predisposition. *Environ. Health Perspect.* 75:109–119.
- Hartmann A, Kiskinis E, Fjaellman A, Suter W. (2001) Influence of cytotoxicity and compound precipitation on test results in the alkaline comet assay. *Mutat. Res.* 497:199–212.
- Haworth S, Lawlor T, Mortelmans K, Speck W, Zeiger E. (1983) Salmonella mutagenicity results for 250 chemicals. *Environ. Mutagen.* 5(Suppl. 1):1–142.
- Hayashi M, Rice RR, MacGregor JT, Anderson D, Blakey DH, Kirsch-Volders M, Oleson FB Jr, Pacchierotti F, Romagna F, Shimada H, Sutou S, Vamnier B. (1994) In vivo rodent erythrocyte micronucleus assay. *Mutat. Res.* 312:293–304.
- Hayashi M, MacGregor JT, Gatehouse DG, Blakey DH, Gertinger SD, Abransson-Zetterberg L, Krishna G, Morita T, Russo A, Asano N, Suzuki H, Ohyama W, Gibson D. (2007) In vivo erythrocyte micronucleus assay. III. Validation and regulatory acceptance of automated scoring and the use of the rat peripheral blood reticulocytes, with discussion of non-hematopoietic target cells and a single dose-level limit test. *Mutat. Res.* 627:10–30.
- Heddle JA. (1973) A rapid in vivo test for chromosomal damage. *Mutat. Res.* 18:187–190.
- Heddle JA, Bodycote DJ. (1970) On the formation of chromosomal aberrations. *Mutat. Res.* 9:117–126.

- Heddle JA, Hite M, Kirkhart B, Mavournin K, MacGregor JT, Newell GW, Salamone M. (1983) The induction of micronuclei as a measure of genotoxicity. A report of the U.S. Environmental Protection Agency Gene-Tox Program. *Mutat. Res.* 123:61–118.
- Herrera G, Urios A, Alexandre V, Blanco M. (1988) UV light induced mutability in *Salmonella* strains containing the umu DC or the muc AB operon: evidence for a umu C function. *Mutat. Res.* 198:9–13.
- Holden HE. (1982) Comparison of somatic and germ cell models for cytogenetic screening. *J. Appl. Toxicol.* 2:196–200.
- Honma M, Hayashi M, Shimada H, Tanaka N, Wakuri S, Awogi T, Yamamoto KI, Kodani N-U, Nishi Y, Nakadate M, Sofuni T. (1999a) Evaluation of the mouse lymphoma tk assay (microwell method) as an alternative to the in vitro chromosomal aberration test. *Mutagenesis.* 14:5–22.
- Honma M, Zhang L, Sakamoto H, Ozaki M, Takeshita K, Momose M, Hayashi M, Sofuni T. (1999b) The need for long-term treatment in the mouse lymphoma assay. *Mutagenesis.* 14:23–29.
- Hozier J, Sawger D, Clieve D, Moore M. (1982) Cytogenetic distinction between the TK+ and TK– chromosomes in L5178Y/TK+/- -3.7.2.C mouse lymphoma cell line. *Mutat. Res.* 105:451–456.
- Hsie AW, Brimer PA, Mitchell TJ, Gosslee DG. (1975) The dose-response relationship for ethyl methane sulfonate-induced mutation at the hypoxanthine-guanine phosphoribosyl transferase locus in Chinese hamster ovary cells. *Somatic Cell Genet.* 1:247–261.
- Hsie AW, Casciano DA, Couch DB, Krahn DF, O'Neil JP, Whitfield BL. (1981) The use of Chinese hamster ovary cells to quantify specific locus mutation and to determine mutagenicity of chemicals: a report of the gene-tox program. *Mutat. Res.* 86:193–214.
- Hubbard SA, Brooks TM, Gonzalez LP, Bridges JW. (1985) Preparation and characterization of S9-fractions. In: *Comparative Genetic Toxicology* (Parry JM and Arlett CG, Eds.). Macmillan, London, pp. 413–438.
- IARC. (1979) *Handling Chemical Carcinogens in the Laboratory; Problems of Safety*, Scientific Publications, No. 33. International Agency for Research on Cancer, Lyons.
- ICEMC. (1983) Committee & final report: screening strategy for chemicals that are potential germ-cell mutagens in mammals. *Mutat. Res.* 114:117–177.
- ICH (2011) *S2(R1) Genotoxicity Testing and Data Interpretation for Pharmaceuticals Intended for Human Use*. International Conference on Harmonisation, Geneva, Switzerland, Step 4, November 9, 2011.
- ICH (2012) *Technical Requirements for Registration of Pharmaceuticals for Human Use Guidance on Genotoxicity Testing and Data Interpretation for Pharmaceuticals Intended for Human Use S2(R1)*. The International Organization for Standardization, Geneva, Switzerland, Recommended for adoption at Step 4 on November 9, 2011. Adopted at Step 5 on June 7, 2012, 77 FR 33748, pp. 33748–33749.
- ISCN. (1985) An international system for human cytogenetic nomenclature. In: *Report of the Standing Committee on Human Cytogenetic Nomenclature* (Harnden DG and Klinger HP, Eds.). Karger, Basel.
- Ishidate M Jr. (1988a) *Data Book of Chromosomal Aberration Tests In Vitro*. Elsevier, Amsterdam.
- Ishidate M Jr. (1988b) A proposed battery of tests for the initial evaluation of the mutagenic potential of medicinal and industrial chemicals. *Mutat. Res.* 205:397–407.
- Ishidate M Jr, Harnois MC. (1987) The clastogenicity of chemicals in mammalian cells. Letter to the editor. *Mutagenesis.* 2:240–243.
- Ishidate M Jr, Sofuni T, Yoshikawa K, Hayashi M, Nohmi T, Sawada M, Matsuoka A. (1984) Primary mutagenicity screening of food additives currently used in Japan. *Food Chem. Toxicol.* 22:623–636.
- Ishidate M Jr, Harnois MC, Sofuni T. (1988) A comparative analysis of data on the clastogenicity of 951 chemical substances tested in mammalian cell cultures. *Mutat. Res.* 195:151–213.
- ISO (2012) *ISO 10993-12—Biological Evaluation of Medical Devices: Part 12: Sample Preparation and Reference Materials*.
- ISO (2014) *ISO 10993-3—Biological Evaluation of Medical Devices: Part 3, Tests for Genotoxicity, Carcinogenicity and Reproductive Toxicity*.
- Japanese Ministry of Health & Welfare (JMHW) (1989) *Guidelines for Toxicity Studies of Drugs, Notification No. 24 of the First Evaluation and Registration Division*. Japanese Ministry of Health, Labour, and Welfare, Tokyo, Japan, September 11, 1989.
- Japanese Ministry of Health and Welfare and Japanese Ministry of International Trade and Industry (1987, March 31) Japanese Ministry of Health, Labour, and Welfare, Tokyo, Japan, Joint Directives of the Japanese Environmental Protection Agency.
- Jordan E, Saedler H, Starlinger P. (1967) Strong polar mutations in the transferase gene of the galactose operon in *E. coli*. *Mol. Gen. Genet.* 100:296–306.
- Kada T. (1981) The DNA damaging activity of 42 coded compounds in the rec-assay. In: *Evaluation of Short-term Tests for Carcinogens—Report of the International Collaborative Program* (de Serres F and Ashby J, Eds.). Elsevier, North Holland/Amsterdam, pp. 175–182.
- Kado NY, Langley D, Eisenstadt E. (1983) A simple modification of the *Salmonella* liquid-incubation assay. Increased sensitivity for detecting mutagens in human urine. *Mutat. Res.* 121:25–32.
- Kalter K. (1977) Correlation between teratogenic and mutagenic effects of chemicals in mammals. In: *Chemical Mutagens: Principles and Methods for Their Detection* (Hollanender A, Ed.), Vol. 6. Plenum Press, New York, pp. 57–82.
- Kato H. (1973) Induction of sister chromatid exchanges by UV light and its inhibition by caffeine. *Exp. Cell Res.* 82:383–390.
- Kato H. (1974) Induction of sister chromatid exchanges by chemical mutagens and its possible relevance to DNA repair. *Exp. Cell Res.* 85:239–247.
- Kato T, Shinoura Y. (1977) Isolation and characterization of mutants of *Escherichia coli* deficient in induction of mutation by ultraviolet light. *Mol. Gen. Genet.* 156:121–132.
- Kennelly JC, Stanton C, Martin CN. (1984) The effect of acetyl-CoA supplementation on the mutagenicity of benzidines in the Ames assay. *Mutat. Res.* 137:39–45.
- Kirkland DJ, Garner RC. (1987) Testing for genotoxicity-chromosomal aberrations in vitro—CHO cells or human lymphocytes? *Mutat. Res.* 189:186–187.
- Kirkland DJ, Muller L. (2000) Interpretation of the biological relevance genotoxicity test: results the importance of thresholds. *Mutat. Res.* 464:137–1147.

- Kirkland DJ, Marshall RR, McEnaney S, Bidgwood J, Rutter A, Mullineux S. (1989) Aroclor-1254 induced rat liver S-9 causes chromosome aberrations in CHO cells but not in human lymphocytes: a role for active oxygen? *Mutat. Res.* 214:115–122.
- Kirkland DJ, Pfuhler S, Tweats D, Asrdema M, Corvi R, Darroudi F, Elhajouji A, Glatt H, Hastwell P, Hayashi M, Kasper P, Kirchner S, Lynch A, Marzin D, Maurici D, Meunier J-R, Muller L, Nohynek G, Parry J, Parry E, Thybaud V, Tice R, van Benthem J, Vanparys P, White R. (2007) How to reduce false positive results when undertaking in vitro genotoxicity testing and thus avoid unnecessary follow-up animal tests: report of the ECVAM Workshop. *Mutat. Res.* 628(1):31–55.
- Korenberg JR, Freedlender EF. (1974) Giemsa technique for the detection of sister chromatid exchanges. *Chromosoma*. 48:355–360.
- Lambert B, Lindblad A, Holmberg K, Francesconi D. (1982) The use of sister chromatid exchange to monitor human populations for exposure to toxicologically harmful agents. In: *Sister Chromatid Exchange* (Wolff S, Ed.). Wiley, New York, pp. 149–182.
- Latt SA. (1973) Microfluorometric detection of deoxyribonucleic acid replication in human metaphase chromosomes. *Proc. Natl. Acad. Sci. U.S.A.* 770:3395–3399.
- Latt SA, Loveday KS. (1978) Characterization of sister chromatid exchange induction by 8-methoxypsoralen plus UV light. *Cytogenet. Cell Genet.* 21:184–200.
- Latt SA, Allen JW, Rogers WE, Jurgens LA. (1977) In vitro and in vivo analysis of sister chromatid exchange formation. In: *Handbook of Mutagenicity Test Procedures* (Kilbey BJ, Legator M, Nichols W and Ramel C, Eds.). Elsevier, Amsterdam, pp. 275–291.
- Latt SA, Schreck RR, Loveday KS, Dougherty CP, Shuler CF. (1980) Sister chromatid exchanges. *Adv. Hum. Genet.* 31:291–298.
- Latt SA, Allen J, Bloom SE, Carrano A, Falke E, Kram D, Schneider E, Schreck R, Tice R, Whitfield B, Wolff S. (1981) Sister chromatid exchanges: a report of the Gene-Tox Program. *Mutat. Res.* 87:17–62.
- Lawley P. (1989) Mutagens as carcinogens: development of current concepts. *Mutat. Res.* 213:3–26.
- Li AP, Carver JH, Choy WN, Bupta RS, Loveday KS, O'Neill JP, Riddle JC, Stankowski LF, Yang LC. (1987) A guide for the performance of the Chinese hamster ovary cell/hypoxanthine guanine phosphoribosyl transferase gene mutation assay. *Mutat. Res.* 189:135–141.
- Li AP, Aaron CS, Auelitta AE, Dearfield KL, Riddle JC, Slesinski RS, Stankowski LF Jr. (1991) An evaluation of the roles of mammalian cell mutation assays in the testing of chemical genotoxicity. *Regul. Toxicol. Pharmacol.* 14:24–40.
- Linblad WJ, Jackim E. (1982) Mechanism for the differential induction of mutation by kS9 activated benzo(a)pyrene employing either a glucose-6-phosphate dependent NADPH-regenerating system or an isocitrate dependent system. *Mutat. Res.* 96:109–118.
- Lindahl T. (1979) DNA glycosylases, endonucleases for apurinic/aprimidinic sites and base excision repair. *Prog. Nucleic Acid Res. Mol. Biol.* 22:109–118.
- Lindahl T, Sedwick B, Sekiguchi M, Nakabeppu Y. (1988) Regulation and expression of the adaptive response to alkylating agents. *Annu. Rev. Biochem.* 57:133–157.
- Little JW. (1984) Autodigestion of lex A and phage T repressors. *Proc. Natl. Acad. Sci. U.S.A.* 81:1375–1379.
- Little JW, Mount DW. (1982) The SOS regulatory system of *Escherichia coli*. *Cell*. 29:11–22.
- Lorge E, Hayashi M, Albertini S, Kirkland D. (2008) Comparison of different methods for an accurate assessment of cytotoxicity in the in vitro micronucleus test. *Mutat. Res.* 655(1–2):4–21.
- MacGregor JT, Wehr CM, Gould DH. (1980) Clastogen-induced micronuclei in peripheral blood erythrocytes: the basis of an improved micronucleus test. *Environ. Mutagen.* 2:509–514.
- MacPhee DG. (1973) Salmonella typhimurium hisG46 (R-Utrecht): possible use in screening mutagens and carcinogens. *Appl. Microbiol.* 26:1004–1005.
- Madle S, Obe G. (1980) Methods for the analysis of the mutagenicity of indirect mutagens/carcinogens in eukaryotic cells. *Hum. Genet.* 56:7–20.
- MAFF (2000) *Test Guidelines for Agricultural Chemicals*. 12-Nousan No. 8147, 2-1-19-1. Ministry of Agriculture, Forestry, and Fisheries, Tokyo.
- Mahon GAT, Green MHL, Middleton B, Mitchell I, Robinson WD, Tweats DJ. (1989) Analysis of data from microbial colon assays. In: *Statistical Evaluation of Mutagenicity Test Data* (Kirkland DJ, Ed.). Cambridge University Press, Cambridge, UK, pp. 26–65.
- Majeska JB, Holden HE. (1994) Procedure for reduction of test chemical requirement in preliminary assays using the L5178Y TK+/- mouse lymphoma forward mutation assay. *Environ. Mol. Mutagen.* 23:150–154.
- Marinus MG, Morris RN. (1974) Biological function for the 6-methyladenine residues in the DNA of *Escherichia coli* K12. *J. Mol. Biol.* 85:309–322.
- Maron DM, Ames BN. (1983) Revised methods for the Salmonella mutagenicity test. *Mutat. Res.* 113:173–215.
- Mavourin KH, Blakey DH, Cimino MC, Salamone MF, Heddle JA. (1990) The in vivo micronucleus assay in mammalian bone marrow and peripheral blood. A report of the U.S. Environmental Protection Agency Gene-Tox Program. *Mutat. Res.* 239:29–80.
- McCann J, Ames BN. (1976) Detection of carcinogens as mutagens in the Salmonella/microsome test: assay of 300 chemicals: discussion. *Proc. Natl. Acad. Sci. U.S.A.* 73:950–954.
- McCann J, Choi E, Yamasaki E, Ames BN. (1975a) Detection of carcinogens as mutagens in the Salmonella/microsome test: assay of 300 chemicals. *Proc. Natl. Acad. Sci. U.S.A.* 72:5135–5139.
- McCann J, Springarn NE, Kobori J, Ames BN. (1975b) Detection of carcinogens as mutagens: bacterial tester strains with R factor plasmids. *Proc. Natl. Acad. Sci. U.S.A.* 72:979–983.
- Merk O, Speit G. (1998) Significance of formaldehyde-induced DNA-protein crosslinks for mutagenesis. *Environ. Mol. Mutagen.* 32:260–268.
- Miller JH. (1985) Pathways of mutagenesis revealed by analysis of mutational specificity. In: *Genetics of Bacteria* (Scaife J, Leach D and Galizzi A, Eds.). Academic Press, New York, pp. 25–40.
- Miller EC, Miller JA. (1971) The mutagenicity of chemical carcinogens: correlations, problems and interpretations. In: *Chemical Mutagens, Principles and Methods for Their Detection* (Hollaender A, Ed.), Vol. 1. Plenum Press, New York, pp. 83–120.
- Modrich P. (1987) DNA mismatch correction. *Annu. Rev. Biochem.* 56:435–466.
- Monticone RE, Schneider EL. (1979) Induction of SCEs in human cells by fluorescent light. *Mutat. Res.* 59:215–221.

- Moore MM, Honma M, Clements J, Awogi T, Bolcsfoldi G, Cole J, Gollapudi B, Harrington-Brock K, Mitchell A, Muster W, Myhr B, O'Donovan M, Ouldehkim M, San R, Shimada H, Stankowski LF. (2000) Mouse lymphoma thymidine kinase locus gene mutation assay: international workshop on genotoxicity test procedures workgroup report. *Environ. Mol. Mutagen.* 35:185–190.
- Moore MM, Honma M, Clements J, Bolcsfoldi G, Burlinson B, Cifone M, Clarke J, Delongschamp R, Durward R, Fellows M, Gollapudi B, Hou S, Jenkinson P, Loyd M, Majeska J, Myhr B, O'Donovan M, Omori T, Riach C, San R, Stankowski LF Jr, Thakur A, Van Goethem F, Wakuri S, Yoshimura I. (2006) Mouse lymphoma thymidine kinase gene mutation assay: follow-up meeting of the International Workshop on Genotoxicity Testing, Aberdeen, Scotland, 2003, assay acceptance criteria, positive controls, and data evaluation. *Environ. Mol. Mutagen.* 47:1–5.
- Moore MM, Honma M, Clements J, Bolcsfoldi G, Burlinson B, Cifone M, Clarke J, Clay P, Doppalapudi R, Fellows M, Gollapudi B, Hou S, Jenkinson P, Kidd D, Lorge E, Loyd M, Muster W, Myhr B, O'Donovan M, Omori T, Pant K, Riach C, Stankowski LF Jr, Thakur A, Van Goethem F, Wakuri S, Yoshimura I. (2007) Mouse lymphoma thymidine kinase gene mutation assay: meeting of the International Workshop on Genotoxicity Testing, San Francisco, 2005 recommendations for 24-hour treatment. *Mutat. Res.* 627:36–40.
- Morimoto K, Sato M, Koizumi A. (1983) Proliferative kinetics of human lymphocytes in culture measure by autoradiography and sister chromatid differential staining. *Epithel. Cell Res.* 145:249–356.
- Mortelmans KE, Dousman L. (1986) Mutagenesis and plasmids. In: *Chemical Mutagens, Principles and Methods for Their Detection* (de Serres FJ, Ed.), Vol. 10. Plenum Press, New York, pp. 469–508.
- Mortelmans KE, Strocker BAD. (1979) Segregation of the mutator property of plasmid R46 from its ultraviolet-protecting property. *Mol. Gen. Genet.* 167:317–328.
- MRC. (1981) *Guidelines for Work with Chemical Carcinogens in Medical Research Council Establishments*. Medical Research Council, London.
- Muller HJ. (1927) Artificial transmutation of the gene. *Science.* 66:84–87.
- Müller L, Kasper P. (2000) Human biological relevance and the use of threshold arguments in regulatory genotoxicity assessment: experience with pharmaceuticals. *Mutat. Res.* 464:9–34.
- Natarajan AT, Obe G. (1982) Mutagenicity testing with cultured mammalian cells: cytogenetic assays. In: *Mutagenicity, New Horizons in Genetic Toxicology* (Heddle JA, Ed.). Academic Press, New York, pp. 172–213.
- Neudecker T, Henschler D. (1985) Allyl isothiocyanate is mutagenic in *Salmonella typhimurium*. *Mutat. Res.* 30:143–148.
- Newton MF, Lilly LJ. (1986) Tissue specific clastogenic effects of chromium and selenium salts in vivo. *Mutat. Res.* 169:61–69.
- O'Neill JP, Brimer PA, Machanoff R, Hirsch GP, Hsie AW. (1977) A quantitative assay of mutation induction at the hgp^rt locus in CHO cells. *Mutat. Res.* 45:91–101.
- Oberly TJ, Bewsey BJ, Probst GS. (1987) A procedure for the CHO/HGPRT mutation assay involving treatment of cells in suspension culture and selection of mutants in soft agar. *Mutat. Res.* 182:99–111.
- OECD (1997a) *OECD Guidelines for the Testing of Chemicals, Section 4: Health Effects*. No. 476. In Vitro Mammalian Cell Gene Mutation Test, Adopted July 21, 1997. Organisation for Economic Co-operation and Development, Paris.
- OECD (1997b) *OECD Guidelines for the Testing of Chemicals, Section 4: Health Effects*. No. 471. Bacterial Reverse Mutation Test, Adopted July 21, 1997. Organisation for Economic Co-operation and Development, Paris.
- OECD (1997c) *OECD Guidelines for the Testing of Chemicals, Section 4: Health Effects*. No. 474. Mammalian Erythrocyte Micronucleus Test, Adopted July 21, 1997. Organisation for Economic Co-operation and Development, Paris.
- OECD (2010) *OECD Guidelines for the Testing of Chemicals, Section 4: Health Effects*. No. 487. In Vitro Mammalian Cell Micronucleus Test (MNvit), Adopted September 26, 2014. Organisation for Economic Co-operation and Development, Paris.
- OECD (2014a) *OECD Guidelines for the Testing of Chemicals, Section 4: Health Effects*. No. 475. Mammalian Bone Marrow Chromosomal Aberration Test, Adopted September 26, 2014. Organisation for Economic Co-operation and Development, Paris.
- OECD (2014b) *OECD Guidelines for the Testing of Chemicals, Section 4: Health Effects*. No. 473. In Vitro Mammalian Chromosomal Aberration Test, Adopted September 26, 2014. Organisation for Economic Co-operation and Development, Paris.
- Ong T, Mukhtar M, Wolf CR, Zeiger E. (1980) Differential effects of cytochrome P450-inducers on promutagen activation capabilities and enzymatic activities of S-9 from rat liver. *J. Environ. Pathol. Toxicol.* 4:55–65.
- Organization for Economic Co-operation and Development (OECD) (1983) *OECD Guidelines for the Testing of Chemicals*. No. 475. Genetic toxicology: *in vivo* mammalian bone marrow cytogenetic test—chromosomal analysis. Adopted April 4, 1984. OECD, Paris.
- Östling O, Johansson KJ. (1984) Microelectrophoretic study of radiation-induced DNA damages in individual mammalian cells. *Biochem. Biophys. Res. Commun.* 123:291–298.
- Pagano DA, Zeiger E. (1985) The stability of mutagenic chemicals tested in solution. *Environ. Mutagen.* 7:293–302.
- Painter RB. (1980) A replication model of sister-chromatid exchange. *Mutat. Res.* 70:337–341.
- Parry JM, Arlett CF, Ashby J. (1985) An overview of the results of the *in vivo* and *in vitro* test systems used to assay the genotoxicity of BZD, DAT, DAB and CDA in the second UKEMS study. In: *Comparative Genetic Toxicology: The Second UKEMS Collaborative Study* (Parry JM and Arlett CF, Eds.). Macmillan, London, pp. 597–616.
- Perry PE, Evans HJ. (1975) Cytological detection of mutagen/carcinogen exposure by sister chromatid exchange. *Nature.* 258:121–125.
- Perry PE, Thomson EJ. (1984) Sister chromatid exchange methodology. In: *Handbook of Mutagenicity Test Procedures* (Kilbey BJ, Legator M, Nichols W and Ramel C, Eds.). Elsevier, Amsterdam, pp. 495–529.
- Perry PE, Wolff S. (1974) New Giemsa method for the differential staining of sister chromatids. *Nature.* 251:156–158.

- Perry P, Henderson L, Kirkland D. (1984) Sister chromatid exchange in cultured cells. In: *UKEMS Subcommittee on Guidelines for Mutagenicity Testing. Report Part IIA*. United Kingdom Environmental Mutagen Society, London, pp. 89–121.
- Pinney RJ. (1980) Distribution among incompatibility groups of plasmids that confer UV mutability and UV resistance. *Mutat. Res.* 72:155–159.
- Preston RJ, Au W, Bender MA, Brewen JG, Carrano AV, Heddle JA, McFee AF, Wolff S, Wassom JS. (1981) Mammalian in vivo and in vitro cytogenetic assays. *Mutat. Res.* 87:143–188.
- Preston RJ, San Sebastian JR, McFee AF. (1987) The in vitro human lymphocyte assay for assessing the clastogenicity of chemical agents. *Mutat. Res.* 189:175–183.
- Prival MJ, Mitchell VD. (1982) Analysis of a method for testing azo-dyes for mutagenic activity in *S. typhimurium* in the presence of FMN in hamster liver S9. *Mutat. Res.* 97:103–116.
- Prival MJ, Bell SJ, Mitchell VD, Peiperi MD, Vaughn VL. (1984) Mutagenicity of benzidine and benzidine-congener dyes and selected monoazo dyes in a modified Salmonella assay. *Mutat. Res.* 136:33–47.
- Revell SH. (1974) The breakage-and-reunion theory and the exchange theory for chromosome aberrations induced by ionizing radiations: a short history. In: *Advances in Radiation Biology* (Lett JT and Zelle M, Eds.), Vol. 4. Academic Press, New York, pp. 367–415.
- Richardson C, Williams DA, Allen JA, Amphlett G, Changer DO, Phillips B. (1989) Analysis of data from in vitro cytogenetic assays. In: *UKEMS Sub-committee on Guidelines for Mutagenicity Testing. Report. Part III. Statistical Evaluation of Mutagenicity Test Data* (Kirkland DJ, Ed.). Cambridge University Press, Cambridge, UK, pp. 141–154.
- Robinson WD, Green MHL, Cole J, Healy MJR, Garner RC, Gatehouse D. (1989) Statistical evaluation of bacterial mammalian fluctuation tests. In: *Statistical Evaluation of Mutagenicity Test Data* (Kirkland DJ, Ed.). Cambridge University Press, Cambridge, UK, pp. 102–140.
- Rossman TG, Klein CB. (1988) From DNA damage to mutation in mammalian cells: a review. *Environ. Mol. Mutagen.* 11:119–133.
- Roth JR. (1974) Frameshift mutations. *Annu. Rev. Genet.* 8:319–346.
- Russell WL. (1951) X-ray induced mutations in mice. *Cold Spring Harb. Symp. Quant. Biol.* 16:327–336.
- Russell LB. (1984) Procedures and evaluation of results of the mouse spot test. In: *Handbook of Mutagenicity Test Procedures* (Kilbey BJ, Legator M, Nichols W and Ramel C, Eds.). Elsevier, Amsterdam, pp. 393–403.
- Russell WL. (1989) Reminiscences of a mouse specific-locus addict. *Environ. Mol. Mutagen.* 14(Suppl. 16):16–22.
- Russell LB, Cumming RB, Hunsicker PR. (1984) Specific-locus mutation rates in the mouse following inhalation of ethylene oxide and the application of the results to estimation of human genetic risk. *Mutat. Res.* 129:381–388.
- Samson L, Cairns J. (1977) A new pathway for DNA repair in *E. coli*. *Nature*. 267:281–282.
- Schmid W. (1975) The micronucleus test. *Mutat. Res.* 1:9–15.
- Scott D, Danford N, Dean B, Kirkland D, Richardson C. (1983) In vitro chromosome aberration assays. In: *UKEMS Subcommittee on Guidelines for Mutagenicity Testing, Report Part I, Basic Test Battery* (Dean BJ, Ed.). UKEMS, Swansea, pp. 63–86.
- Scott D, Dean BJ, Danford ND, Kirkland DJ. (1990) Metaphase chromosome aberration assays in vitro. In: *UKEMS Subcommittee on Guidelines for Mutagenicity Testing. Report Part I: Revised Basic Mutagenicity Tests, UKEMS Recommended Procedures* (Kirkland DJ, Ed.). Cambridge University Press, Cambridge, UK, pp. 63–86.
- Searle AG. (1984) The specific locus test in the mouse. In: *Handbook of Mutagenicity Test Procedures* (Kilbey BJ, Legator M, Nichols W and Ramel C, Eds.). Elsevier, Amsterdam, pp. 373–391.
- Sedgwick SG, Goodwin PA. (1985) Differences in mutagenic and recombinational DNA repair in enterobacteria. *Proc. Natl. Acad. Sci. U.S.A.* 82:4172–4176.
- Seifried HE, Seifried RM, Clarke JJ, Junghans TB, San RHC. (2006) A compilation of two decades of mutagenicity test results with the Ames *Salmonella typhimurium* and L5178Y mouse lymphoma cell mutation assays. *Chem. Res. Toxicol.* 19(5):627–644.
- Selby PB, Olson WH. (1981) Methods and criteria for deciding whether specific-locus mutation-rate data in mice indicates a positive, negative or inconclusive result. *Mutat. Res.* 83:403–418.
- Shanabroch WG, Walker GC. (1980) Localization of the plasmid (pKM101) gene(s) involved in *recA⁺ lexA⁺*-dependent mutagenesis. *Mol. Gen. Genet.* 129:289–297.
- Sinha AK, Linscombe VA, Gollapudi BB, McClintock ML, Flake RE, Bodner KM. (1984) The incidence of spontaneous cytogenetic aberrations in lymphocytes cultured from normal humans for 48 to 72 hours. *Can. J. Genet. Cytol.* 26:528–531.
- Sinha AK, Gollapudi BB, Linscombe VA, Bruce RJ, McClintock ML. (1985) Analyses for in vitro growth conditions, cytokinetics, and sister chromatid exchanges in lymphocytes cultured from the Sprague–Dawley rat. *In Vitro*. 21:583–587.
- Sinha AK, Bruce RJ, Gollapudi BB. (1988) Cytokinetics and SCE baselines in rat and human lymphocytes during successive divisions in the presence of different culture media. *In Vitro*. 24:927–930.
- Sinha AK, Gollapudi BB, Linscombe VA, McClintock ML. (1989) Utilization of rat lymphocytes for the in vitro chromosomal aberration assay. *Mutagenesis*. 4:147–153.
- Slater EE, Anderson MD, Rosenkranz HS. (1971) Rapid detection of mutagens and carcinogens. *Cancer Res.* 31:970–973.
- Snee RD, Irr JD. (1981) Design of a statistical method for the analysis of mutagenesis at the hypoxanthine-guanine phosphoribosyl transferase locus of cultured Chinese hamster ovary cells. *Mutat. Res.* 85:77–93.
- Snyder R, Green J. (2001) A review of the genotoxicity of marketed pharmaceuticals. *Mutat. Res.* 488:151–169.
- Speit G, Hartmann A. (1995) The contribution of excision repair to the DNA effects seen in the alkaline single cell gel test (comet assay). *Mutagenesis*. 10(6):555–559.
- Speit G, Hartmann A. (2005) The comet assay: a sensitive genotoxicity test for the detection of DNA damage. *Methods Mol. Biol.* 291:85–95.
- Steel RGD, Torrie JH. (1996) *Principles and Procedures of Statistics: A Biometrical Approach*. McGraw-Hill, New York.

- Streisinger G, Okada Y, Emrich J, Newton J, Tsugita A, Terzaghi E, Inouye M. (1966) Frameshift mutations and the genetic code. *Cold Spring Harb. Symp. Quant. Biol.* 31:77–84.
- Tates AD, Neuteboom I, Hofker M, den Engelese L. (1980) A micronucleus technique for detecting clastogenic effects of mutagens/carcinogens (DEN, DMN) in hepatocytes of rat liver in vivo. *Mutat. Res.* 74:11–20.
- Taylor JH. (1958) Sister chromatid exchanges in tritium labeled chromosomes. *Genetics.* 43:515–529.
- Tennant RW, Margolin BH, Shelby MD, Zeiger E, Haseman JK, Spalding J, Caspary W, Resnick M, Stasiewicz S, Anderson B, Minor R. (1987) Prediction of chemical carcinogenicity in rodents from in vitro genetic toxicity assays. *Science.* 236:933–941.
- The Collaborative Study Group for the Micronucleus Test. (1986) Sex differences in the micronucleus test. *Mutat. Res.* 172:151–163.
- The Collaborative Study Group for the Micronucleus Test. (1988) Strain differences in the micronucleus test. *Mutat. Res.* 204:307–316.
- Thomas SM, Sedgwick SG. (1989) Cloning of *Salmonella typhimurium* DNA encoding mutagenic DNA repair. *J. Bacteriol.* 171:5776–5782.
- Thompson ED. (1986) Comparison of in vivo and in vitro cytogenetic assay results. *Environ. Mutagen.* 8:753–767.
- Thust R, Mendel J, Schwarz H, Warzoki R. (1980) Nitrosated urea pesticide metabolites and other nitrosamides. Activity in clastogenicity and SCE assays, and aberration kinetics in Chinese hamster V79-E cells. *Mutat. Res.* 79:239–248.
- Tice RR, Andrews PW, Hirai O, Singh NP. (1991) The single cell gel (SCG) assay: an electrophoretic technique for the detection of DNA damage in individual cells. In: *Biological Reactive Intermediates IV, Molecular and Cellular Effects and Their Impact on Human Health* (Witmer CR, Snyder RR, Jollow DJ, Kalf GF, Kocsis JJ and Sipes JG, Eds.). Plenum Press, New York, pp. 157–164.
- Tice RR, Agurell E, Anderson D, Burlinson B, Hartmann A, Kobayashi H, Miyamae Y, Rojas E, Ryu JG, Sasaki YF. (2000) Single cell gel/comet assay: guidelines for in vitro and in vivo genetic toxicology testing. *Environ. Mol. Mutagen.* 35:206–221.
- Topham J, Albanese R, Bootman J, Scott D, Tweats D. (1983) In vivo cytogenetic assays. In: *Report of UKEMS Sub-committee on Guidelines for Mutagenicity Testing. Part I* (Dea B, Ed.). United Kingdom Environmental Mutagen Society, London, pp. 119–141.
- Tsuchimoto T, Matter BE. (1979) In vivo cytogenetic screening methods for mutagens with special reference to the micronucleus test. *Arch. Toxicol.* 42:239–248.
- Tweats DJ, Gatehouse DG. (1988) Discussion forum: further debate of testing strategies. *Mutagenesis.* 3:95–102.
- Tweats DK, Blakey D, Heflich RH, Jacobs A, Jacobsen SD, Morita T, Nohmi T, O'Donovan MR, Sasaki YF, Sofuni T, Tice R. (2007a) Report of the IWGT working group on strategies and interpretation of regulatory in vivo tests. I. Increases in micronucleated bone marrow cells in rodents that do not indicate genotoxic hazards. *Mutat. Res.* 627:79–91.
- Tweats DK, Blakey D, Heflich RH, Jacobs A, Jacobsen SD, Morita T, Nohmi T, O'Donovan MR, Sasaki YF, Sofuni T, Tice R. (2007b) Report of the IWGT working group on strategy/interpretation on regulatory in vivo tests. II. Identification of in vivo-only positive compounds in the bone marrow micronucleus test. *Mutat. Res.* 627:92–105.
- Venitt S, Crofton-Sleigh C. (1981) Mutagenicity of 42 coded compounds in a bacterial assay using *Escherichia coli* and *Salmonella typhimurium*. In: *Evaluation of Short-term Tests for Carcinogens. Report of the International Collaborative Program. Progress in Mutational Research* (de Serres FJ and Ashby J, Eds.), Vol. 1. Elsevier, New York, pp. 351–360.
- Venitt S, Parry JM. (1984) Background to mutagenicity testing. In: *Mutagenicity Testing: A Practical Approach* (Venitt S and Parry JM, Eds.). IRL Press, Oxford, pp. 1–24.
- Villani G, Boiteux S, Radman M. (1978) Mechanisms of ultraviolet-induced mutagenesis: extent and fidelity of in vitro DNA synthesis on irradiated template. *Proc. Natl. Acad. Sci. U.S.A.* 75:3037–3041.
- deVries JH. (1901) *The Mutation Theory*. Verlag von Veit & Co., Leipzig.
- Wagner R, Meselson M. (1976) Repair tracts in mismatched DNA heteroduplexes. *Proc. Natl. Acad. Sci. U.S.A.* 73:4135–4139.
- Walker GC. (1984) Mutagenesis and inducible responses to deoxyribonucleic acid damage in *Escherichia coli*. *Microbiol. Rev.* 48:60–93.
- Walker GC, Dobson PP. (1979) Mutagenesis and repair deficiencies of *Escherichia coli* umu C mutants are suppressed by the plasmid pKM101. *Mol. Gen. Genet.* 172:17–24.
- Wilcox P, Naidoo A, Wedd DJ, Gatehouse DG. (1990) Comparison of *Salmonella typhimurium* TA102 with 56h *Escherichia coli* WP2 tester strains. *Mutagenesis.* 5:285–291.
- Wolff S. (1977a) Lesions that lead to SCEs are different from those that lead to chromosome aberrations. *Mutat. Res.* 46:164.
- Wolff S. (1977b) *In vitro inhibition of mono-oxygenase dependent reactions by organic solvents*. International Conference on Industrial and Environmental Xenobiotics, Prague, Czechoslovakia, September 13–15, 1977.
- Wolff S, Perry P. (1974) Differential staining of sister chromatids and the study of sister chromatid exchange without autoradiography. *Chromosoma.* 48:341–353.
- Yahagi T, Degawa M, Seino Y, Matsushima T, Nagao M. (1975) Mutagenicity of carcinogen azo dyes and their derivatives. *Cancer Lett.* 1:91–96.
- Yahagi T, Nagao M, Seino Y, Matsushima T, Sugimura T, Okada M. (1977) Mutagenicities of N-nitrosamines on *Salmonella*. *Mutat. Res.* 48:120–130.
- Yanofsky C. (1971) Mutagenesis studies with *Escherichia coli* mutants with known amino acid (and base-pair) changes. In: *Chemical Mutagens, Principles and Methods for Their Detection* (Hollaender A, Ed.), Vol. 1. Plenum Press, New York, pp. 283–287.

QSAR TOOLS FOR DRUG SAFETY

10.1 STRUCTURE–ACTIVITY RELATIONSHIPS

Structure–activity relationship (SAR) methods first became a legitimate and useful part of toxicology in the mid-1970s, though the tools had limited utility. These methods were various forms of mathematical or statistical models which seek to predict the adverse biological effects of chemicals based on their structure. The predictions could be of either a qualitative (mutagenic/not mutagenic, carcinogen/noncarcinogen) or quantitative (LD_{50}) nature, with the second group usually being denoted as quantitative structure–activity relationship (QSAR) models. At the start, the basic techniques utilized to construct such models combined modeling, extrapolation, and reduction of dimensionality methods.

The concept that the biological activity of a compound is a direct function of its chemical structure is now at least a century and a half old (Crum Brown and Fraser, 1869). During most of the twentieth century, the development and use of SARs were the domain of pharmacology and medicinal chemistry. These two fields are responsible for the beginnings of all the basic approaches in SAR work, usually with the effort being called drug design. An introductory medicinal chemistry text (such as Foye et al., 1995 or Thomas et al., 2012) is strongly recommended as a starting place for SAR. Additionally, Burger's Medicinal Chemistry (Abraham and Rotella, 2010; Selassie and Verma, 2010), with its excellent overview of drug structures and activities, should enhance at least the initial stages of identifying the potential biological actions of de novo compounds using a pattern recognition approach. Most recently, Cronin and Livingston (2004), Helma (2005), Seger-Hartmann (2015), Roy et al. (2015), and Instituto di Ricerche... Milano (2015a) have published excellent overviews of methods and applications.

Having already classified SAR methods into qualitative and quantitative, it should also be pointed out that both of these can be approached on two different levels. The first is on a local level, where prediction of activity (or lack of activity) is limited to other members of a congeneric series or structural near neighbors. The accuracy of predictions via this approach is generally greater but is of value only if one has sufficient information on some of the structures within a series of interest.

The second approach is prediction of activity over a wide range, generally based on the presence or absence of particular structural features (functional groups).

For toxicology, SARs had a small but important and increasing number of uses until the last decade, when ICH's incorporation of modern SAR/QSAR methods to identify mutagens and nonmutagens among impurity structures gave a huge boost to their use. These can all be generalized as identifying potentially toxic effects or restated as three main uses:

1. For the selection and design of toxicity tests to address end points of possible concern.
2. If a comprehensive or large testing program is to be conducted, SAR predictions can be used to prioritize the tests so that outlined questions (the answers to which might preclude the need to do further testing) may be addressed first.
3. As an alternative to testing at all. Though in general it is not believed that the state of the art for SAR methods allows such usage, in certain special cases (such as selecting which of several alternative candidate compounds to develop further and then test) this use may be valid and valuable.

It is the regulatory acceptance of QSAR methods in place of actual testing in some cases (ICH, 2006; Kruhlak et al., 2012;

OECD, 2014) which has truly “exploded” the use of the methods and which now drives efforts to both expand their scope and improve their performance.

10.1.1 Basic Assumptions

Starting with the initial assumption that there is a relationship between structure and biological activity, we can proceed to more readily testable assumptions.

First, that the dose of chemical is subject to a number of modifying factors (such as membrane selectivities and selective metabolic actions) which are each related in some manner to chemical structure. Indeed, absorption, metabolism, pharmacologic activity, and excretion are each subject to not just structurally determined actions but also to (in many cases) stereospecific differential handlings.

Given these assumptions, actual elucidation of SARs requires the following:

1. Knowledge of the biological activities of existing structures
2. Knowledge of structural features which serve to predict activity (also called molecular parameters of interest)
3. One or more models which relate two to one with some degree of reliability

There are now extensive sources of information as to both toxic properties of chemicals and, indeed, biological activities. These include books, journals, and manual and computerized databases. The reader is directed to Chapter 3 of this text and Wexler et al. (1999) as a guide to accessing the different sources of toxicology information but is cautioned to remember that there is also extensive applicable information in the realms of medicinal chemistry and pharmacology, as exemplified by Burger's (Abraham and Rotella, 2010).

10.1.2 Molecular Parameters of Interest

Which structural and physicochemical properties of a chemical are important in predicting its toxicologic activity are both open to considerable debate (Kaufman et al., 1983; Tamura, 1983; Tute, 1983). Table 10.1 presents a partial list of such parameters. The reader is referred to a biologically oriented physical chemistry text (such as Chang, 1981) both for explanations of these parameters and for references to sources from which specific values may be obtained.

There are now several systems available to study the three-dimensional structural aspects of molecules and their interactions. Various molecular modeling sets, molecular design and analysis packages, and molecular graphics software packages are available for personal computers and larger systems. Use of such forms of graphic structural

TABLE 10.1 Molecular Parameters of Interest

| |
|---|
| Electronic Effects |
| Ionization constants |
| Sigma substituent constant |
| Distribution constant |
| Resonance effect |
| Field effect |
| Molecular orbital indices |
| Atomic/electron net charge |
| Nucleophilic superdelocalizability |
| Electrophilic superdelocalizability |
| Free radical superdelocalizability |
| Energy of the lowest empty molecular orbital |
| Energy of the highest occupied molecular orbital |
| Frontier atom–atom polarizability |
| Intermolecular coulombic interaction energy |
| Electric field created at point (A) by a set of charges on a molecule |
| Hydrophobic Parameters |
| Partition coefficients |
| Pi substituent constants |
| <i>R_m</i> value in liquid–liquid chromatography |
| Elution time in high-pressure liquid chromatography (HPLC) |
| Solubility |
| Solvent partition coefficients |
| p <i>K_a</i> |
| Steric Effects |
| Intramolecular steric effects |
| Steric substituent constant |
| Hyperconjugation correction |
| Molar volume |
| Molar refractivity, MR substituent constants |
| Molecular weight |
| Van der Waals radii |
| Interatomic distances |
| Substructural Effects |
| Three-dimensional geometry |
| Fragment and molecular properties (see Kubinyi, 1995 for substituent effects) |
| Chain lengths |

examination as a tool or method in SAR analysis has been discussed by Cohen et al. (1974) and Gund et al. (1980). Such methods are generally called topological methods.

10.2 SAR MODELING METHODS

A detailed review of even the major methodologies available for SAR/QSAR modeling in toxicology is beyond the scope of this book. Though we will briefly discuss the major approaches, the reader is directed to one of the several very readable introductory articles (Chang, 1981) or books (Olson and Christoffersen, 1979; Topliss, 1983; Goldberg, 1983; Mackay et al., 2009; or Venkatapathy et al., 2009) for somewhat detailed presentations.

To begin with, it should be made clear that all the actual techniques involved in the performance of SAR analysis are presented elsewhere in this text (Chapter 3 and statistical methods in Chapter 30). It is only their actual application to data which sets such analysis apart from the forms of modeling we have previously looked at.

All the current major SAR methods used in toxicology can be classified based on what kinds of compound-related or structural data they use and what method is used to correlate this structural data with the existing biological data. While most used currently to predict mutagenicity (under ICH M7 or REACH guidances) in lieu of testing, these methods currently offer much additional utility such as predicting carcinogenicity or the potency of carcinogens (see Venkatapathy et al., 2009 and Instituto de Ricerche Farmacologiche Mario Negri Milano, 2015b) or sensitizers or providing an initial screen to support candidate molecule selections.

The more classical approaches use physicochemical data (such as molecular weight, free energies, etc.) as a starting point. The major approaches to this are by manual pattern recognition methods, cluster analysis, or by regression analysis. It is this last in the form of Hansch or linear free energy relationships (LFER) which actually launched all SAR work (other than that on limited congeneric cases) into the realm of being a useful approach. Indeed, still foremost among the QSAR methods is the model proposed by Hansch and his coworkers (Hansch, 1971). It was the major contribution of this group to propose the incorporation of earlier observations of the importance of the relative lipophilicity to biologic activity into the formal LFER approach to provide a general QSAR model for biological effects. As a suitable measure of lipophilicity, the partition coefficient ($\log P$) between 1-octanol and water was proposed, and it was demonstrated that this was an approximately additive and constitutive property and that it was therefore calculable, in principle, from molecular structure. Using a probabilistic model for the Hansch equation, which can be expressed as

$$\log\left(\frac{1}{C}\right) = -k\pi^2 + k'\pi + p\sigma + k''$$

or

$$\log\left(\frac{1}{C}\right) = -k(\log P)^2 + k'(\log P) + p\sigma + k''$$

where C is the dose that elicits a constant biological response (e.g., ED_{50} , LD_{50}), p is the substituent lipophilicity, $\log P$ is the partition coefficient, σ is the substituent electronic effect of Hammett, and k , k' , p , and k'' are the regression coefficients derived from the statistical curve fitting, the partition coefficient ($\log P$) can be calculated as an estimate of lipophilicity. The reciprocal of the concentration reflects the fact

that higher potency is associated with lower dose and the negative sign for the π_0 or $\log P_0$.

The statistical method used to determine the aforementioned coefficients is multiple linear regression. A number of statistics are derived in conjunction with such a calculation, which allow the statistical significance of the resulting correlation to be assessed. The most important of these are s , the standard deviation; r^2 , the coefficient of determination or percentage of data variance accounted for by the model (r , the correlation coefficient is also commonly cited); and F , a statistic for assessing the overall significance of the derived equation, values, and confidence intervals (usually 95%) for the individual regression coefficients in the equation. These must be low to assure true "independence" or orthogonality of the variables, a necessary condition for meaningful results.

In a like manner, there are a number of approaches for using structural and substructural data and correlating these to biological activities. Such approaches are generally classified as regression analysis methods, pattern recognition methods, and miscellaneous others (such as factor analysis, principal components, and probabilistic analysis).

The regression analysis methods which use structural data have been, as we will see when we survey the state of the art in toxicology, the most productive and useful. "Keys"—or fragments of structure—are assigned weights as predictors of an activity, usually in some form of the Free-Wilson model (Free and Wilson, 1964) which was developed at virtually the same time as the Hansch. According to this method, the molecules of a chemical series are structurally decomposed into a common moiety (or core) that may be substituted in multiple positions. A series of linear equations of the form

$$BA_i = \sum_j a_j X_j + \mu$$

are constructed where BA is the biological activity, X_j is the j th substituent with a value of 1 if present and 0 if not, a_j is the contribution of the j th substituent to BA , and μ is the overall average activity. All activity contributions at each position of substitution must sum to zero. The series of linear equations thus generated is solved by the method of least squares for the a_j and μ . There must be several more equations than unknowns and each substituent should appear more than once at a position in different combinations with substituents at other positions. The favorable aspects of this model are:

1. Any set of quantitative biological data may be employed as the dependent variable.
2. No independently determined substituent constants are required.

3. The molecules comprising a sample of interest may be structurally dismembered in any desired or convenient manner.
4. Multiple sites of variable substitution are readily handled by the model.

There are also several limitations: a substantial number of compounds with varying substituent combinations are required for a meaningful analysis; the derived substituent contributions give no reasonable basis for extrapolating predictions from the substituent matrix analyzed; and the model will break down if nonlinear dependence on substituent properties is important or if there are interactions between the substituents.

Pattern recognition methods comprise yet another approach to examining structural features and/or chemical properties for underlying patterns that are associated with differing biological effects (Cronin, 2004; Elkins, 2007). Accurate classification of untested molecules is again the primary goal. This is carried out in two stages. First, a set of compounds, designated the training set, is chosen for which the correct classification is known. A set of molecular or property descriptors (features) is generated for each compound. A suitable classification algorithm is then applied to find some combination and weight of the descriptors that allow perfect classification. Many different statistical and geometric techniques for this purpose have been used and were presented in earlier chapters. The derived classification function is then applied in the second step to compounds not included in the training set to test predictability. In published work these have generally been other compounds of known classification also. Performance is judged by the percentage of correct predictions. Stability of the classification function is usually tested by repeating the procedure several times with slightly altered, but randomly varied, sets or samples.

The main difficulty with these methods is in “decoding” the QSAR in order to identify particular structural fragments

responsible for the expression of a particular activity. And even if identified as “responsible” for activity, far harder questions for the model to answer are whether the structural fragment so identified is “sufficient” for activity, whether it is always “necessary” for activity, and to what extent its expression is modified by its molecular environment. Most pattern recognition methods use as weighting factors either the presence or absence of a particular fragment or feature (coded 1 or 0) or the frequency of occurrence of a feature. They may be made more sophisticated by coding the spatial relationship between features.

Enslein et al. (1984) have published a good brief description of the problems involved in applying these methods in toxicology. His “toxicity prediction by computer-assisted technology” (TOPKAT) program was a program running on early desktop computers that offered utility in predicting acute toxicity end points and mutagenicity. The program is still (in a limited number) commercially available.

10.3 APPLICATIONS IN TOXICOLOGY

True SAR methods have been developed to predict a number of toxicological end points (mutagenesis, carcinogenesis, dermal sensitization, lethality (LD_{50} values), biological oxygen demands, and teratogenicity) with varying degrees of accuracy, and models for the prediction of other end points continue under development. An earlier “semi-SAR” tool was the Cramer decision tree in the 1970s (Cramer et al., 1978). The purpose of the decision tree was to categorize the mass of chemical structures for which there was limited or no data as either potentially very toxic (and therefore of high priority to test—Category 1) to Category 3 (unlikely to be toxic). The current version of the Cramer Tree (ToxTree) still serves primarily the same purpose. Some of these existing models are presented by category of use in Table 10.2. Additionally, both Environmental Protection Agency (EPA)

TABLE 10.2 Existing SAR Models for Toxicology End Points

| End Point | Prediction | | Reference |
|--------------------------------|--------------|-------------|---|
| | Quantitative | Qualitative | |
| Mutagenicity | X | | Asher and Zervos (1977) |
| Carcinogenicity | | X | Niculescu-Duvaz et al. (1981) |
| | X | | Enslein et al. (1983) |
| | X | | Franke (1973) |
| | X | | Asher and Zervos (1977) |
| | X | X | Niculescu-Duvaz et al. (1981) |
| | X | | Enslein and Craig (1982) |
| Sensitization | X | | Dupuis and Benezra (1982) |
| LD_{50} | | X | Enslein et al. (1983) |
| Developmental toxicity | X | | Enslein et al. (1983) |
| Biological oxygen demand (BOD) | | X | Enslein et al. (1984) |
| Reproductive toxicity | | X | Bernstein (1984), Enslein et al. (1983) |

and Food and Drug Administration (FDA) have models for mutagenicity/carcinogenicity that they utilize to “flag” possible problem compounds.

Note that all models are only as good as the data that they are based on. Issues of data quality and range of possible structural features incorporated into the data set utilized remain an issue (Sahigara et al., 2012).

It should be expected that qualitative models are more “accurate” than quantitative ones and that the more possible mechanisms associated with an end point, the less accurate (or more difficult) a prediction. Table 10.3 contains end points which each of the major programs evaluates.

10.3.1 Metabolism

A central tenet of both toxicology and pharmacology is that any assessment of either adverse or desirable effects of a drug must incorporate an assessment of the absorption, distribution, elimination, and metabolism of the administered drug molecule. While Chapter 17 looks at the generating and evaluation of such ADME data, the prediction of values of absorption (particularly from the skin—see Baba et al., 2015) and of metabolic routes and rates (Kirchmair et al., 2015) are areas of both active technology development and of great interest. Kirchmair presents a table summarizing the aspects and capabilities of 34 different available programs for use in metabolic predictions, including Meteor which will be discussed further later. It should be noted that most of the current toxicology predictive programs do not explicitly incorporate such predictions.

10.3.2 Reproductive

SARs have not been as successful for reproductive and developmental toxicology. Data are available that suggest SARs for certain classes of chemicals (e.g., glycol ethers, phthalate esters, heavy metals). Yet, for other agents, nothing in their structure would have identified them as male reproductive toxicants (e.g., chlordecone). Bernstein (1984) reviewed the literature and has offered a set of classifications relating structure to reported male reproductive activity. Although limited in scope and in need of rigorous validation, such schemes do provide hypothesis that can be tested. Enslein et al. (1983) have proposed a commercial computer model for this and developmental end points. Comparison of the chemical or physical properties of an agent with those of known male reproductive toxicants may provide some indication of a potential for reproductive toxicity. Such information may be helpful in setting priorities for testing of agents or for the evaluation of potential toxicity when only minimal data are available.

10.3.3 Eye Irritation

QSAR analysis, widely used to predict various physiological and biochemical activities of novel chemicals, also has been used to predict eye irritancy of structurally related chemicals. Using QSAR, Sugai et al. (1990) examined the eye irritancy (opacity and conjunctivitis) of 131 chemically heterogeneous chemicals. The accuracy was 86.3% for classifying irritancy of the chemicals. Overall accuracy rates as high as 91% have

TABLE 10.3 Comparison of Commercial Software Features

| Category | Derek (v3.4) | MCASE (v3.45) | TOPKAT (v5.01) | Leadscope |
|--|-----------------------------------|------------------------|---------------------------------------|----------------|
| <i>Models</i> | | | | |
| Genotoxicity | Yes | Yes | Yes | Yes |
| Carcinogenicity | Yes | Yes | Yes | Yes |
| Teratogenicity/developmental toxicity | Yes ^a | Yes | Yes | Yes |
| Structural rules | Yes | Yes | No | No |
| QSAR | No | Yes | Yes | No |
| User editable models | Yes | | No | No |
| Model appropriateness alert | No | No | Yes | No |
| Integrated metabolism prediction | Yes | Yes | No | No |
| Inorganic metal compounds | Yes | No | No | No |
| <i>Interface</i> | | | | |
| Batch mode | Yes | Yes | Yes | Yes |
| Output file | Yes | Yes | No | Yes |
| Operating system | Window/UNIX | Open VMS | Windows | Mainframe |
| Compound input formats | ChemDraw/ISIS draw, MOL and SD | SMILES strings and MOL | SMILES strings | SMILES strings |
| Additional toxicity/compound information | References and rules descriptions | Training set | Training set compounds and references | No |

^aOnly seven structural alerts implemented for teratogenicity.

been reported. Although this approach may provide useful information on structurally related chemicals, its utility is limited for formulated products.

10.3.4 Lethality

Analysis of the SARs within a class of chemicals can yield valuable information and may reduce the number of bioassays conducted. QSAR analysis is particularly useful during the discovery stage for selection of chemicals for further development. QSAR also can be used for prioritization of chemicals for various actions related to health and safety and environmental assessment. The elements generally needed for QSAR include a verified bioassay database for the end point to be predicted, a set of chemical–physical parameters which described the chemical structures so that the end point can be modeled in terms of these parameters, statistical techniques, that is, principally multivariate regression and discriminant analysis for weighing these parameters in a near-optimum fashion for the explanation of the end point, and computer technology to make it all practical. Using QSAR, Enslein et al. (1983) has analyzed 2066 chemicals of various chemical structures and found that the oral rat LD₅₀ of almost 50% of the compounds examined was predicted within a factor of 2 and 95% within a factor of 8. Obviously, there are limitations for the QSAR approach to predict a complex toxic response in whole animals. These include limited database on which to base a QSAR model, the temptation to extrapolate beyond the confines of the model, and the noise inherent in the bioassays on which the models are based. The results from QSAR have to be used with caution, and at this stage, QSAR is useful during the discovery stage and for prioritizing chemicals.

10.3.4.1 Oral Rat LD₅₀ Oral Rat LD₅₀ is the administered dose of chemical in milligram per kilogram body weight that causes 50% of rats to die after oral ingestion. The data set for this end point contains 7420 chemicals, and predictions by the EPA Toxicity Estimation Software Tool (TEST) and TOTKAT (both predict oral LD₅₀ values) methods have a good correlation with experimental data. Table 10.4 presents a summary of the Predictive Performance using this data set.

TEST allows you to estimate the value for several toxicity end points using several different advanced QSAR methodologies. For predicting mutagenicity, a data set of 6512

chemicals was compiled from several different sources. The final data set consists of 5743 chemicals (after removing salts, mixtures, ambiguous compounds, and compounds without CAS numbers).

The Ames mutagenicity data set statistical external for validation (Bercu et al., 2009) has shown that the consensus method achieved the best prediction accuracy (concordance) and prediction coverage. The consensus method's predicted toxicity is estimated by taking an average of the predicted toxicities from the hierarchical clustering, FDA, and nearest neighbor QSAR methods (provided the predictions are within the respective applicability domains (ADs)):

- *Hierarchical method* The toxicity for a given query compound is estimated using the weighted average of the predictions from several different cluster models.
- *FDA method* The prediction for each test chemical is made using a new model that is fit to the chemicals that are most similar to the test compound.
- *Nearest neighbor method* The predicted toxicity is estimated by taking an average of the three chemicals in the training set that are most similar to the test chemical.

The prediction accuracy is then evaluated in terms of the fraction of compounds that are predicted accurately. The prediction accuracy is evaluated in terms of three different statistics—concordance, sensitivity, and specificity:

- Concordance is the fraction of all compounds that are predicted correctly (i.e., experimentally active compounds that are predicted to be active and experimentally inactive compounds that are predicted to be inactive).
- Sensitivity is the fraction of experimentally active compounds that are predicted to be active.
- Specificity is the fraction of experimentally inactive compounds that are predicted to be inactive.

10.3.5 Carcinogenicity

Given the 1–2 million dollars in cost and the 3–5 years required to test a single chemical in a lifetime rodent carcinogenicity bioassay, initial decisions on whether to continue the development of a chemical, submit a premanufacturing

TABLE 10.4 Statistical External Validation for Oral LD₅₀ Prediction Using Data Set

| Method | R^2 | $(R^2 - R_0^2)/R^2$ | k | RMSE | MAE | Coverage |
|------------------|-------|---------------------|-------|-------|-------|----------|
| Hierarchical | 0.568 | 0.200 | 0.955 | 0.652 | 0.460 | 0.826 |
| FDA | 0.542 | 0.220 | 0.950 | 0.668 | 0.492 | 0.980 |
| Nearest neighbor | 0.519 | 0.340 | 0.954 | 0.677 | 0.491 | 0.994 |
| Consensus | 0.604 | 0.277 | 0.952 | 0.604 | 0.441 | 1.000 |

notice (PMN), or require additional testing may be based largely on SARs and limited short-term assays. A test agent's structure, solubility stability, pH sensitivity, electrophilicity, and chemical reactivity can represent important information for hazard identification. Historically, certain key molecular structures have provided regulators with some of the most readily available information on which to assess hazard potential. For example, 8 of the first 14 occupational carcinogens were regulated together by the Occupational Safety and Health Administration (OSHA) as belonging to the aromatic amine chemical class. The EPA Office of Toxic Substances relies on SARs to meet deadlines for responding to PMN for new chemical manufacture under the Toxic Substances Control Act (TSCA). Structural alerts such as *n*-nitroso or aromatic amine groups, amino azo dye structures, and phenanthrene nuclei are clues to prioritizing agents for additional evaluation as potential carcinogens. The limited database of known developmental toxicants limits SARs to only a few chemical classes, including chemicals with structures related to those of valproic acid and retinoic acid.

More than 100 years ago, Richet (1893), Meyer (1899), and Overton (1901) independently found linear relationships between lipophilicity, expressed as solubility or oil–water partition coefficients, and biological effects, like toxicity and narcotic activity (Gad, 2014). From here the use of predictive models grew to what is now commonly known as QSAR methodology.

The transition from the classic computational approaches and the current more useful methods started with the Cramer rules, which has since evolved into ToxTree.

SARs are useful in assessing the relative toxicity of chemically related compounds. The EPA's (1994) reassessment of the risk of 2,3,7,8-tetrachlorodibenzo-*p*-dioxin and related chlorinated and brominated dibenzo-*p*-dioxins, dibenzofurans, and planar biphenyls might have relied too heavily on toxicity equivalence factors (TEFs) based on induction of the Ah receptor (EPA, 1994). The estimated toxicity of environmental mixtures containing those chemicals is a product of the concentration of each chemical times its TEF value. However, it is difficult to predict activity across chemical classes and especially across multiple toxic end points by using a single biological response. Many complex chemical–physical interactions are not easily understood and may be oversimplified by researchers. Several computerized SAR methods gave disappointing results in the National Toxicology Program's (NTP) 44-chemical rodent carcinogenicity prediction challenge (Ashby and Tennant, 1994).

If we take a series of chemicals and attempt to form a quantitative relationship between the biological effects (i.e., the activity) and the chemistry (i.e., the structure) of each of the chemicals, then we are able to form a QSAR. Less complex, or quantitative, understanding of the role of structure to govern effects, that is, that a fragment or substructure could result in a certain activity, is often simply

termed an SAR. Together SARs and QSARs can be referred to as QSARs and fall within a range of techniques known as *in silico* approaches (Puzyn, 2010; Fjodorova and Novič, 2012; Pfannkuch and Suter-Dick, 2015).

QSAR predictions use both statistically based and knowledge-based methods. The statistically based methods rely on techniques such as multivariate analysis, rule induction, cluster analysis, and pattern recognition. They use limited or no prior chemical or biological classification according to mechanism of carcinogenicity. The knowledge-based (or rule-based) methods include toxicological knowledge, expert judgment, and fuzzy logic which considers toxicokinetics, toxicodynamics, and metabolism related to processes with cellular macromolecules or receptors (Serafimova et al., 2010).

10.4 GENOTOXICITY

According to EMA and ICH guidelines on genotoxic impurities, substances which show “alerting structure” in terms of genotoxicity which are not shared with the active substance should be considered and discussed (CHMP, 2008; EMEA, 2006). EMA states that the absence of a structural alert based on a well-performed assessment will be sufficient to conclude that the impurity is of no concern with respect to genotoxicity and no further “qualification” studies or justification will be required. When an impurity is found above the ICH identification threshold, but below the qualification threshold, and the structure gives rise to a structural alert, this can be negated by carrying out an Ames test on the active ingredient containing the impurity as long as the impurity is present at a minimum concentration of 250 mg plate⁻¹.

10.4.1 QSAR for Mutagenicity

As part of a hazard assessment, database and literature searches should be performed for carcinogenicity and bacterial mutagenicity data. To limit a possible human cancer risk associated with the exposure to potentially mutagenic impurities, the bacterial mutagenicity assay is used to assess met mutagenic potential and the need for controls. Other types of genotoxicants that are nonmutagenic typically have threshold mechanisms and usually do not pose carcinogenic risk in humans at the level ordinarily present as impurities (ICH M7).

One simple and well-known approach to predict genotoxicity and carcinogenicity for chemicals is based on the use of structure alerts (SAs). SAs for carcinogenicity are molecular functional groups or substructures known to be linked to the carcinogenic activity of chemicals. Benigni and Bossa authored one of the most recent and complete lists of SAs, 33 in all (Benigni, 2000; Benigni et al., 2008).

TABLE 10.5 ICH M7 Classification of Genotoxic/Mutagenic Risk of Impurity

| | |
|---------|--|
| Class 1 | Known mutagenic carcinogen |
| Class 2 | Known mutagen with unknown carcinogenic potential |
| Class 3 | Alerting structure, unrelated to the structure of the drug substance; no mutagenicity data |
| Class 4 | Alerting structure, same alert in drug substance which has been tested and is nonmutagenic |
| Class 5 | No structural alerts, or alerting structure with sufficient data to demonstrate lack of mutagenicity |

Current regulatory guidance on mutagenic impurities in pharmaceuticals states that the absence of structural alerts from two complementary QSAR methodologies, one expert or rule based and one statistical based, is sufficient to conclude that an impurity is of no concern and no further testing is required (ICH, 2014). This guideline is not intended to be used for leachables and extractables but the idea can also be extended to cover these compounds as a first step in risk management. The ICH M7 guideline calls for using QSAR results and any available associated data to classify impurities as belonging to one of five categories, as summarized in Table 10.5.

If a compound/impurity is predicted to be mutagenic, then further testing may be necessary in order to determine the safety of the drug substance or product. If the compound does not elicit a positive prediction for mutagenicity, then further steps should be taken to determine its safety.

When carcinogenic and mutagenic structural alerts are present for an impurity, it is necessary that the levels of this impurity be kept to less than 1.5 mg daily intake, an amount that poses less than one in 100 000 risks to patients. Exposure at or above this level may require additional testing in the appropriate species (Ames *in vitro* mutagenicity would be quite effective).

When more than one genotoxic impurity is present in the drug substance, the relevant TTC can be applied to each individual impurity if the impurities are structurally unrelated.

Three additional specific areas of potential toxicity in drug development are (i) adverse immune system responses (sensitization and anaphylaxis, though only the former—as dermal sensitization) is addressed in current QSAR models; (ii) hepatic toxicity (frequently called delayed liver toxicity or DLT); and (iii) cardiotoxicity (both arrhythmias and significant changes in blood pressure). Some of both the expert and statistical-based system address these, as summarized in Table 10.6.

10.4.1.1 Sensitization There are four classes of hypersensitivity responses in the Coombs and Gell scale, but two of these (I, delayed contact hypersensitivity, usually a dermal response, and IV, most seriously expressed as anaphylaxis) are generally significant concerns.

Derek, Leadscope modeler, and TOPKAT predict the potential type I responses well, but no model currently predicts the type IV.

10.4.1.2 Hepatotoxicity For the last two decades, post-marketing occurrences of DLT have been the single most common cause of postmarket drug withdrawal due to safety—a wider range of available commercial models predict the potential for these responses moderately well, which given the widespread belief that such are idiosyncratic and not predicted well is a promising situation.

10.4.1.3 Cardiotoxicity Cardiotoxicity is second only to hepatic toxicity as a cause of safety concern withdrawal of drugs post market. While the concern is usually expressed as being due to potential arrhythmia (many associated with disruption of the hERG channel), this is the case less than half the time. hERG channel effects are, however, what QSAR models currently seek to predict.

10.5 COMPARISON OF AVAILABLE MODELS/APPLICATIONS

As a start it should be noted that the current available toxicity prediction programs have two areas of limitation—not currently differentiated stereoisomers and all struggle with predictions involving organometallic ions due to issues associated with metal ions having multiple possible valence states (Walker et al., 2012).

10.5.1 QSAR of Metabolism

As predictive use of QSAR for toxicology end points has matured, the prediction of metabolism is an area of rapidly increasing sophistication (Kirchmair et al., 2015).

10.5.2 Meteor

Meteor is a computer program that helps scientists who need information about the metabolic fate of chemicals and want to be more efficient, be more effective, and make better decisions.

The program uses expert knowledge rules in metabolism to predict the metabolic fate of chemicals, and the predictions are presented in metabolic trees. The only information needed by the program to make its prediction is the molecular structure of the chemical.

Meteor provides comments and literature citations as evidence to support its predictions and is used by scientists with an interest in understanding the metabolic fate of chemicals. Meteor can be used:

- To provide information for decision making when there is little or no experimental metabolism data available
- To help interpret data from mass spectrometry metabolism studies more speedily and accurately

TABLE 10.6 Available QSAR Predictive Modeling Systems

| Type | Derek | Leadscope Modeler | Multicase ^a | EPA TEST | TOPKAT | VEGA (REACH Tool) | OECD Toolbox | ToxTree | CAESAR | Read Across |
|-------------------------------|-------------|---|------------------------|----------|----------------------|--------------------------------------|-----------------|-------------------------------|--------|-------------|
| | Expert rule | Statistical | Statistical | | Statistical | | | Expert rule classification | | Expert rule |
| Notes | X | Gives probability. Each end point (“suite”) billed separately | No | | Predicts probability | Combination of CAESAR and TEST | | No | | |
| Predicts | | | | | | | | | | |
| Mutagenicity | X | X | X | X | X | X | | No | | |
| Clastogenicity | X | X | X | No | X | No | | No | | |
| Carcinogenicity | X | X | X | No | X | X | | No | X | |
| Sensitivity | X | X | No | No | X | No | | No | | |
| Hepatotoxicity | X | X | X | No | No | No | | No | | |
| Developmental Tox | X | X | X | X | X | X | | No | X | |
| Rat LD ₅₀ | No | No | No | X | X | X | | No | | |
| Metabolism (pathways) | No | No | No | No | No | No | | No | | |
| Explains (ID’s mechanisms) | X | No | No | No | No | No | | No | | |

^aMulticase now offers Case ULTRA, an expert rule system.

Meteor also:

- Predicts the metabolic fate of chemicals, displays results as a metabolic tree, and allows you to filter results and see only “likely” metabolites
- Covers phase I and phase II biotransformations
- Links directly to MetaboLynx to speed up analysis of mass spectrometry data from metabolism studies
- Evaluates the metabolic fate of potential new chemical products, existing products, and impurities to identify potentially harmful metabolites
- Makes decisions about which chemicals are likely to have “more favorable” metabolic profiles when you do not have all the experimental information you would like to have about the metabolism of each chemical
- Carries out metabolism experiments where you could save time by knowing in advance what metabolites are likely to be observed
- Improves the properties of a chemical in the R&D pipeline by slightly redesigning its molecular structure

Meteor is used to provide information about the metabolic fate of a chemical when there is no experimental data available and to help analyze mass spectrometry data from metabolism studies:

- Meteor can be used to suggest the most likely molecular structure of a metabolite, or group of metabolites, when only empirical formula data is available.

Within Meteor, you can predict the likely chemical structures and metabolic pathways of metabolites detected by a mass spectrometer. Simply input the masses or formulae of your identified metabolites along with the molecular structure of the parent compound to obtain a Meteor prediction.

Meteor bases its decisions on metabolism on a mechanism very similar to Derek:

- Substructures of the molecule that are the known or suspected points of metabolism, enzymatic or chemical action, or conjugation are identified.
- The likelihood of the reaction taking place is estimated and metabolites or resulting compounds are identified.
- The resulting compounds are then compared under the same rules to generate second order metabolites. This process is repeated up to the level the user specifies.
- The metabolite tree is created, along with an indicator of how likely/prevalent the metabolite is.

Similar to Derek, rules for Meteor are proposed both by the users and experts hired by Lhasa, and the rules are implemented only after being reviewed by a panel of experts.

Therefore, there is research and documentation available for each Meteor prediction. For the same reason, idiosyncratic reactions will not be anticipated unless they have been reported in research, proposed to Lhasa, and reviewed by the Lhasa experts.

It is always recommended that a thorough search of available data on a compound is performed and to remember that data itself, when good, is always better than a prediction.

10.5.3 Derek

Derek is the premier expert rule-based QSAR model system, and both cover a wide (and expanding) range of end points and undergo regular updates. Current revision predicts the probability of an end point (called an alert) being operative in a classification rating scale, shown in Table 10.7. “Negatives”—no alerts—are called (Lhasa, 2015).

Derek predicts genotoxicity based on substructure toxicity and a complex structure of decision rules. It will predict a positive if a particular substructure in the compound has reported genotoxic effects. The individual substructures in Derek predictions are chosen based on a review by a panel of experts, and the results are not numerical but instead are hierarchical: probable is more serious than plausible, plausible is more serious than doubted, etc. The identification of “alerting” structures adds a valuable aspect to interpreting the results. Lhasa, Ltd., the developer of Derek, does now also offer a complementary statistical-based system so that the two-method requirement for impurity evaluation under ICH M7 is met.

Pricing for Derek evaluations is for assessment (all end points), unlike Leadscape and multiple computer-automated structural evaluation (Multicase) where each end point has an additional cost.

TABLE 10.7 Terminology Used in the Outputs Provided by Derek for Each End Point

| Terminology | Description |
|--------------|--|
| Certain | There is proof that the proposition is true |
| Probable | There is at least one strong argument that the proposition is true and there are no arguments against it |
| Plausible | The weight of evidence supports the proposition |
| Equivocal | There is an equal weight for and against the proposition |
| Doubted | The weight of evidence opposes the proposition |
| Improbable | There is at least one strong argument that the proposition is false and there are no arguments that it is true |
| Impossible | There is proof that the proposition is false |
| Open | There is no evidence that supports or opposes that proposition |
| Contradicted | There is proof both that the proposition is true and that it is false |

10.5.4 Leadscope

Leadscope rose of the CAS/RTEC database and performs a nonlinear multiple regression against a group of agents, some classified as genotoxic and some classified as nongenotoxic. The regression is based on structural similarity but does not use specific substructures. The known agents were selected for their predictive ability against a larger database of agents whose genotoxicity is known (Leadscope, 2015).

The genetic toxicity modeler provides either a probability of an alert being relevant or an error message of “not in domain” for evaluated structures. If the prediction indicates “not in domain” compound was not within the domain of applicability (compounds/structures currently used in the prediction model) and no prediction can thus be calculated. The domain of applicability determines whether the test compound can be predicted by the model. If the test structure is not at least 30% similar to one of the training set compounds using a fingerprint of Leadscope structural features and at least one model feature is present in the test structure, it is not in the domain. Most use a criterion of 60% in taking an alert as relevant.

10.5.4.1 Multiple Computer-Automated Structural Evaluation Multicase arose out of an effort, funded by EPA, by Klopman (1992) and Klopman and Rosenkranz (1994) to develop a QSAR model for prediction of potential genotoxicity and carcinogenicity associated with chemical structures. It uses a multivariate model to predict proximity to positive structures, expressed as a probability of an alert being relevant (here a value of 0.6 is usually employed). The system has (or has added) additional end points beyond the original two (mutagenicity and carcinogenicity). Structures need to be submitted as Mole files. FDA uses a modified version. Structures that are organometallic or with molecular weights above 5000Da cannot be evaluated by Multicase.

10.5.4.2 Toxicity Prediction by Computer-Assisted Technology Developed by Kurt Enslein in the 1970s, (Enslein and Craig, 1982; Enslein et al., 1983, 1984; Enslein, 1984) TOPKAT is a hybrid between a traditional system using calculations based on physical chemical aspects of structures and a multivariate statistical system. Unlike Derek, Leadscope, and Multicase, in addition to mutagenicity, it also predicts rat oral LD₅₀ values, developmental toxicity, skin sensitization, skin and ocular irritation, octanol–water partition coefficients, and aerobic biodegradability. It is a PC-based program owned by Accelrys since 2001. Structures must be entered in simplified molecular-input line-entry system (SMILES—Weininger, 1988).

Cramer Rule (ToxTree) Also an older system, originally intended to help prioritize chemical for testing, ToxTree is not actually a QSAR system in that it does not predict the presence or absence of potential activity for specific end points. Rather, it classifies chemicals (based on the presence

or absence of substructural features) as to relative toxicity, ranging from class I (highly toxic) to class III (practically nontoxic) (Cramer et al., 1978; Curios-IT, 2009).

Toxicity Estimation Software Tool TEST was developed by EPA (EPA, 2011). While it predicts end points of concern for candidate drugs (Ames mutagenicity, rat oral LD₅₀, and developmental toxicity), many of its end points (such as 96 h fathead minnow LC₅₀) are more of environmental concern.

10.5.5 VEGA

The model provides a qualitative prediction of mutagenicity on *Salmonella typhimurium* (Ames test). VEGA combines QSAR models and read across tools. The models have been taken from CAESAR or TEST or have been developed later by the contributors to VEGA.

A completely independent algorithm is at the basis of the evaluation for read across. This algorithm shows similar compounds, assesses the QSAR results on the similar compounds, and analyzes some relevant chemical features in the target compound and its related compounds. The QSAR models available for regulatory purposes have been optimized in accordance with the REACH requirements.

The AD of predictions is assessed using an Allowable Daily Intake (ADI) that has values from 0 (worst case) to 1 (best case). The ADI is calculated by grouping several other indices, each one taking into account a particular issue of the AD.

Most of the indices are based on the calculation of the most similar compounds found in the training and test set of the model, calculated by a similarity index that considers molecule's fingerprint and structural aspects (count of atoms, rings, and relevant fragments).

10.5.5.1 Global AD Index The global index takes into account all other indices in order to give a general global assessment on the AD for the predicted compound. Defined intervals are:

$1 \geq \text{index} \geq 0.9$ predicted substance is into the AD of the model.

$0.9 > \text{index} \geq 0.7$ predicted substance could be out of the AD of the model.

$\text{index} < 0.7$ predicted substance is out of the AD of the model.

In the following, all AD components are reported along with their explanation and the intervals used.

10.5.5.2 Similar Molecules with Known Experimental Value This index takes into account how similar are the first three most similar compounds found. Values near 1 mean that the predicted compound is well represented in the data set used to build the model; otherwise the prediction could be an extrapolation.

10.5.5.3 Accuracy of Prediction for Similar Molecules

This index takes into account the error in prediction for the three most similar compounds found. Values near 0 mean that the predicted compound falls in an area of the model's space where the model gives reliable predictions.

10.5.5.4 Concordance for Similar Molecules This index takes into account the difference between the predicted value and the experimental values of the three most similar compounds. Values near 0 mean that the prediction made disagrees with the values found in the model's space; thus the prediction could be unreliable.

10.5.5.5 Atom-Centered Fragments Similarity Check

This index takes into account the presence of one or more fragments that aren't found in the training set or that are rare fragments.

10.5.5.6 Model Descriptors Range Check This index checks if the descriptors calculated for the predicted compound are inside the range of descriptors of the training and test set. The index has value 1 if all descriptors are inside the range, 0 if at least one descriptor is out of the range.

The model goes through a first step in which a set of 12 SAs related to mutagenicity is checked. If one or more fragments are found, compound is predicted as "mutagen."

Statistics obtained applying the model to its original data set:

- Training set: $n=3275$; accuracy=0.92; specificity = 0.87; sensitivity=0.96
- Test set: $n=805$; accuracy=0.82; specificity=0.74; sensitivity=0.89

Furthermore, the statistics for the test set considering the ADI is here reported; the ADI is used, as in the final model's assessment, in order to divide results in three groups (into AD, possibly out of AD, out of AD), showing that compounds considered into AD have better performance than the others:

- Test set with ADI greater than 0.9 (compounds into the AD):
 $n=338$; accuracy=0.94; specificity=0.85; sensitivity = 0.98
- Test set with ADI between 0.9 and 0.7 (compounds could be out of AD):

$n=320$; accuracy=0.82; specificity=0.80; sensitivity = 0.84

- Test set with ADI lower than 0.7 (compounds out of the AD):
 $n=147$; accuracy=0.56; specificity=0.48; sensitivity=0.66

Results given as PDF file consist of a PDF document containing all the information about the prediction. Several pages are dedicated to each molecule, in the following order:

1. *Prediction* Here it is reported a depiction of the compound and the final assessment of the prediction (i.e., the prediction made together with the analysis of the AD). In the following, all information related to the prediction are reported. Note that if some problems were encountered while processing the molecule structure, some warnings are reported in the last field (Remarks).
2. *Similar compounds* Here it is reported the list of the six most similar compounds found in the training and test set of the model, along with their depiction and relevant information (mainly experimental value and predicted value).
3. *Applicability domain* Here it is reported the list of all ADIs, starting with the global ADI. For each index, it is reported its value and a brief explanation of the meaning of that value.
4. *Reasoning on fragments/structural alerts* If some SAs are found, they are reported here (one for each page) with a brief explanation of their meaning and the list of the three most similar compounds that contain the same fragment. Note that if no SAs are found, this section is not shown

TEST allows you to estimate the value for several toxicity end points using several different advanced QSAR methodologies.

A data set of 6512 chemicals was compiled from several different sources. The final data set consists of 5743 chemicals (after removing salts, mixtures, ambiguous compounds, and compounds without CAS numbers).

Ames mutagenicity data set statistical external validation for demonstrating that the consensus method achieved the best prediction accuracy (concordance) and prediction coverage.

The prediction results for the Ames mutagenicity were as shown in Table 10.8.

TABLE 10.8 Prediction Results for Ames Mutagenicity Characteristics

| Method | Concordance | Sensitivity | Specificity | Coverage |
|------------------|-------------|-------------|-------------|----------|
| Hierarchical | 0.765 | 0.775 | 0.753 | 0.951 |
| FDA | 0.771 | 0.771 | 0.771 | 0.962 |
| Nearest neighbor | 0.798 | 0.833 | 0.772 | 0.802 |
| Consensus | 0.807 | 0.817 | 0.775 | 1.000 |

TABLE 10.9 Comparison of MCASE versus Derek versus Leadscope versus TOPKAT—Current State of Predictive Toxicology FDA Purchased LEADSCOPE in 2001

| | Derek | TOPKAT | MCASE | LEADSCOPE |
|------|---|---|--|---|
| Pros | <ul style="list-style-type: none"> • Predictions can be traced back to substructures making further information available (provides structural alerts) • Uses both bacterial mutagenicity data and all other available genotoxic data • Rule-based system • Has suggested permissible limit of distance from the OPS, precalculated and stored within database • Able to process more molecules than TOPKAT • Indicates presence of aromatic amines • Prediction extends to genotoxicity, toxicity, carcinogenicity, developmental and reproductive effects, and local time | <ul style="list-style-type: none"> • Uses electrotopological descriptors rather than chemical structures, atomic size- adjusted E-states computed from rescaled count of valence electrons, molecular weight, topological shape indices, and symmetry indices (extension of QSAR) • Correlative structure–activity relationship (SAR) system | <ul style="list-style-type: none"> • Breaks molecule into fragments (biophores) • Considers physicochemical properties and potentially deactivating fragments • Accounts for aqueous solubility and quantum chemical parameters such as highest occupied molecular orbital and lowest occupied molecular orbital energies • Able to identify the potential of existence of baseline activity due to a specific physical attribute • Can deal with noncongeneric data sets • Uses binary information to discriminate among structural features associated with active and inactive compounds • Focuses on genotoxicity and carcinogenicity | <ul style="list-style-type: none"> • Able to predict nongenotoxic substances • Results numerical • Predicts genotoxicity at different levels, not clear-cut like Derek (either it is or it is not) • Allows for easy visualization of large sets of chemical compounds, their properties, and biological activities • Organizes into functional groups, aromatics, and heterocycles with common substituents (2D histograms (frequencies of each class in data set) and scatter plots) • Dynamically constructs series of detailed cross sections of the data set by expanding more specific structural classes, creating subprojects, and using interactive controls for properties • Free to explore any subset, not limited to the statistically significant feature • Predictions too complex to be traced back for explanations and/or causes • Generates false positives • Does not predict idiosyncratic toxicity well, occasionally reporting false positives. Although it usually reports no prediction possible |
| Cons | <ul style="list-style-type: none"> • Does not specifically call out nongenotoxic substances • Results are hierarchical, so it provides no numerical probability (categorical) • Predicts something or nothing • Can generates false positives, but not false negatives • Does not predict idiosyncratic toxicity unless the toxicity has already been reported or included in database • If molecular weight is too high, it cannot generate genotoxicity predictions unless structure • When processing hundreds of compounds, compound-by-compound evaluation becomes impractical • Overpredicts the mutagenicity of aromatic amines, additional research required to refine the rules • Cannot take advantage of provided training data without programming input • Unknown SARs are not contained in database | <ul style="list-style-type: none"> • Uses only bacterial mutagenicity data, not other genotoxic data • Not useful in predicting the carcinogenicity of chemicals • Database too small and decisions rules not refined enough to accurately predict carcinogenicity • Molecule must be within all dimensions to be acceptable; otherwise, it depends on the distance of the query from the OPS • Cannot process if (i) a charge on a molecule other than anionic oxygen or cationic nitrogen, (ii) compounds with manganese, and (iii) mixtures • Has defined chemical space increasing the changes of “unprocessability” • Cannot process molecules in batch mode making evaluations of large data sets unfeasible • Compound names and structures not searchable or organized alphabetically (finding one structure among hundreds becomes strenuous) • Struggles with predicting mutagenicity of aromatic amines | <ul style="list-style-type: none"> • If substructure not related to observed activity, will be randomly placed among active/inactive compounds in the database • Cannot take advantage of provided training data • Yields lower prediction accuracy when the software cannot adapt to a specific chemical space | <ul style="list-style-type: none"> • Predictions too complex to be traced back for explanations and/or causes • Generates false positives • Does not predict idiosyncratic toxicity well, occasionally reporting false positives. Although it usually reports no prediction possible |

Neither TOPKAT nor Derek can reduce false-positive rates or false-negative rate simultaneously (reducing one allows for more specific use).

The consensus method The predicted toxicity is estimated by taking an average of the predicted toxicities from the hierarchical clustering, FDA, and nearest neighbor QSAR methods (provided the predictions are within the respective ADs):

- *Hierarchical method* The toxicity for a given query compound is estimated using the weighted average of the predictions from several different cluster models.
- *FDA method* The prediction for each test chemical is made using a new model that is fit to the chemicals that are most similar to the test compound
- *Nearest neighbor method* The predicted toxicity is estimated by taking an average of the three chemicals in the training set that are most similar to the test chemical.

The prediction accuracy is then evaluated in terms of the fraction of compounds that are predicted accurately. The prediction accuracy is evaluated in terms of three different statistics: concordance, sensitivity, and specificity.

Concordance is the fraction of all compounds that are predicted correctly (i.e., experimentally active compounds that are predicted to be active and experimentally inactive compounds that are predicted to be inactive).

Sensitivity is the fraction of experimentally active compounds that are predicted to be active, while specificity is the fraction of experimentally inactive compounds that are predicted to be inactive.

10.5.6 Derek versus Leadscope

Please note that Derek and Leadscope Model Applier use very different algorithms. Derek predicts genotoxicity based on substructure toxicity, that is, it will predict a positive if a particular substructure in the compound has reported genotoxic effects. The individual substructures in Derek predictions are chosen based on a review by a panel of experts, and the results are not numerical but instead are hierarchical: probable is more serious than plausible, plausible is more serious than doubted, etc. Leadscope instead performs a nonlinear regression against a group of agents, some classified as genotoxic and some classified as nongenotoxic. The regression is based on structural similarity but does not use specific substructures. The known agents were selected for their predictive ability against a larger database of agents whose genotoxicity is known. Since regression is performed numerically, you get numerical results instead of classifications. Strength of association is based both on the probability that measures a degree of “closeness” to the agents in question and on the number of agents the structure is close to. This leads to a number of differences in results, specifically:

1. Derek will not predict a substance as nongenotoxic while Leadscope will.
2. Similarly, Derek either predicts something or not. Leadscope generates positives (structurally “close” to one or more known genotoxic agents), negatives (structurally close to one or more known nongenotoxic agents), or no prediction possible (not close enough to any agents in the database for prediction).
3. Derek predictions can always be traced back to a particular substructure and further information is available. Leadscope predictions are usually too complex for ready explanation.
4. Both will generate false positives.
5. Neither is good with idiosyncratic toxicities: Derek will not predict idiosyncratic toxicities unless the toxicity has been reported and included in their database. Leadscope usually reports no prediction possible but also has a false-positive rate for idiosyncratic toxicities.

Tables 10.3 and 10.9 contain comparisons between Derek, Leadscope, and other QSAR programs.

REFERENCES

- Abraham DJ, Rotella DP. (2010) *Burger's Medicinal Chemistry, Drug Discovery and Development*, 7th Ed. John Wiley & Sons, Hoboken, NJ.
- Ashby J, Tennant RW. (1994) Prediction of rodent carcinogenicity of 44 chemicals: results. *Mutagenesis*. 9:7–15.
- Asher IM, Zervos C. (1977) *Structural Correlates of Carcinogenesis and Mutagenesis*. Office of Science, FDA, Washington, DC.
- Baba H, Takahara J, Matitsuka H. (2015) In silico predictions of skin permeability using nonlinear quantitative structure property relationship models. *Pharm. Res.* 32:2360–2371.
- Benigni R. (Ed.) (2000) *Quantitative Structure-Activity Relationship (QSAR) Models of Mutagens and Carcinogens*. CRC Press, Washington, DC.
- Benigni R, Bossa C, Jeliaskova N, Netzevac T, Worth A. (2008) *The Benigni/Bossa Rulebase for Mutagenicity and Carcinogenicity—A Module of Toxtree*. European Commission Joint Research Centre, Ispra, Italy.
- Bercu JP, Dobo KL, Gocke E, McGovern TJ. (2009) Overview of genotoxic impurities in pharmaceutical development. *Int. J. Toxicol.* 28(6):468–478.
- Bernstein ME. (1984) Agents affecting the male reproductive system—effects of structure on activity. *Drug Metab. Rev.* 15:941–996.
- Chang R. (1981) *Physical Chemistry with Application to Biological Systems*. MacMillan, New York.
- CHMP Safety Working Party (SWP) (2008) *Questions & Answers on the CHMP Guideline on the Limits of Genotoxic Impurities*. London, June 26, 2008.

- Cohen JL, Lee W, Lien EJ. (1974) Dependence on toxicity on molecular structure: group theory analysis. *J. Pharm. Sci.* 63:1068–1072.
- Cramer GM, Ford RA, Hall RA. (1978) Estimation of toxic hazard—a decision tree approach. *Food Chem. Toxicol.* 16:255–276.
- Cronin MTD, Livingston DJ. (2004) *Predicting Chemical Toxicity and Fate*. CRC Press, Washington, DC.
- Crum Brown A, Fraser TR. (1869) On the connection between chemical constitution and physiological action. Part II. On the physiological action of the ammonium bases derived from atropia and conia. *Trans. R. Soc. Edinb.* 25:693–739.
- Curios-IT. (2009) *Cramer Rules with Extensions*. Curios-IT, Leiden, NE.
- Dupuis G, Benezra C. (1982) *Allergic Contact Dermatitis to Simple Chemicals—A Molecular Approach*. Marcel Dekker, New York.
- Ekins S. (Ed.) (2007) *Computational Toxicology, Risk Assessment for Pharmaceutical and Environmental Chemicals*. Wiley Interscience, Hoboken, NJ.
- EMA (2006) *Guideline on the Limits of Genotoxic Impurities*. London, June 28, 2006.
- Enslein K. (1984) Estimation of toxicological endpoints by structure-activity relationships. *Pharmacol. Rev.* 36:131–134.
- Enslein K, Craig PN. (1982) Carcinogenesis: a predictive structure activity model. *J. Toxicol. Environ. Health.* 10:521–530.
- Enslein K, Lander TR, Tomb ME, Landis WG. (1983) Mutagenicity (Ames): a structure-activity model. *Teratog. Carcinog. Mutagen.* 3:503–514.
- Enslein K, Tomb ME, Lander TR. (1984) Structure-activity models of biological oxygen demand. In: *QSAR in Environmental Toxicology* (Kaiser KLE, Ed.). Reidel, Dordrecht.
- EPA (1994) *Health Assessment Document of 2,3,7,8-Tetrachlorodibenzo-p-dioxin (TCDD) and Related Compounds*. External Review Draft. EPA/600/BP-921001b, Washington, DC.
- EPA (2016) *User's Guide for T.E.S.T. Version 4.0, Toxicity Estimation Software Tool*. United States Environmental Protection Agency, Washington, DC.
- Fjodorova N, Novič M. (2012) Integration of QSAR and SAR methods for the mechanistic interpretation of predictive models for carcinogenicity. *Comput. Struct. Biotechnol. J.* 1(2):e201207003.
- Foye WO, Lemke TL, Williams DA. (1995) *Principles of Medicinal Chemistry*, 4th Ed. Lippincott Williams & Wilkins, Philadelphia, PA.
- Franke R. (1973) Structure-activity relationships in polycyclic aromatic hydrocarbons: induction of microsomal aryl hydrocarbon hydroxylase and its possible importance in chemical carcinogenesis. *Chem. Biol. Interact.* 6:1–17.
- Free SM, Wilson JW. (1964) A mathematical contribution to structure-activity studies. *J. Med. Chem.* 7:395–399.
- Gad SC. (2014) QSAR. In: *Encyclopedia of Toxicology* (Wexler P, Ed.), 3rd Ed. Elsevier, Cambridge, UK.
- Goldberg L. (1983) *Structure-Activity Correlations as a Predictive Tool in Toxicology*. Hemisphere, New York.
- Gund P, Andosf JD, Rhodes JB, Smith GM. (1980) Three-dimensional molecular modeling and drug design. *Science.* 208:1425–1431.
- Hansch C. (1971). In: *Drug Design* (Ariens EJ, Ed.), Vol. I. Academic Press, New York.
- Helma C. (Ed.) (2005) *Predictive Toxicology*. Taylor & Francis, Boca Raton, FL.
- ICH (2006) *Q3B(R2) Impurities in New Drug Products*. International Conference on Harmonisation, Geneva, Switzerland.
- ICH (2014) *M7(R4) Assessment and Control of DNA Reactive (Mutagenic) Impurities in Pharmaceuticals to Limit Potential Carcinogenic Risk*. International Conference on Harmonisation, Geneva, Switzerland.
- Instituto di Ricerche Farmacologiche Mario Negri Milano (2015a) *About QSAR/Read-Across*. Available at: <http://www.vega-qsar.eu/about-qsar.html> (accessed January 10, 2016).
- Instituto di Ricerche Farmacologiche Mario Negri Milano (2015b) *QSAR Regulation and Research*. Available at: <http://www.vega-qsar.eu/regulation.html> (accessed January 10, 2016).
- Kaufman JJ, Lewchenko V, Hariharan PC, Kuski WS. (1983) Theoretical predictions of toxicity. In: *Product Safety Evaluation* (Goldberg AM, Ed.). Mary Ann Liebert, New York, pp. 333–355.
- Kirchmair J, Göller AH, Lang D, Kunze J, Testa B, Wilson ID, Glen RC, Schneider G. (2015) Predicting drug metabolism: experiment and/or computation. *Nat. Rev. Drug Discov.* 14:387–404.
- Klopman G. (1992) MULTICASE: a hierarchical computer automated structure evaluation program. *Quant. Struct.-Act. Relat.* 11:176–184.
- Klopman G, Rosenkranz HS. (1994) Approaches to SAR in carcinogenesis and mutagenesis: prediction of carcinogenicity/mutagenicity using MULTI-CASE. *Mutat. Res.* 305:33–46.
- Kruhlak NL, Benz RD, Zhou H, Colatsky TJ. (2012) (Q)SAR modeling and safety assessment in regulatory review. *Clin. Pharmacol. Ther.* 91(3):529–534.
- Kubinyi H. (1995) The quantitative analysis of structure-activity relationships. In: *Burger's Medicinal Chemistry* (Wolff ME, Ed.), Vol. I. John Wiley & Sons, New York, pp. 495–572.
- Leadscope Hosted Client (2015), *Version 3.2.6-1*, CTC Life Science Corporation, Japan. Available at: <http://www.leadscope.com/index.php> (accessed May 7, 2015).
- Lhasa Limited (2015). *Derek Nexus: 4.1.0, Nexus: 2.0.0*. Leeds, UK. Available at: <http://www.lhasalimited.org/> (accessed April 17, 2015).
- Mackay D, Arnot JA, Petkova EP, Wallace KB, Call DJ, Brooke LT, Veith GD. (2009) The physicochemical basis of QSARs for baseline toxicity. *SAR QSAR Environ. Res.* 20(3–4):393–414.
- Meyer H. (1899) Zur Theorie der Alkoholnarkose. *Arch. Exp. Pathol. Pharmacol.* 42:109–118.
- Niculescu-Duvaz I, Craescu T, Tugulea M, Croisy A, Jacquignon PC. (1981) A quantitative structure-activity analysis of the mutagenic and carcinogenic action of 43 structurally related heterocyclic compounds. *Carcinogenesis.* 2:269–275.
- OECD (2014) *Guidance on Grouping of Chemicals*, Second Edition. Series on Testing and Assessment, No. 194, April 14, 2014, Paris. Available at: <http://www.oecd.org/officialdocuments/publicdisplaydocumentpdf/?cote=env/jm/mono%282014%294&doclanguage=en> (accessed January 10, 2016).

- Olson EC, Cristoffersen RE. (1979) *Computer Assisted Drug Design*. ACS, Washington, DC.
- Overton CE. (1901) *Studien über die Narkose, zugleich ein Beitrag zur allgemeinen Pharmakologie*. Gustav Fischer Verlag, Jena, Switzerland.
- Pfannkuch F, Suter-Dick L. (Eds.) (2015) *Predictive Toxicology: From Vision to Reality*. Wiley-VCH, Weinheim.
- Puzyn T. (Ed.) (2010) *Recent Advances in QSAR Studies Methods and Applications*. Springer, New York.
- Richet MC. (1893) Note sur le rapport entre la toxicité et les propriétés physiques des corps. *Compt. Rend. Soc. Biol.* 45:775–776.
- Roy K, Kar S, Das RN. (2015) *Understanding the Basics of QSAR for Applications in Pharmaceutical Sciences and Risk Assessment*. Academic Press, San Diego, CA.
- Sahigara F, Mansouri K, Ballabio D, Mauri A, Consonni V, Todeschini R. (2012) Comparison of different approaches to define the applicability domain of QSAR models. *Molecules*. 17:4791–4810.
- Seeger-Hartmann T. (2015) In silico toxicology—current approaches and future perspectives to predict toxic effects with computational tools. In: *Predictive Toxicology* (Pfannkuch F and Suter-Dick L, Eds.). Wiley-VCH, Berlin, pp. 11–32.
- Selassie CR, Verma RP. (2010) History of quantitative structure-activity relationships. In: *Burger's Medicinal Chemistry, Drug Discovery and Development* (Abraham DJ and Potella DP, Eds.), 7th Ed. John Wiley & Sons, Inc., Hoboken, NJ.
- Serafimova R, Fuat Gatnik M, Worth A. (2010) *Review of QSAR Models and Software Tools for Predicting of Genotoxicity and Carcinogenicity: JRC Scientific and Technical Reports*. European Commission Joint Research Centre, Ispra, Italy.
- Sugai S, Murata K, Kitagahi T, Tomita I. (1990) Studies on eye irritation caused by chemicals in rabbits, I: a quantitative structure-activity relationship approach to primary eye irritations of chemicals in rabbits. *J. Toxicol. Soc.* 15:245–262.
- Tamura RM. (1983) A model for toxicologic prediction. In: *Product Safety Evaluation* (Goldberg AM, Ed.). Mary Ann Liebert, New York, pp. 317–330.
- Thomas L, Lemke TL, Williams DA, Roche VF, Zito SW. (Eds.) (2012) *Foye's Principles of Medicinal Chemistry*, 7th Ed. Philadelphia, Lippincott Williams & Wilkins.
- Topliss JG. (1983) *Quantitative Structure-Activity Relationships of Drugs*. Academic Press, New York.
- Tute MS. (1983) Mathematical modeling. In: *Animals and Alternatives in Toxicity Testing* (Balls M, Riddell RJ and Worden AN, Eds.). Academic Press, New York, pp. 137–166.
- Venkatapathy R, Wang CY, Martin TM, Harten PF, Young D. (2009) Structure-activity relationships for carcinogenic potential. In: *General and Applied Toxicology* (Ballantyne B, Marre TC and Syversen T, Eds.), 3rd Ed, Vol. 4. John Wiley & Sons, Ltd, Chichester, UK, pp. 1771–1803.
- Walker JD, Newman MC, Enache M. (2012) *Fundamental QSARs for Metal Ions*. CRC Press, Boca Raton, FL.
- Weininger DJ. (1988) SMILES, a chemical language for information systems: I. Introduction to methodology and encoding rules. *J. Chem. Inf. Comput. Sci.* 28:31–36.
- Wexler P, Hakkinen PJ, Kennedy G, Stoss FW. (1999) *Information Resources in Toxicology*, 3rd Ed. Academic Press, San Diego, CA.

IMMUNOTOXICOLOGY IN DRUG DEVELOPMENT

11.1 INTRODUCTION

Unlike many of the other areas covered in this volume, immunotoxicity evaluation of drugs has undergone fundamental changes since the first edition. Before, there was no specific guidance. Now, with ICH S8A in place and in force, there are both a requirement and a road map, though biologics are excluded from coverage.

This is in particular a concern as biologics have attained their therapeutic promise with their hyperpharmacology providing the largest portion of safety concern for them (from the “cytokine storm” of TGN-1412 to the unintended immune issues associated with most mAbs). With immune modulation of some form being the primary intended therapeutic effects and with one-third of all new approved therapeutics being biologics, this is particularly a concern, while specifically not covered by ICH guidelines approaches and requirements for immunotoxicity evaluations of these moieties are addressed, however, in this chapter.

All three ICH regions have made strong efforts to harmonize the immunotoxicity risk assessment for investigational new drugs. These efforts culminated in the release of the ICH S8 guideline, which was adopted by the CHMP in October 2005 and came into force in the United States in April 2006, as well as the MHLW in April 2006. According to this current guideline, initial immunotoxicity assessment should be based on the evaluation of data already available from standard toxicity studies and other characteristics of the drug substance like pharmacological properties, the intended patient population, known drug class effects, clinical data for the drug, and its disposition. The need for additional immunotoxicity testing should be decided on the basis of a weight of evidence assessment, taking into account

all available information. Testing should thus be a tiered approach, triggered and determined by concerns from the weight of evidence assessment.

Immunotoxicity is defined in the ICH S8 specifically guideline as unintended immunosuppression or enhancement. It must be noticed that drug-induced hypersensitivity and autoimmunity are not in the scope of ICH S8. Surprisingly, the ICH S8 guideline excludes biological and biotechnology-derived products, probably the larger area of concern for pharmaceuticals, which are now the bigger area of concern.

The immune system is a highly complex system of organ systems, cells, and soluble factors distributed throughout the body and involved in a multitude of functions including antigen presentation and recognition, amplification, and cell proliferation with subsequent differentiation and secretion of lymphokines and antibodies. These are, in health, in a state of balance, and there are extensive mechanisms to maintain this balance. The resulting integrated system is responsible for defense against foreign pathogens and spontaneously occurring neoplasms and is readily triggered to response. To be effective, the immune system must be able to both recognize and destroy foreign antigens. To accomplish this, cellular and soluble components of diverse function and specificity circulate through blood and lymphatic vessels, thus allowing them to act at remote sites and tissues. For this system to function in balance and harmony requires regulation through cell-to-cell communications and precise recognition of self versus nonself. There are multiple opportunities for immunotoxicants to upset this balance by selectively disabling one or more of the cell types or alter membrane morphology and receptors. There are several undesired immune system responses that may occur upon repeated

therapeutic administration of a pharmaceutical that may ultimately present barriers to its development, including:

- Down-modulation of the immune response (immunosuppression), which may result in an impaired ability to deal with neoplasia and infections. This is of particular concern if the therapeutic agent is intended to be used in patients with preexisting conditions such as cancer, severe infection, or immunodeficiency diseases.
- Up-modulation of the immune system (i.e., autoimmunity).
- Direct adverse immune responses to the agent itself in the form of hypersensitivity responses (anaphylaxis and delayed contact hypersensitivity).
- Direct immune responses to the agent that limit or nullify its efficacy (i.e., the development of neutralizing antibodies).

Immune-modulated responses to drugs (“drug allergies”) are a major problem and cause of discontinuance of use by patients who do need access to the therapeutic benefits (Patterson et al., 1986), and there remains no adequate pre-clinical methodology for identifying/predicting these responses to orally administered small molecule drugs (Hastings, 2001).

It is the intent of this chapter to provide an understanding of these adverse immunological effects, the types of preclinical tests that may be used to detect them, and approaches for testing and interpreting test results.

Immunotoxicology has evolved over the last 30 years as a specialty within toxicology that brings together knowledge from basic immunology, molecular biology, microbiology, pharmacology, and physiology. As a discipline, immunotoxicology involves the study of adverse effects that xenobiotics have on the immune system. As listed earlier, several different types of adverse immunological effects may occur, including immunosuppression, autoimmunity, and hypersensitivity. Although these effects are clearly distinct, they are not mutually exclusive. For example, immunosuppressive drugs that suppress suppressor cell activity can also induce autoimmunity (Hutchings et al., 1985), and drugs that are immunoenhancing at low doses may be immunotoxic at high doses. Chemical xenobiotics may be in the form of natural or man-made environmental chemicals—pharmaceuticals and biologicals that are pharmacologically, endocrinologically, or toxicologically active. Although, in general, xenobiotics are not endogenously produced, immunologically active biological response modifiers that naturally occur in the body should also be included, since many are not known to compromise immune function when administered in pharmacologically effective doses (Koller, 1987).

Although the types of immunological responses to various xenobiotics may be similar, the approach taken for screening potential immunological activity will vary depending on the

application of the compound. Thus, this chapter will primarily focus on the immunotoxicology of pharmaceuticals. In contrast to potential environmental exposures, pharmaceuticals are developed with intentional but restricted human exposure and their biological effects are extensively studied in surveillance. Pharmaceuticals are developed to be biologically active and, in some cases, intentionally immunomodulating or immunosuppressive. Many will react with biological macromolecules or require receptor binding in order to be pharmacologically active. By their nature, these interactions may result in toxicity to the cells of the immune system, may adversely alter the appearance of “self” to produce an autoimmune response, or may form a hapten, which may then elicit a hypersensitivity response. Because of the fast-expanding development of new drugs that can potentially impact the immune responsiveness of humans, immunotoxicity testing of new pharmaceutical products has become a growing concern.

Until recently, immunotoxicology in pharmaceutical safety assessment has been poorly addressed both by regulatory requirements/guidelines and by existing practice. Notable exceptions are the testing requirements for delayed contact hypersensitivity for dermally administered agents and antigenicity/anaphylaxis testing for drugs to be registered in Japan. The most recently announced regulatory expectation for parenterally administered protein or peptide agents produced by biotechnology is that the development of antibodies (neutralizing and otherwise) should be evaluated in at least one (preferably two) of the animal models used to assess general systemic toxicity.

Unanticipated immunotoxicity is infrequently observed with drugs that have been approved for marketing. With the exception of drugs that are intended to be immunomodulatory or immunosuppressive as part of their therapeutic mode of action, there is little evidence that drugs cause unintended functional immunosuppression in man (Gleichmann et al., 1989). However, hypersensitivity (allergy) and autoimmunity are frequently observed and are serious consequences of some drug therapies (DeSwarte, 1986; Patterson et al., 1986; Choquet-Kastylevsky et al., 2001; Pieters, 2001). An adverse immune response in the form of hypersensitivity is one of the most frequent safety causes for withdrawal of drugs that have already made it to the market (see Table 11.1) and accounts for approximately 15% of adverse reactions to xenobiotics (deWeck, 1983; Bakke et al., 1984). In addition, adverse immune responses such as this (usually urticaria and frank rashes) are the chief “unexpected” finding in clinical studies. These findings are unexpected in that they are not predicted by preclinical studies because there is a lack of good preclinical models for predicting systemic hypersensitivity responses, especially to orally administered agents. As a consequence, the unexpected occurrence of hypersensitivity in the clinic may delay, or even preclude, further development and commercialization. Thus, a primary purpose for preclinical

TABLE 11.1 Drugs Withdrawn from the Market Due to Dose- and Time-Unrelated Toxicity Not Identified in Animal Experiments

| Compound | Adverse Reaction | Year of Introduction | Years on Market |
|------------------|----------------------------------|----------------------|-----------------|
| Aminopyrine | Agranulocytosis | ca. 1900 | 75 |
| Phenacetin | Interstitial nephritis | ca. 1900 | 83 |
| Dipyrone | Agranulocytosis | ca. 1930 | 47 |
| Clioquinol | Subacute myelo-optic neuropathy | ca. 1930 | 51 |
| Oxyphenisatin | Chronic active hepatitis | ca. 1955 | 23 |
| Nialamide | Liver damage | 1959 | 19 |
| Phenoxypropazine | Liver damage | 1961 | 5 |
| Mebanazine | Liver damage | 1963 | 3 |
| Ibufenac | Hepatotoxicity | 1966 | 2 |
| Practolol | Oculomucocutaneous syndrome | 1970 | 6 |
| Alcolofenace | Hypersensitivity | 1972 | 7 |
| Azaribine | Thrombosis | 1975 | 1 |
| Ticrynafen | Nephropathy | 1979 | 1 |
| Benoxaprofen | Photosensitivity, hepatotoxicity | 1980 | 2 |
| Zomepirac | Urticaria, anaphylactic shock | 1980 | 3 |
| Zirnelidine | Hepatotoxicity | 1982 | 2 |
| Temafloxacin | Hepato- and renal toxicity | 1990 | 2 |
| Tronan | Hepato- and renal toxicity | 1997 | 3 |
| Renzalin | Hepatotoxicity | 1996 | 4 |

Source: Adapted from Bakke et al. (1984, pp. 559–567).

immunotoxicology testing is to help us detect these adverse effects earlier in development, before they are found in clinical trials (Burns-Naas and Pallardy, 2013).

11.2 OVERVIEW OF THE IMMUNE SYSTEM

A thorough review of the immune system is not the intent of this chapter, but a brief description of the important components of the system and their interactions is necessary for an understanding of how xenobiotics can affect immune function. A breakdown at any point in this intricate and dynamic system can lead to immunopathology.

The immune system is divided into two defense mechanisms: nonspecific, or innate, and specific, or adaptive, mechanisms that recognize and respond to foreign substances. Some of the important cellular components of nonspecific and specific immunity are described in Table 11.2. The nonspecific immune system is the first line of defense against infectious organisms. Its cellular components are the phagocytic cells such as the monocytes, macrophages, and polymorphic neutrophils (PMNs).

The innate immune system has been identified in immune defense in insects and alone in the phylogenetic sequence and serves the same purpose in all these organisms, including humans. The innate immune system encompasses all physical, chemical, and cellular barriers that protect the individual from microbial infections without the need to learn to discriminate self from nonself. The body protects itself from dangerous actions of the innate immune system by the lack of expression of molecular patterns of microor-

ganisms and by the abundant expression of inhibitors. Most components of the innate immune system can work independently and in parallel to destroy microorganisms. Due to its redundancy, the innate immune system is rather robust regarding its actions.

The specific, or adaptive, immune system is present only in vertebrates and is characterized by memory, specificity, and the ability to distinguish “self” from “nonself.” The important cells of the adaptive immune system are the lymphocytes and antigen-presenting cells (APCs) that are part of nonspecific immunity. The lymphocytes, which originate from pluripotent stem cells located in the hematopoietic tissues of the liver (fetal) and bone marrow, are composed of two general cell types: T and B cells. The T cells differentiate in the thymus and are made up of three subsets: helper, suppressor, and cytotoxic. The B cells, which have the capacity to produce antibodies, differentiate in the bone marrow or fetal liver. The various functions of the T cells include presenting antigen to B cells, helping B cells to make antibody, killing virally infected cells, regulating the level of the immune response, and stimulating cytotoxic activity of other cells such as macrophages (Male et al., 2012).

Activation of the immune system is thought to occur when APCs such as macrophages and dendritic cells take up antigen via F_c or complement receptors, process the antigen, and present it to T cells (see Figure 11.1). Macrophages release soluble mediators such as interleukin-1 (IL-1), which stimulate T cells to proliferate. APCs must present antigen to T cells in conjunction with the class II major histocompatibility complex (MHC) proteins that are located on the surfaces

TABLE 11.2 Cellular Components of the Immune System and Their Functions

| Cell Subpopulations | Markers ^a | Functions |
|---|----------------------|---|
| <i>Nonspecific immunity</i> | | |
| Granulocytes | | Degranulate to release mediators |
| Neutrophils (blood) | | |
| Basophils (blood) | — | |
| Eosinophils (blood) | | |
| Mast cells (connective tissue) | | |
| Natural killer (NK) cells | — | Nonsensitized lymphocytes; directly kill target cells |
| Reticuloendothelial | CD14; HLA-DR | Antigen processing, presentation, and phagocytosis (humoral and some cell-mediated responses) |
| Macrophage (peritoneal, pleural, alveolar spaces) | | |
| Histiocytes (tissues) | | |
| Monocytes (blood) | | |
| <i>Specific immunity</i> | | |
| Humoral immunity | CD19; CD23 | Proliferate; form plasma cells |
| Activated B cells | — | Secrete antibody; terminally differentiated |
| Plasma cells | | |
| Resting | — | Secrete IgM antibodies (primary response) |
| Memory | — | Secrete IgG antibodies (secondary response) |
| <i>Cell-mediated immunity</i> | | |
| T-cell types | | |
| Helper (T _h) | CD4; CD25 | Assists in humoral immunity; required for antibody production |
| Cytotoxic (T _k) | CD8; CD25 | Targets lysis |
| Suppressor (T _s) | CD8; CD25 | Suppresses/regulates humoral and cell-mediated responses |

^aActivation surface markers detected by specific monoclonal antibodies; can be assayed with flow cytometry.

of T cells. The receptor on the T cell is a complex of the Ti molecule that binds antigen, the MHC proteins, and the T3 molecular complex, which is often referred to as the CD3 complex. Upon stimulation, T cells proliferate, differentiate, and express interleukin-2 (IL-2) receptors. T cells also produce and secrete IL-2, which, in turn, acts on antigen-specific B cells, causing them to proliferate and differentiate into antibody-forming (plasma) cells.

The adaptive arm of the immune system which depends on delicate tolerance mechanisms to shape the repertoire of the randomly produced antigen receptors makes this part of the immune system susceptible for the induction of autoimmunity. In contrast to innate immune responses, induction of an adaptive immune response depends on a well-regulated temporal and spatial sequence of activation events, which requires the cooperation of several cell types, and the migration of cells between different organs. As a result, activation of adaptive immune responses is a rather slow process, as compared with activation of innate defense mechanisms. All cellular interactions during the induction of an adaptive immune response are based on differentiated intercellular communications pathways involving a host of cytokines and surface receptors.

Antibodies circulate freely in the blood or lymph and are important in neutralizing foreign antigens. The various types

of antibodies involved in humoral immunity and their functions are described in Table 11.3. There are multiple genes (polymorphisms) that encode diversity to the variable region of the antibody. B cells are capable of generating further diversity to antibody specificity by a sequence of molecular events involving somatic mutations, chromosomal rearrangements during mitosis, and recombination of gene segments (Roitt et al., 1985; Ladics, 2005).

The immune system is regulated in part by feedback inhibition involving complex interactions between the various growth and differentiation factors listed in Table 11.4. Since antigen initiates the signal for the immune response, elimination of antigen will decrease further stimulation (Male et al., 2012). T suppressor cells (T_s) also regulate the immune response and are thought to be important in the development of tolerance to self antigens. In addition to the humoral immune system or the branch that is modulated by antibody, cell-mediated immunity and cytotoxic cell types play a major role in the defense against virally infected cells, tumor cells, and cells of foreign tissue transplants. Cytotoxic T_k cells (T killer cells) recognize antigen in association with class I molecules of the MHC, while natural killer (NK) cells are not MHC restricted. Cell killing results in a sequence of events following activation of the effector cell, lysosomal degranulation, and calcium influx into the targeted cell. The

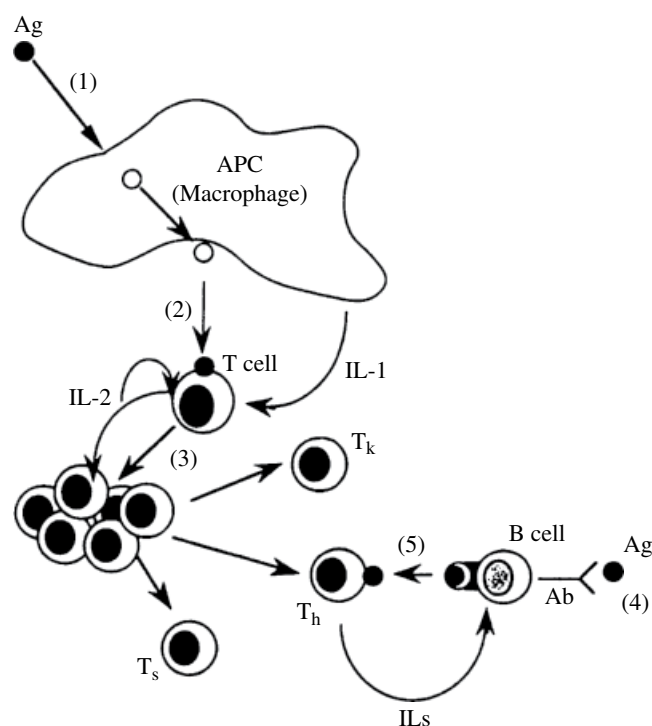


FIGURE 11.1 Simplified schematic of immunoregulatory circuit that regulates activation of T cells and B cells involved in humoral (T-cell-dependent) and cell-mediated immunity. (1) Antigen (Ag) is processed by the APCs expressing class II MHC molecules. (2) Antigen plus class II MHC is then presented to antigen-specific T helper cells ($CD4^+$), which stimulates secretion of IL-2. (3) IL-2 in turn stimulates proliferation (clonal expansion) of T cells and differentiation into T suppressor (T_s), T killer (T_k), and T helper (T_h) effector cells. The expanded clone has a higher likelihood of finding the appropriate B cell that has the same antigen and class II molecules on its surface. (4) Next, the antigen binds to an antibody (Ab) on the surface of a specific B cell. (5) The B cell in turn processes the antigen and presents it (plus class II MHC) to the specific T_h cell. The T_h cell is then stimulated to secrete additional ILs that stimulate clonal expansion and differentiation of the antigen-specific B cell.

various types of cells involved in cell-mediated cytotoxicity and their mechanisms of action are outlined in Table 11.5.

11.3 IMMUNOTOXIC EFFECTS

The immune system is a highly integrated and regulated network of cell types that requires continual renewal to achieve balance and immunocompetence. The delicacy of this balance makes the immune system a natural target for cytotoxic drugs or their metabolites. Since renewal is dependent on the ability of cells to proliferate and differentiate, exposure to agents that arrest cell division can subsequently lead to reduced immune function or immunosuppression. This concept has been exploited in the development of therapeutic drugs intended to treat leukemias, autoimmune diseases, and chronic inflammatory diseases and to prevent transplant rejection (Brunton et al., 2006). However, some drugs may adversely modulate the immune system secondarily to their therapeutic effects.

Two broad categories of immunotoxicity have been defined on the basis of suppression or stimulation of normal immune function (Burlison et al., 1995; Luebke et al., 2007; Burns-Naas and Pallardy, 2013). Immunosuppression is a down-modulation of the immune system characterized by cell depletion, dysfunction, or dysregulation that may subsequently result in increased susceptibility to infection and tumors. By contrast, immunostimulation is an increased or exaggerated immune responsiveness that may be apparent in the form of a tissue-damaging allergic hypersensitivity response or pathological autoimmunity. However, as knowledge of the mechanisms involved in each of these conditions has expanded, the distinction between them has become less clear. Some agents can cause immunosuppression at one dose or duration of exposure and immunostimulation at others. For instance, the chemotherapeutic drug cyclophosphamide is in most cases immunosuppressive; however, it can also induce autoimmunity (Hutchings et al., 1985).

TABLE 11.3 Antibodies Involved in the Humoral Immune Response

| Antibodies | Serum Concentration | | Characteristics/Functions |
|------------|---------------------|-------|---|
| | mg mL ⁻¹ | % | |
| IgG | 10–12 | 80 | Monomeric structure (γ -globulin); secreted from B cells during secondary response; binds complement; can cross placenta |
| IgM | 1–2 | 5–10 | Pentameric structure; secreted from B cells during primary response; potent binder of complement; high levels indicative of systemic lupus erythematosus or rheumatoid arthritis; cannot cross placenta |
| IgA | 3–4 | 10–15 | Dimeric or monomeric structures; found in seromucous secretions (breast milk); secreted by B cells associated with epithelial cells in GI tract, lung, etc. |
| IgD | 0.03 | <1 | Monomer; extremely labile; functions not well known |
| IgE | <0.0001 | | Reaginic antibody involved in immediate hypersensitivity; antihelminthic; does not bind complement |

Source: Adapted from Clark (1983, pp. 1–453).

TABLE 11.4 Growth and Differentiation Factors of the Immune System

| Factors | Cell of Origin | Primary Immune Functions |
|---|--|--|
| Interleukins^a | | |
| IL-1 | Macrophage, B and T cells | Lymphocyte-activating factor; enhances activation of T & B cells, NK cells, and macrophages |
| IL-2 | T cells (T _h) | T-cell growth factor; stimulates T-cell growth and effector differentiation; stimulates B-cell proliferation/differentiation |
| IL-3 | T cells (T _h) | Mast cell growth factor; stimulates proliferation/differentiation of mast cells, neutrophils, and macrophages |
| IL-4 | T cells (T _h), mast cells, B cells | B-cell growth factor; induces proliferation/differentiation of B cells and secretion of IgA, IgG ₁ , and IgE; promotes T-cell growth; activates macrophages |
| IL-5 | T cells (T _h) | Stimulates antibody secretion (IgA), proliferation of B cells, and eosinophil differentiation |
| IL-6 | T cells, fibroblasts, monocytes | Stimulates growth/differentiation of B cells and secretion of IgG; promotes IL-2-induced growth of T cells |
| IL-7 | Bone marrow stromal cells | Stimulates pre-B- and pre-T-cell growth/differentiation; enhances thymocyte proliferation |
| IL-8 | Monocytes, fibroblasts | Neutrophil chemotaxis |
| IL-9 | T cells | Stimulates T cells and mast cells |
| IL-10 | T cells | Stimulates mast cells and thymocytes; induction of class II MHC |
| Interferons (INF) | | |
| A-INF | Leukocytes and mast cells | Antiviral; increases NK-cell function, B-cell differentiation, potentiates macrophage production of IL-1 |
| B-INF | Fibroblasts, epithelial cells | Antiviral; potentiates macrophage production of IL-1; increases NK-cell function |
| Γ-INF | T cells (T _h), cytotoxic T cells | Antiviral; activates macrophages; induces MHC class II expression on macrophages, epithelial, and endothelial cells |
| Tumor necrosis factors (TNF) | | |
| TNFα | Macrophage, B and T cells | Catectin; promotes tumor cytotoxicity; activates macrophages and neutrophils; enhances IL-2 receptor expression on T cells; inhibits antibody secretion |
| TNFβ | T cells (T _h) NK cells | Lymphotoxin; promotes T-cell-mediated cytotoxicity B-cell activation |
| Colony stimulating factors (CSF) | | |
| | Stem cells | Promotes growth and differentiation of |
| Granulocyte CSF | Myeloid | Granulocytes and macrophages |
| Macrophage CSF | Myeloid | Macrophages and granulocytes |
| Granulocyte-macrophage CSF | Myeloid | Granulocytes, macrophages, eosinophils, mast cells, and pluripotent progenitor cells |

Source: Adapted from Golub and Green (1991).

^a Includes lymphokines, monokines, and cytokines produced by T cells, macrophages, and other cells, respectively.

TABLE 11.5 Cells and Mechanisms Involved in Cell-Mediated Cytotoxicity

| Cell Type | Mechanism of Cytotoxicity |
|----------------------|--|
| T _k cells | T _k cells that are specifically sensitized to antigens on target cells interact directly with target cells to lyse them |
| T _D | Cells involved in delayed hypersensitivity that act indirectly to kill target cells; T _D cells react with antigen and release cytokines that can kill target cells |
| NK cells | Nonspecific T cells that react directly with target cells (tumor cells) without prior sensitization |
| Null cells | Antibody-dependent cell-mediated cytotoxicity (ADCC) involving non-T/non-B cells (null cells) with F _c receptors specific for antibody-coated target cells |
| Macrophages | Nonspecific, direct killing of target by phagocytosis; also involved in presenting antigen to specific T _k cells that can then mediate cytotoxicity as described previously |

Likewise, dimethylnitrosamine, a nitrosamine detected in some foods, has been shown to have both suppressing and enhancing effects on the immune system (Yoshida et al., 1989).

11.4 IMMUNOSUPPRESSION

The various cells of the immune system may differ in their sensitivity to a given xenobiotic. Thus, immunosuppression may be expressed as varying degrees of reduced activity of a single cell type or multiple populations of immunocytes. Several lymphoid organs such as the bone marrow, spleen, thymus, and lymph nodes may be affected simultaneously or the immunodeficiency may be isolated to a single tissue, such as the Peyer's patches of the intestines. The resulting deficiency may in turn lead to an array of clinical outcomes of varying ranges of severity. These outcomes include increased susceptibility to infections, increased severity or persistence of infections, or infections with unusual organisms (e.g., systemic fungal infections). Immunosuppression can be induced in a dose-related manner by a variety of therapeutic agents at dose levels lower than those required to produce overt clinical signs of general toxicity. In addition, immunosuppression can occur without regard to genetic predisposition, given that a sufficient dose level and duration of exposure have been achieved.

Humoral immunity is characterized by the production of antigen-specific antibodies that enhance phagocytosis and destruction of microorganisms through opsonization. Thus, deficiencies of humoral immunity (B lymphocytes) may lead to reduced antibody titers and are typically associated with acute gram-positive bacterial infections (i.e., *Streptococcus*). Although chronic infection is usually associated with dysfunction of some aspect of cellular immunity, chronic infections can also occur when facultative intracellular organisms such as *Listeria* or *Mycobacterium* evade antibodies and multiply within phagocytic cells.

Since cellular immunity results in the release of chemotactic lymphocytes that in turn enhance phagocytosis, a deficiency in cellular immunity may also result in chronic infections. Cellular immunity is mediated by T cells, macrophages, and NK cells involved in complex compensatory networks and secondary changes. Immunosuppressive agents may act directly by lethality to T cells or indirectly by blocking mitosis, lymphokine synthesis, lymphokine release, or membrane receptors to lymphokines. In addition, cellular immunity is involved in the production and release of interferon, a lymphokine that ultimately results in blockage of viral replication. Viruses are particularly susceptible to cytolysis by T cells since they often attach to the surface of infected cells. Thus, immunosuppression of any of the components of cellular immunity may result in an increase in protozoan, fungal, and viral infections as well as opportunistic bacterial infections.

Immune depression may result unintentionally as a side effect of cancer chemotherapy or intentionally from therapeutics administered to prevent graft rejection. In fact, both transplant patients administered immunosuppressive drugs and cancer patients treated with chemotherapeutic agents have been shown to be at high risk of developing secondary cancers, particularly of lymphoreticular etiology (Penn, 1977). Most of these drugs are alkylating or cross-linking agents that by their chemical nature are electrophilic and highly reactive with nucleophilic macromolecules (protein and nucleic acids). Nucleophilic sites are quite ubiquitous and include amino, hydroxyl, mercapto, and histidine functional groups. Thus, immunotoxic agents used in chemotherapy may induce secondary tumors through direct genotoxic mechanisms (i.e., DNA alkylation).

Reduced cellular immunity may result in increased malignancy and decreased viral resistance through indirect mechanisms as well, by modulating immune surveillance of aberrant cells. T lymphocytes, macrophage cells, and NK cells are all involved in immunosurveillance through cytolysis of virally infected cells or tumor cells, each by a different mechanism (Table 11.5) (Burnet, 1970). In addition to the common cell types described in Table 11.5, at least two other types of cytotoxic effector cells of T-cell origin have been identified, each of which has a unique lytic specificity phenotype and activity profile (Merluzzi, 1985). Of these, both LAK and TIL cells have been shown to lyse a variety of different tumor cells. However, TIL cells have 50–100 times more lytic activity than LAK cells. Most tumor cells express unique surface antigens that render them different from normal cells. Once detected as foreign, they are presented to the T helper cells in association with MHC molecules to form an antigen–MHC complex. This association elicits a genetic component to the immunospecificity reaction. T helper cells subsequently direct the antigen complex toward the cytotoxic T lymphocytes, which possess receptors for antigen–MHC complexes. These cells can then proliferate, respond to specific viral antigens or antigens on the membranes of tumor cells, and destroy them (Yoshida et al., 1989).

In contrast, the macrophages and NK cells are involved in nonspecific immunosurveillance in that they do not require prior sensitization with a foreign antigen as a prerequisite for lysis and are not involved with MHC molecules. The enhancement of either NK-cell function or macrophage function has been shown to reduce metastasis of some types of tumors. Macrophage cells accumulate at the tumor site and have been shown to lyse a variety of transformed tumor cells (Volkman, 1984). NK cells are involved in the lysis of primary autochthonous tumor cells. Migration of NK cells to tumor sites has been well documented. Although not clearly defined, it appears that they can recognize certain proteinaceous structures on tumor cells and lyse them with cytolysin.

11.4.1 Immunosuppressive Drugs

Table 11.6 lists numerous types of drugs that are immunosuppressive and describes their immunotoxic effects (Brunton et al., 2006). Several classes of drugs that characteristically depress the immune system are further discussed in the following.

11.4.1.1 Antimetabolites This class of drugs includes purine, pyrimidine, and folic acid analogs that have been successfully used to treat various carcinomas, autoimmune diseases, and dermatological disorders such as psoriasis. Because of their structural similarities to normal components of DNA and RNA synthesis, they are capable of competing with the normal macromolecules and alkylating biological nucleophiles.

Thioguanine and mercaptopurine are purine analogs structurally similar to guanine and hypoxanthine that have been used to treat malignancies. Azathioprine, an imidazolyl derivative of mercaptopurine, has been used as an immunosuppressive therapeutic in organ transplants and to treat severe refractory rheumatoid arthritis (Hunter et al., 1975) and autoimmune disorders including pemphigus vulgaris and bullous pemphigoid. These drugs act as antimetabolites to block *de novo* purine synthesis through the erroneous incorporation of thioinosinic acid into the pathway in place of inosine. The antimetabolite can bind to the inosine receptor, which in turn will inhibit the synthesis of DNA, RNA, protein synthesis, and ultimately T-cell differentiation (Hadden et al., 1984). For example, both thioguanine and mercaptopurine can act as substrates for the HGPRT enzyme to produce T-IMP (thioinosine monophosphate) and T-GMP (thioguanine monophosphate), respectively. Thioinosine monophosphate is a poor substrate for guanylyl kinase, which would normally catalyze the conversion of GMP to GDP (Calabresi and Chabner, 1990). Thus, T-IMP can accumulate in the cell and inhibit several vital metabolic reactions. At high doses, these drugs can suppress the entire immune system. However, at clinical dosages, only the T-cell response is affected, without an apparent decrease in T-cell numbers (Spreafico and Anaclerio, 1977).

Pentostatin (2'-deoxycoformycin) is an adenosine analog that is a potent inhibitor of adenosine deaminase. Pentostatin is particularly useful for treating T-cell leukemia since malignant T cells have higher levels of adenosine deaminase than most cells. Similar to individuals that are genetically deficient in adenosine deaminase, treatment with pentostatin produces immunosuppression of both T and B lymphocytes, with minimal effect on other tissues. As a result, severe opportunistic infections are often associated with its clinical use.

5-Fluorouracil (5-FU), adenosine arabinoside (AraA), and cytosine arabinoside (AraC) are pyrimidine analogs to uracil, adenine, and cytosine, respectively. 5-FU is used primarily to treat cancer of the breasts and gastrointestinal

tract as well as severe recalcitrant psoriasis (Alper et al., 1985). AraC is predominantly indicated for the treatment of acute leukemia and non-Hodgkin's lymphomas. Although high-dose therapy with AraC has a good likelihood of producing complete remission, it is often accompanied by severe leukopenia, thrombocytopenia, and anemia (Barnett et al., 1985). Likewise, myelosuppression is the major toxicity associated with bolus-dose regimens of 5-FU.

11.4.1.2 Glucocorticosteroids Corticosteroids are commonly used to reduce inflammation, treat autoimmune diseases such as systemic lupus erythematosus (SLE), and as a prophylactic measure to prevent transplant rejection. The adrenocorticosteroid prednisone is often coadministered with other immunosuppressives such as cyclosporine and azathioprine (Elion and Hitchings, 1975). Glucocorticosteroids act pharmacologically by modulating the rate of protein synthesis. The molecule reacts with specific receptors to form a complex that crosses into the nucleus of the cell and regulates transcription of specific mRNA. The corticosteroid complex releases inhibition of transcription, thus enhancing protein synthesis (Hollenberg et al., 1987). This may lead to the initiation of *de novo* synthesis of the phospholipase A2-inhibiting protein, lipocortin, which blocks the synthesis of arachidonic acid and its prostaglandin and leukotriene metabolites (Wallner et al., 1986). Glucocorticosteroids induce immunosuppression and anti-inflammation as a result of the inhibition of specific leukocyte functions such as lymphokine activity. Glucocorticoids can also inhibit recruitment of leukocytes and macrophages into the site of inflammation. In addition, amplification of cell-mediated immunity can be suppressed by inhibiting the interaction of IL-2 with its T-cell receptors. However, the immunosuppression is reversible and immune function recovers once therapy has ceased.

11.4.1.3 Cyclosporine Cyclosporin A (cyclosporine) is an 11-amino-acid cyclic peptide residue of fungal origin isolated from the fermentation products of *Trichoderma polysporum* and *Cylindrocarpon lucidum*. In addition to having a very narrow range of antibiotic activity, it was also found to inhibit proliferation of lymphocytes, which made it unsuitable as an antibiotic. Cyclosporine inhibits the early cellular response of helper T cells to antigens (Kay and Benzie, 1984) primarily by inhibiting production of IL-2 (Elliot et al., 1984), and at higher doses it may inhibit expression of IL-2 receptors (Herold et al., 1986). Cyclosporine does not prevent the stimulation of helper T-cell clonal expansion by IL-2, only its activation. Since it is not myelosuppressive at therapeutic dosages, the incidence of secondary infection is lower than that induced by other classes of immunosuppressives. Thus, cyclosporine is ideal as an immunosuppressive agent to prevent transplant rejection and graft-versus-host disease (Kahan and Bach, 1988). Cyclosporine has also

TABLE 11.6 Immunosuppressive Drugs and Their Effects

| Drugs | Biological Activity and Indications | Immunotoxic Effects |
|---|---|--|
| <i>Hormones and Antagonists</i> | | |
| Corticosteroids (prednisone) | Anti-inflammatory; systemic lupus erythematosus; leukemias; rheumatoid arthritis; breast cancer | Depresses T- and B-cell function; reduces lymphokines; alters macrophage function; increases infections |
| Diethylstilbestrol | Synthetic estrogen; cancer chemotherapy | Depletes or functionally impairs T cells; enhances macrophage suppressor cell; increases infections and tumorigenesis |
| Estradiol | Synthetic estrogen; dysmenorrhea; osteoporosis | Decreases T _h cells and IL-2 synthesis; increases T _s cell function, infections, and tumorigenesis |
| <i>Antibiotics</i> | | |
| Cephalosporins | β -Lactam antimicrobial | Granulocytopenia; cytopenia |
| Chloramphenicol | Wide-spectrum antimicrobial | Pancytopenia, leukopenia (idiosyncratic) |
| Penicillins | β -Lactam antimicrobial | Granulocytopenia; cytopenia |
| Rifampin | Macrocyclic antibiotic | Suppresses T-cell function |
| Tetracyclines | Antimicrobial | Decreased migration of granulocytes |
| <i>Chemotherapeutics and Immunomodulators</i> | | |
| Arabinoside (AraA and AraC) | Antimetabolites; antivirals; leukemias; lymphomas | Leukopenia; thrombocytopenia |
| Azathioprine | Antimetabolite; leukemia; arthritis; transplant rejection | Inhibits protein synthesis; bone marrow suppression |
| Busulfan | Alkylating agent; chronic granulocytic leukemia | Leukopenia; myelosuppressive; granulocytopenia |
| Carmustine and lomustine (BCNU and CCNU) | Alkylating agents; Hodgkin's disease; lymphomas | Delayed hematopoietic depression; leukopenia; thrombocytopenia |
| Chlorambucil | Alkylating agent; leukemia; lymphomas; vasculitis | Bone marrow suppression; myelosuppressive |
| Cyclophosphamide (cytotoxin) | Alkylating agent; cancer chemotherapy; transplant rejection; rheumatoid arthritis | Decreased T _s cells, B cells, and NK cells |
| Cyclosporin A | Transplant rejections | Depresses T cells; inhibits IL-2 production |
| Interferon | Immunomodulator; antiviral, hairy cell leukemia | Bone marrow suppression; granulocytopenia; leukopenia |
| Melphalan (L-PAM) | Alkylating agent; breast and ovarian cancer | Leukopenia; bone marrow suppression; granulocytopenia; pancytopenia |
| 6-Mercaptopurine | Antimetabolite; acute leukemias; arthritis | Decreased T-cell function; bone marrow suppression |
| Methotrexate | Folic acid analog; cancer chemotherapy, arthritis | Inhibits proliferation; T-cell suppression; granulocytopenia; lymphocytopenia |
| Pentostatin | Adenosine analog; T-cell leukemia | Inhibits adenosine deaminase; suppresses T and B cells |
| Zidovudine (AZT) | Antiviral (HIV) | Decreases T _h cells and granulocytes |
| <i>Miscellaneous</i> | | |
| Colchicine | Antimitotic; gout; anti-inflammatory | Inhibits migration of granulocytes; leukopenia; agranulocytosis |
| Diphenylhydantoin (phenytoin) | Antiepileptic | Leukocytopenia; neutropenia |
| Indomethacin (Indocin) | Nonsteroidal anti-inflammatory; analgesic; antipyretic | Neutropenia |
| Procainamide | Antiarrhythmic | Agranulocytosis; leukopenia (rare) |
| Sulfasalazine | Antimicrobial anti-inflammatory; ulcerative colitis/inflammatory bowel diseases | Suppresses NK cells; impaired lymphocyte function |

Source: Adapted from Brunton et al. (2006).

been used as an antihelminthic and as an anti-inflammatory agent to treat rheumatoid arthritis and other autoimmune-type diseases.

11.4.1.4 Nitrogen Mustards Nitrogen mustards characteristically consist of a bis(2-chloroethyl) group bonded to nitrogen. These molecules are highly reactive bifunctional alkylating agents that have been successfully used in cancer chemotherapy. Included in this group are mechlorethamine, L-phenylalanine mustard (melphalan), chlorambucil, ifosfamide, and cyclophosphamide. The cytotoxic effects of each on the bone marrow and lymphoid organs are similar; however, their pharmacokinetic and toxic profiles can vary on the basis of the substituted side group. For example, the side group may consist of a simple methyl group, as is the case of mechlorethamine, or substituted phenyl groups, in the cases of melphalan and chlorambucil.

Cyclophosphamide, which contains a cyclic phosphamide group bonded to the nitrogen mustard, is representative of this class. The parent compound itself is not active *in vitro* unless treated in conjunction with an exogenous P450 microsomal enzyme system (Colvin, 1982) such as rat liver S9 homogenate, which metabolizes it to a highly reactive alkylating agent (4-hydroxy-cyclophosphamide). Thus, *in vivo*, cyclophosphamide is not toxic until it is metabolically activated in the liver. Cyclophosphamide has been the most widely used nitrogen mustard where it has been effective as a cancer chemotherapeutic and to treat autoimmune-type diseases including SLE, multiple sclerosis, and rheumatoid arthritis (Calabresi and Parks, 1985). Treatment with cyclophosphamide suppresses all classes of lymphoid cells, which may result in reduced lymphocyte function as well as lymphopenia and neutropenia (Webb and Winkelstein, 1982). Thus, it has also been administered as a large single dose prior to bone marrow transplants to suppress cellular immunity and subsequently inhibit rejection (Shand, 1979).

11.4.1.5 Estrogens β -Estradiol (Luster et al., 1984; Pung et al., 1984) and therapeutics with estrogenic activity, such as diethylstilbestrol (DES), have also been shown to be immunosuppressive (Luster et al., 1985). Estrogens have been shown to increase T suppressor cell activity in splenocytes, decrease numbers of T helper cells, inhibit IL-2 synthesis, and modulate production of immunoregulatory factors (Luster et al., 1987). These effects have been particularly characterized in studies with DES, a nonsteroidal synthetic estrogen used widely in the treatment of prostate and breast cancers, as well as administered to pregnant women as a "morning-after" contraceptive. Decreased mitogenicity of human peripheral blood lymphocytes has been observed in men treated with DES for prostate cancer and women exposed *in utero* (Haukaas et al., 1982; Ways et al., 1987). In mice, thymic involution and atrophy with depletion of the cortical lymphocytes have been observed histologically.

Function is also modulated, as evident by depressed mixed lymphocyte responses, mitogenicity, and T-cell release of IL-2 (Pung et al., 1985). Dean et al. (1980) speculated that DES treatment selectively depletes or functionally impairs T cells and/or the induction of suppressor macrophages, resulting in immunosuppression. Macrophage suppressor cell activity is enhanced (Luster et al., 1980) and PMN cells accumulate following bacterial challenge. Although macrophage functions of phagocytosis and tumor growth inhibition are potentiated, defects in macrophage migration and decreased bactericidal activity contribute to decreased host resistance with resulting increased susceptibility to bacterial infections.

11.4.1.6 Heavy Metals Some heavy metals such as gold and platinum are used pharmacologically as immunomodulators to treat rheumatoid arthritis and as antineoplastic drugs, respectively. Most heavy metals inhibit mitogenicity, antibody responses, and host resistance to bacterial or viral challenge and tumor growth. Platinum has been shown to suppress humoral immunity, lymphocyte proliferation, and macrophage function (Lawrence, 1985). Clinically, mild to moderate myelosuppression may also be evident with transient leukopenia and thrombocytopenia. Mercury salts can induce autoimmune reactions (Bellon et al., 1982).

Likewise, injectable gold salts such as gold sodium thiomalate affect a variety of immune responses in man (Bloom et al., 1987). Severe thrombocytopenia occurs in 1% of patients as a result of an immunological disturbance that accelerates the degradation of platelets. Leukopenia, agranulocytosis, and fatal aplastic anemia may also occur. Although better tolerated than parenteral preparations, the organic gold compound, auranofin, administered orally is also immunosuppressive. In a dog study, auranofin was shown to produce thrombocytopenia similar to that described in humans administered parenteral preparations (Bloom et al., 1985a). Long-term toxicity studies with these compounds in dogs show evidence of immune-modulating activity, possible drug-induced immunotoxicity, and treatment-related changes in immune function (e.g., lymphocyte activation).

11.4.1.7 Antibiotics β -Lactam-containing antibiotics such as the cephalosporins may also induce significant immunosuppressive effects (Caspritz and Hadden, 1987) in a small percentage of human patients. Adverse effects including anemia, neutropenia, thrombocytopenia, and bone marrow depression were observed in dogs administered high doses of cefonicid for 6 months (Bloom et al., 1985b). A similar syndrome has been characterized in cefazedone-treated dogs expressing an agglutinating red cell antibody. Further studies with this drug indicated that both cytopenia (Bloom et al., 1985b) and suppression of bone marrow stem cell activity appear to be antibody mediated (Deldar et al., 1985).

11.5 IMMUNOSTIMULATION

A variety of drugs as well as environmental chemicals have been shown to have immunostimulatory or sensitizing effects on the immune system, and these effects are well documented in humans exposed to drugs (DeSwarte, 1986). The drug or metabolite can act as a hapten and covalently bind to a protein or other cellular constituent of the host to appear foreign and become antigenic. Haptens are low molecular weight substances that are not in themselves immunogenic but will induce an immune response if conjugated with nucleophilic groups on proteins or other macromolecular carriers. In both allergy and autoimmunity, the immune system is stimulated or sensitized by the drug conjugate to produce specific pathological responses. An allergic hypersensitivity reaction may vary from one which results in an immediate anaphylactic response to one which produces a delayed hypersensitivity reaction or immune complex reaction. Allergic hypersensitivity reactions result in a heightened sensitivity to nonself antigens, whereas autoimmunity results in an altered response to self antigens. Unlike immunosuppression, which nonspecifically affects all individuals in a dose-related manner, both allergy and autoimmunity have a genetic component that creates susceptibility in those individuals with a genetic predisposition. Susceptible individuals, once sensitized, can respond to even minute quantities of the antigen. Several examples of drugs that can stimulate the immune system are presented in Table 11.7.

11.5.1 Hypersensitivity (or Allergenicity)

The four types of hypersensitivity reactions as classified by Coombs and Gell (1975) are outlined in Table 11.8. The first three types are immediate antibody-mediated reactions, whereas the fourth type is a cellular-mediated delayed-type response that may require 1–2 days to occur after a secondary exposure. Type I reactions are characterized by an anaphylaxis response to a variety of compounds, including proteinaceous materials and pharmaceuticals such as penicillin. Various target organs may be involved depending on the route of exposure. For example, the gastrointestinal tract is usually involved with food allergies, the respiratory system with inhaled allergens, the skin with dermal exposure, and smooth muscle vasculature with systemic exposure. The type of response elicited often depends on the site of exposure and includes dermatitis and urticaria (dermal), rhinitis and asthma (inhalation), increased gastrointestinal emptying (ingestion), and systemic anaphylactic shock (parenteral).

11.5.1.1 Type I Hypersensitivity During an initial exposure, IgE antibodies are produced and bind to the cell surface of mast cells and basophils. Upon subsequent exposures to the antigen, reaginic IgE antibodies bound to

TABLE 11.7 Drugs That Produce Immunostimulation

| Drug | Type of Response |
|--|--|
| <i>Hypersensitivity</i> | |
| Antibiotics | |
| Cephalosporins | Anaphylaxis, urticaria, rash, granulocytopenia |
| Chloramphenicol | Rash, dermatitis, urticaria |
| Neomycin | Dermal exposure—rash, dermatitis |
| Sulfathiazole | Rash, dermatitis, urticaria |
| Spiramycin | Rash, dermatitis, urticaria |
| Quinolones | Photosensitivity |
| Tetracyclines | Photosensitivity, anaphylaxis, asthma, dermatitis |
| Others | |
| Allopurinol | Rash, urticaria, fever, eosinophilia |
| Avridine | Delayed-type hypersensitivity; increases NK cells, T cells, IL-1, and IL-2 |
| Isoprinosine | Delayed-type hypersensitivity; increases T-lymphocytes |
| Indomethacin | Rash, urticaria, asthma, granulocytopenia |
| Quinidine | Fever, anaphylaxis, asthma |
| Salicylates | Rash, urticaria |
| <i>Autoimmunity</i> | |
| Amiodarone | Thyroiditis |
| Captopril | Autoimmune hemolytic anemia, pemphigus, granulocytopenia |
| Chlorpromazine | Granulocytopenia |
| Halothane | Autoimmune chronic active hepatitis |
| Hydralazine | Autoimmune hemolytic anemia, drug-induced SLE, myasthenia gravis, pemphigus, glomerulonephritis, Goodpasture's disease |
| Methyldopa | Autoimmune hemolytic anemia, leukopenia, drug-induced SLE, pemphigus |
| Nitrofurantoin | Peripheral neuritis |
| D-penicillamine | Autoimmunity; drug-induced SLE, myasthenia gravis, pemphigus, glomerulonephritis, Goodpasture's disease |
| Propranolol | Autoimmunity |
| Procainamide | Autoimmunity, drug-induced SLE, rash, vasculitis, myalgias |
| Pyrithioxine | Pemphigus |
| <i>Hypersensitivity and Autoimmunity</i> | |
| Antibiotics | |
| Isoniazid | Rash, dermatitis, vasculitis, arthritis, drug-induced SLE |
| Penicillins | Anaphylaxis, dermatitis; vasculitis, serum sickness, hemolytic anemia |
| Sulfonamides | Dermatitis, photosensitivity; pemphigus, hemolytic anemia, serum sickness, drug-induced SLE |
| Others | |
| Acetazolamide | Rash, fever, autoimmunity |
| Lithium | Dermatitis; autoimmune thyroiditis, vasculitis |
| Thiazides | Hypersensitivity, photosensitivity; autoimmunity (diabetes) |
| Phenytoin | Rash; drug-induced SLE, hepatitis |

TABLE 11.8 Types of Hypersensitivity Responses

| Type and Designation | Agents: Clinical Manifestations | Components | Effects | Mechanism |
|------------------------------|---|------------------------------------|---|--|
| I, Immediate (reaginic) | Food additives (GI allergies; anaphylactic) Penicillin: urticaria and dermatitis | Mast cells; IgE | Anaphylaxis, asthma, urticaria, rhinitis, dermatitis | IgE binds to mast cells to stimulate release of humoral factors |
| II, Cytotoxic | Cephalosporin: hemolytic anemia Quinidine: thrombocytopenia | IgG, IgM | Hemolytic anemia, Goodpasture's disease | IgG and IgM bind to cells (e.g., RBCs), fix complement (opsonization), then lyse cells |
| III, Immune complex (Arthus) | Methicillin: chronic glomerulonephritis | Antigen-antibody complexes (Ag-Ab) | SLE, rheumatoid arthritis, glomerular nephritis, serum sickness, vasculitis | Ag-Ab complexes deposit in tissues, and may fix complement |
| IV, Delayed hypersensitivity | Penicillin: contact dermatitis | T _D cells; macrophages | Contact dermatitis, tuberculosis | Sensitized T cells induce a delayed-hypersensitivity response upon challenge |

Source: Based on classification system of Coombs and Gell (1975, p. 761).

TABLE 11.9 Proteins and Soluble Mediators Involved in Hypersensitivity

| Factor | Origin | Characteristics/Functions |
|----------------|-----------------------|--|
| Histamine | Mast cells, basophils | Contraction of smooth muscle; increases vascular permeability |
| Serotonin | Mast cells, basophils | Contraction of smooth muscle; leukotriene |
| SRS-A | Lung tissue | (Slow-reacting substance of anaphylaxis); contraction of smooth muscle; acidic polypeptide |
| ECF-A | Mast cells | (Eosinophilic chemotactic factor of anaphylaxis); attracts eosinophils; small peptide |
| Prostaglandins | Various tissues | Modifies release of histamine and serotonin from mast cells and basophils |

Source: Adapted from Clark (1983, pp. 1-453).

the surface of target cells at the F_c region (mast cells and basophils) become cross-linked (at the F_{ab} regions) by the antigen. Cross-linking causes distortion of the cell surface and IgE molecule, which, in turn, activates a series of enzymatic reactions, ultimately leading to degranulation of the mast cells and basophils. These granules contain a variety of pharmacological substances (Table 11.9), such as histamines, serotonins, prostaglandins, bradykinins, and leukotrienes (SRS-A and ECR-A). Upon subsequent challenge exposures, these factors are responsible for eliciting an allergic reaction through vasodilation and increased vascular permeability. The nasal passages contain both mast cells and plasma cells that secrete IgE antibodies. Allergic responses localized in the nasal mucosa result in dilation of the local blood vessels, tissue swelling, mucus secretion, and sneezing. Reactions localized in the respiratory tract, also rich in mast cells and IgE, result in an allergic asthma response. This condition is triggered by the release of histamine and SRS-A, which induce constriction of the bronchi and alveoli, pulmonary edema, and mucous secretions that block the bronchi and alveoli, together resulting in severe difficulty in breathing. In the case of a challenge dose of a drug

administered systemically, the reactive patient may have difficulty breathing within minutes of exposure and may experience convulsions, vomiting, and low blood pressure. The effects of anaphylactic shock and respiratory distress, if severe, may ultimately result in death. The best models for this use guinea pigs (Mazzone and Canning, 2002).

Antibiotics containing β -lactam structures, such as penicillin and cephalosporins, are the most commonly occurring inducers of anaphylactic shock and drug hypersensitivity in general. Other hypersensitivity reactions may include urticarial rash, fever, bronchospasm, serum sickness, and vasculitis with reported incidences of all types varying from 0.7 to 10% (Idsøe et al., 1968) and the incidence of anaphylactoid reactions varying from 0.04 to 0.2%. When the β -lactam ring is opened during metabolism, the penicilloyl moiety can form covalent conjugates with nucleophilic sites on proteins. The penicilloyl conjugates can then act as haptens to form the determinants for antibody induction. Although most patients that have received penicillin produce antibodies against the metabolite benzylpenicilloyl, only a fraction experience allergic reaction (Garraty and Petz, 1975), which suggests a genetic component to susceptibility.

11.5.1.2 Type II Hypersensitivity Type II cytolytic reactions are mediated by IgG and IgM antibodies that can fix complement, opsonize particles, or induce antibody-dependent cellular cytotoxicity reactions. Erythrocytes, lymphocytes, and platelets of the circulatory system are the major target cells that interact with the cytolytic antibodies causing depletion of these cells. Hemolytic anemia (penicillin, methyl dopa), leukopenia, thrombocytopenia (quinidine), and/or granulocytopenia (sulfonamide) may result. Type II reactions involving the lungs and kidneys occur through the development of antibodies (autoantibodies) to the basement membranes in the alveoli or glomeruli, respectively. Prolonged damage may result in Goodpasture's disease, an autoimmune disease characterized by pulmonary hemorrhage and glomerulonephritis. Several other autoimmune-type diseases have been associated with extended treatments with D-penicillamine and other pharmaceuticals. Various types of autoimmune responses and examples of drug-induced autoimmunity are discussed in further detail later in this section.

11.5.1.3 Type III Hypersensitivity Type III reactions (Arthus) are characterized as an immediate hypersensitivity reaction initiated by antigen-antibody complexes that form freely in the plasma instead of at the cell surface. Regardless of whether the antigens are self or foreign, complexes mediated by IgG can form and settle into the tissue compartments of the host. These complexes can then fix complement and release C3a and C5a fragments that are chemotactic for phagocytic cells. Polymorphonuclear leukocytes are then attracted to the site, where they phagocytize the complexes and release hydrolytic enzymes into the tissues. Additional damage can be caused by binding to and activating platelets and basophils, which, in the end, results in localized necrosis, hemorrhage, and increased permeability of local blood vessels. These reactions commonly target the kidney, resulting in glomerulonephritis through the deposition of the complexes in the glomeruli.

Some antibiotics (β -lactam) have been reported to produce glomerular nephritis in humans that has been attributed to circulating immune complexes. These complexes have also been observed in preclinical toxicology studies with baboons treated with a β -lactam antibiotic, prior to the appearance of any biochemical or clinical changes (Descotes and Mazue, 1987). In addition, immunoglobulin complexes have been observed in rats treated with gold and autologous immune complex nephritis has been observed in guinea pigs (Ueda et al., 1980). Similar evidence of immunomediated nephrotoxicity has been reported in rheumatoid arthritis patients administered long-term treatments with gold compounds; proteinuria has been observed in approximately 10% of these patients.

Other target organs such as the skin with lupus, the joints with rheumatoid arthritis, and the lungs with pneumonitis may be affected. The deposition of antigen-antibody complexes

through the circulatory system results in a syndrome referred to as serum sickness, which was quite prevalent prior to 1940 (Clark, 1983), when serum therapy for diphtheria was commonly used. Serum sickness occurs when the serum itself becomes antigenic as a side effect from passive immunization with heterologous antiserum produced from various sources of farm animals. The antitoxin for diphtheria was produced in a horse and administered to humans as multiple injections of passive antibody. As a consequence, these people often became sensitized to the horse serum and developed a severe form of arthritis and glomerulonephritis caused by deposition of antigen-antibody complexes. Clinical symptoms of serum sickness present as urticarial skin eruptions, arthralgia or arthritis, lymphadenopathy, and fever. Drugs such as sulfonamides, penicillin, and iodides can induce a similar type of reaction. Although uncommon today, transplant patients receiving immunosuppressive therapy with heterologous anti-lymphocyte serum or globulins may also exhibit serum sickness.

11.5.1.4 Type IV Delayed-Type Hypersensitivity (DTH) Delayed-type hypersensitivity (DTH) reactions are T-cell mediated with no involvement of antibodies. However, these reactions are controlled through accessory cells, suppressor T cells, and monokine-secreting macrophages, which regulate the proliferation and differentiation of T cells. The most frequent form of DTH manifests itself as contact dermatitis. The drug or metabolite binds to a protein in the skin or the Langerhans cell membrane (class II MHC molecules) where it is recognized as an antigen and triggers cell proliferation. After a sufficient period of time for migration of the antigen and clonal expansion (latency period), a subsequent exposure will elicit a dermatitis reaction. A 24–48 h delay often occurs between the time of exposure and onset of symptoms to allow time for infiltration of lymphocytes to the site of exposure. The T cells ($CD4^+$) that react with the antigen are activated and release lymphokines that are chemotactic for monocytes and macrophages. Although these cells infiltrate to the site via the circulatory vessels, an intact lymphatic drainage system from the site is necessary since the reaction is initiated in drainage lymph nodes proximal to the site (Clark, 1983). The release (degranulation) of enzymes and histamines from the macrophages may then result in tissue damage. Clinical symptoms of local dermal reactions may include a rash (not limited to sites of exposure), itching, and/or burning sensations. Erythema is generally observed in the area around the site, which may become thickened and hard to the touch. In severe cases, necrosis may appear in the center of the site followed by desquamation during the healing process. The immune-enhancing drugs isoprinosine and avridine have been shown to induce a DTH reaction in rats (Exon et al., 1986).

A second form of DTH response is similar to that of contact dermatitis in that macrophages are the primary

effector cells responsible for stimulating CD4⁺ T cells; however, this response is not necessarily localized to the epidermis. A classical example of this type of response is demonstrated by the tuberculin diagnostic tests. To determine if an individual has been exposed to tuberculosis, a small amount of fluid from tubercle bacilli cultures is injected subcutaneously. The development of induration after 48 h at the site of injection is diagnostic of prior exposure.

Shock, similar to that of anaphylaxis, may occur as a third form of a delayed systemic hypersensitivity response. However, unlike anaphylaxis, IgE antibodies are not involved. This type of response may occur 5–8 h after systemic exposure and can result in fatality within 24 h following intravenous or intraperitoneal injection.

A fourth form of delayed hypersensitivity results in the formation of granulomas. If the antigen is allowed to persist unchecked, macrophages and fibroblasts are recruited to the site to proliferate, produce collagen, and effectively “wall off” the antigen. A granuloma requires a minimum of 1–2 weeks to form.

11.5.2 Photosensitization

Regardless of the route of exposure, some haptens (photoantigens) that are absorbed locally into the skin, or reach the skin through systemic absorption, can be photoactivated by ultraviolet (UV) light between 320 and 400 nm. Once activated, the hapten can bind to the dermal receptors to initiate sensitization (photoallergy). Subsequent exposures to the hapten in the presence of UV light can result in a hypersensitivity response. Clinical symptoms of photoallergy may occur within minutes (immediate hypersensitivity) of exposure to sunlight, or 24 h or more after exposure (DTH). Symptoms may range from acute urticarial reactions to eczematous or papular lesions. Although both phototoxic and photoallergic reactions require the compound to be exposed to sunlight in order to elicit a response, their mechanisms of action are quite different. Since photosensitization is an immune-mediated condition, repeated exposures with a latency period between the initial exposure and subsequent exposures is required, the response is not dose related (small amounts can produce a response once sensitized), and not all individuals exposed to the compound will necessarily respond (genetic component to susceptibility). Although both conditions can present similar symptoms (erythema), phototoxicity is limited mainly to erythema, whereas photoallergy can result in erythema, edema, and dermatitis as described previously.

Several drug classes, including tetracycline, sulfonamide, and quinolone antibiotics, as well as chlorothiazide, chlorpromazine, and amiodarone hydrochloride, have been shown to be photoantigens. Photosensitivity may persist even after withdrawal of the drug, as has been observed with

the antiarrhythmic drug amiodarone hydrochloride, since it is lipophilic and can be stored for extended periods in the body fat (Unkovic et al., 1984). In addition, it is quite common for cross-reactions to occur between structurally related drugs of the same class.

11.5.3 Autoimmunity

In autoimmunity, as with hypersensitivity, the immune system is stimulated by specific responses that are pathogenic, and both tend to have a genetic component that predisposes some individuals more than others. However, as is the case with hypersensitivity, the adverse immune response of drug-induced autoimmunity is not restricted to the drug itself but also involves a response to self antigens.

Autoimmune responses directed against normal components of the body may consist of antibody-driven humoral responses and/or cell-mediated, DTH responses. T cells can react directly against specific target organs, or B cells can secrete autoantibodies that target “self.” Autoimmunity may occur spontaneously as the result of a loss of regulatory controls that initiate or suppress normal immunity causing the immune system to produce lymphocytes reactive against its own cells and macromolecules such as DNA, RNA, or erythrocytes.

Although autoantibodies are often associated with autoimmune reactions, they are not necessarily indicative of autoimmunity (Russel, 1981). Antinuclear antibodies can occur normally with aging in some healthy women without autoimmune disease, and all individuals have B cells with the potential of reacting with self antigens through Ig receptors (Dighiero et al., 1983). The presence of an antibody titer to a particular immunogen indicates that haptenization of serum albumin has occurred as part of a normal immune response. However, if cells are stimulated to proliferate and secrete autoantibodies directed against a specific cell or cellular component, a pathological response may result. The tissue damage associated with autoimmune disease is usually a consequence of type II or III hypersensitivity reactions that result in the deposition of antibody–antigen complexes.

Several diseases have been associated with the production of autoantibodies against various tissues. For example, an autoimmune form of hemolytic anemia can occur if the antibodies are directed against erythrocytes. Similarly, antibodies that react with acetylcholine receptors may cause myasthenia gravis, those directed against glomerular basement membranes may cause Goodpasture’s syndrome, and those that target the liver may cause hepatitis. Other forms of organ-specific autoimmunity include autoimmune thyroiditis (as seen with amiodarone) and juvenile diabetes mellitus, which can result from autoantibodies directed against the tissue-specific antigens thyroglobulin and cytoplasmic components of pancreatic islet cells, respectively.

In contrast, systemic autoimmune diseases may occur if the autoantibodies are directed against an antigen that is ubiquitous throughout the body, such as DNA or RNA. For example, SLE occurs as the result of autoimmunity to nuclear antigens that form immune complexes in the walls of blood vessels and basement membranes of tissues throughout the body.

The etiology of drug-induced autoimmunity is not well established and is confounded by factors such as age, sex, and nutritional state, as well as genetic influences on pharmacological and immune susceptibility. Unlike idiopathic autoimmunity, which is progressive or characterized by an alternating series of relapses and remissions, drug-induced autoimmunity is thought to subside after the drug is discontinued. However, this is not certain since a major determining factor for diagnosis of a drug-related disorder is dependent on the observation of remission upon withdrawal of the drug (Bigazzi, 1988).

One possible mechanism for xenobiotic-induced autoimmunity involves xenobiotic binding to autologous molecules, which then appear foreign to the immunosurveillance system. If a self antigen is chemically altered, a specific T helper (T_h) cell may see it as foreign and react to the altered antigenic determinant portion, allowing an autoreactive B cell to react to the unaltered hapten. This interaction results in a carrier-hapten bridge between the specific T_h and autoreactive B cell, bringing them together for subsequent production of autoantibodies specific to the self antigen that was chemically altered (Weigle, 1980). Conversely, a xenobiotic may alter B cells directly, including those that are autoreactive. Thus, the altered B cells may react to self antigens independent from T_h -cell recognition and in a non-tissue-specific manner.

Another possible mechanism is that the xenobiotic may stimulate nonspecific mitogenicity of B cells. This could result in a polyclonal activation of B cells with subsequent production of autoantibodies. Alternatively, the xenobiotic may stimulate mitogenicity of T cells that recognize self, which in turn activate B-cell production of antibodies in response to "self" molecules. There is also evidence to suggest that anti-DNA autoantibodies may originate from somatic mutations in lymphocyte precursors with antibacterial or antiviral specificity. For example, a single amino acid substitution resulting from a mutation in a monoclonal antibody to polyphorylcholine was shown to result in a loss of the original specificity and an acquisition of DNA reactivity similar to that observed for anti-DNA antibodies in SLE (Talal, 1987).

The mechanisms of autoimmunity may also entail interaction with MHC structures determined by the HLA alleles. Individuals carrying certain HLA alleles have been shown to be predisposed to certain autoimmune diseases, which may account in part for the genetic variability of autoimmunity. In addition, metabolites of a particular drug may vary

between individuals to confound the development of drug-induced autoimmunity. Dendritic cells, such as the Langerhans cells of the skin and B lymphocytes that function to present antigens to T_h cells, express class II MHC structures. Although the exact involvement of these MHC structures is unknown, Gleichmann et al. (1989) have theorized that self antigens rendered foreign by drugs such as D-penicillamine may be presented to T_h cells by MHC class II structures. An alternate hypothesis is that the drug or a metabolite may alter MHC class II structures on B cells, making them appear foreign to T_h cells.

A number of different drugs have been shown to induce autoimmunity in susceptible individuals. A syndrome similar to that of SLE was described in a patient administered sulfadiazine in 1945 by Hoffman (see Bigazzi, 1988). Sulfonamides were one of the first classes of drugs identified to induce an autoimmune response, while to date, over 40 other drugs have been associated with a similar syndrome.

Autoantibodies to red blood cells and autoimmune hemolytic anemia have been observed in patients treated with numerous drugs, including procainamide, chlorpropamide, captopril, cefalexin, penicillin, and methyldopa (Logue et al., 1970; Kleinman et al., 1984). Hydralazine- and procainamide-induced autoantibodies may also result in SLE. Approximately 20% of patients administered methyldopa for several weeks for the treatment of essential hypertension developed a dose-related titer and incidence of autoantibodies to erythrocytes, 1% of which presented with hemolytic anemia. Methyldopa does not appear to act as a hapten but appears to act by modifying erythrocyte surface antigens. IgG autoantibodies then develop against the modified erythrocytes.

D-penicillamine is used to treat patients with rheumatoid arthritis, to reduce excess cystine excretion in patients with cystinurias, and as a chelating agent for copper in patients with Wilson's disease. D-penicillamine can cause multiple forms of autoimmunity including SLE, myasthenia gravis, pemphigus, and autoimmune thyroiditis. This drug is thought to act as immunomodulator in patients by initiating or even potentiating anti-DNA antibody synthesis (Mach et al., 1986). The highly reactive thiol group may react with various receptors and biological macromolecules to induce autoantibodies. Long-term (many months) treatment has been shown to induce autoimmunity resulting in myasthenia gravis in 0.5% of patients (Bigazzi, 1988) and SLE in approximately 2% of patients as exhibited by varying degrees of joint pain, synovitis, myalgia, malaise, rash, nephritis, pleurisy, and neurological effects. In patients exhibiting myasthenia gravis, D-penicillamine may act to alter the acetylcholine receptors. Autoantibodies to acetylcholine receptors have been detected in these patients and have been shown to decrease gradually after drug withdrawal concomitant with reversibility of the clinical syndrome.

However, myasthenia gravis may persist for long periods of time after D-penicillamine therapy has ceased.

Although rare, cases of renal lupus syndrome and pemphigus blisters have also been reported as a consequence of D-penicillamine-induced immune complexes (Ntoso et al., 1986; Bigazzi, 1988), as well as with other drugs. With renal lupus syndrome, secondary glomerulonephritis may result if granular IgG antibodies are produced and deposited on the basement membranes. In patients with pemphigus blisters, autoantibodies to the intercellular substance of the skin have been recovered from the sera, and dermal biopsies have demonstrated intracellular deposits or immunoglobulin deposits on the basement membranes. Pemphigus has also been observed in patients treated with sulfhydryl compounds such as captopril and pyrithioxine (Bigazzi, 1988).

Some metals that are used therapeutically have also been shown to induce autoimmune responses. Gold salts used to treat arthritis may induce formation of antiglomerular basement membrane antibodies, which may lead to glomerulonephritis similar to that seen in Goodpasture's disease (see type II hypersensitivity). Since gold is not observed at the site of the lesions (Druet et al., 1982), it has been hypothesized that the metal elicits an antiself response. Lithium, used to treat manic depression, is thought to induce autoantibodies against thyroglobulin, which in some patients results in hypothyroidism. In studies with rats, levels of antibodies to thyroglobulin were shown to increase significantly in lithium-treated rats compared to controls immediately after immunization with thyroglobulin; however, rats that were not immunized with thyroglobulin did not produce circulation antithyroglobulin antibodies upon receiving lithium, and there was no effect of lithium on lymphocytic infiltration of the thyroid in either group (Hassman et al., 1985).

Some drugs such as penicillin have been shown to induce autoimmunity as well as anaphylaxis (Gleichmann et al., 1989). The carbonyl of the β -lactam ring of penicillin can form a covalent penicilloyl conjugate with nucleophilic sites on proteins, particularly the amino groups of lysine residues. This conjugate, which acts as the major immunogenic determinant, may become biotransformed to other isomeric forms of clinical relevance (Batchelor et al., 1965).

A genetic predisposition to drug-induced development of SLE has been shown to occur in some individuals treated with the drugs hydralazine, isoniazid, procainamide, and sulfamethazine. A polymorphism, which is known to exist for the genes responsible for expression of hepatic *N*-acetyltransferase enzymes, determines the rate of acetylation of these drugs to regulate the rate of drug inactivation. Individuals that are relatively slow acetylators of these drugs are more likely to develop antinuclear antibodies and are at a higher risk for developing SLE (Perry et al., 1970). Other

predisposing factors, such as HLA phenotype (HLA-DR4 and/or C4 allele), may also play a genetic role in determining susceptibility to hydralazine-induced SLE (Spears and Batchelor, 1987).

In addition, silicone-containing medical devices, particularly breast prostheses, have been reported to cause serum sickness-like reactions, scleroderma-like lesions, and an SLE-like disease termed human adjuvant disease (Guillaume et al., 1984; Kumagai et al., 1984). Some patients may also present with granulomas and autoantibodies. Human adjuvant disease is a connective tissue or autoimmune disease similar to that of adjuvant arthritis in rats and rheumatoid arthritis in humans. Autoimmune disease-like symptoms usually develop 2–5 years after implantation in a small percentage of people that receive implants, which may indicate that there is a genetic predisposition similar to that for SLE. An early hypothesis is that the prosthesis or injected silicone plays an adjuvant role by enhancing the immune response through increased macrophage and T-cell helper function. There is currently controversy as to whether silicone, as a foreign body, induces a nonspecific inflammation reaction, a specific cell-mediated immunological reaction, or no reaction at all. However, there is strong support to indicate that silicone microparticles can act as haptens to produce a delayed hypersensitivity reaction in a genetically susceptible population of people. It should be noted that there are currently no known drug-induced type I autoimmunities.

11.6 REGULATORY POSITIONS

The pharmaceutical and medical device industries are increasingly concerned with whether preclinical testing of their products should include routine immunotoxicologic screening or be done on an "as-needed basis," triggered by the toxicological profile of the xenobiotic established in routine preclinical safety testing (Bloom et al., 1987). The FDA has released a guideline for immunotoxicity testing of pharmaceuticals (CDER, 2006). Recent drug development efforts in the areas of biotechnology, prostaglandins, interleukins, and recombinant biological modifiers have elicited the expectation that the development of antibodies (neutralizing and otherwise) should be evaluated in at least one of the animal models used to assess general systemic toxicity. And more to the point, draft guidelines have been released for devices (Center for Devices and Radiological Health (CDRH), 1999). The other available guidances have been the draft guidelines in the revision of the "Redbook" (FDA, 1993) and the EMEA (2012). The ICH S8 guidance (2006) supersedes all of them (for pharmaceuticals) (Pattels and Taylor, 2008). Table 11.10 presents a summary comparison of different regulatory requirements for drugs.

TABLE 11.10 Comparison of Current ICH and Former European Union (EU) and US Immunotoxicity

| | ICH S8 (in Operation in All ICH Regions Since 2006) | U.S. FDA CDER (Still in Operation Since 2002) | EU CPMP (Still in Operation Since 2000) |
|--|--|---|---|
| Specific immunotoxicity guideline | Yes | Yes | No, included in guidance on repeat-dose toxicity |
| Drug-induced hypersensitivity, immunogenicity, and autoimmunity excluded | Yes | No, these categories are included in the guideline | Yes. (Note: Skin sensitizing potential addressed in CPMP Note for Guidance on Non-Clinical Local Tolerance Testing, 2001) |
| Screening study(ies) required | Yes, the initial screen for potential immunotoxicity involves standard toxicity studies (STSs) from short-term to chronic repeat-dose studies in rodents and nonrodents | Yes, including all standard repeat-dose toxicology studies that have been performed | Yes, screening required for all new active substances in at least one repeat-dose toxicity study (duration ideally should be 28 days). Rats or mice are species of choice |
| Screening study(ies) immunotoxicity parameters | Changes in hematology, lymphoid organ weights, histopathology of immune system, and serum globulins and increased incidences of infections and tumors should be evaluated for signs of immunotoxic potential in the STSs | Changes in hematology lymphoid organ weights, gross pathology and histopathology of immune system, and serum globulins and increased incident of infections and tumors should be evaluated for signs of immunotoxic potential | Hematology, lymphoid organ weights, histopathology of lymphoid tissues, bone marrow, cellularity, distribution of lymphocyte subsets, and NK-cell activity (if latter two unavailable, primary antibody response to T-cell-dependent antigen) |
| Other factors to consider in evaluation of potential immunotoxicity and the need for additional immunotoxicity studies | Pharmacological properties of drug; patient population; structural similarities to known immunomodulators; drug disposition; clinical data | Patient population; known drug class effects (including SARs); drug pharmacokinetics; clinical data. If drug intended for HIV, immune function studies required | None specifically included in the guideline |
| “Follow-on”/“Additional” immunotoxicity studies | “Additional” studies may be required depending on the “weight of evidence review” of STSs and “other factors.” “Additional” studies addressed in 3.2, 3.3, and Appendix of guideline | Not specified | Not specified |

The FDA had drafted a similar two-leveled approach (Hinton, 1992) for assessing immunotoxicity of food colors, additives, and drugs, and these are reflected in the S8 guidance. In all of these testing schemes, the initial tier generally includes a fundamental standard toxicity assessment with emphasis on histopathology of the major components of the immune system (Snyder, 2012). Additional tiers are then added to more precisely evaluate the functionality of the components that appeared to be adversely affected in the first tier of tests. These test strategies are primarily geared toward the detection of small molecule pharmaceutical-induced immunosuppression; thus, the effectiveness of these test schemes for detecting immunostimulation has still not been determined (Spreafico, 1988). Table 11.11 presents items that should be considered in such an initial evaluation and Table 11.12

presents guidance as to evaluating findings in clinical pathology parameters.

The FDA, ICH, and EMEA guidelines are not currently entirely aligned. The ultimate immune test would be to examine the effects of xenobiotics on the intact animal's response to challenge by viral, bacterial, or parasitic pathogens or neoplastic cells. The ability of the immune system to compensate or, conversely, its inability to compensate for loss or inhibition of its components is fully examined through host-resistance mechanisms. This tiered test approach has been validated with 50 selected compounds, and results from these studies have shown that the use of only two or three immune tests is sufficient to predict known immunotoxic compounds in rodents with a greater than 90% concordance (Luster et al., 1992a, b). Specifically, the use of either a humoral response assay for plaque-forming

TABLE 11.11 Typical Indicators of Immunotoxicity, Which May Be Observed during Regulatory Repeat-Dose Toxicity Studies

| Findings | Possible Indicator of |
|--|--|
| <i>During the in-life phase</i> | |
| Increased frequencies of infectious disease | Immunosuppression |
| Increased frequencies of tumors in long-term studies in the absence of genotoxicity or nongenotoxic indicators of tumorigenicity (e.g., endocrine) | Immunosuppression |
| Unexpected pathological symptoms or deaths shortly after administration | Hypersensitivity |
| Strong inflammatory reactions at the site of administration | Hypersensitivity |
| <i>Gross necropsy</i> | |
| Significant increase or decrease of size and weight of lymphatic organs | Unintended immunostimulation or immunosuppression |
| <i>Hematology</i> | |
| Changes in total or differential blood counts | Unintended immunostimulation or immunosuppression |
| Anemia | Type II hypersensitivity |
| Altered frequencies of lymphocyte subsets (flow cytometry) ^a | Unintended immunostimulation or immunosuppression |
| <i>Clinical chemistry</i> | |
| Altered total globulin levels or albumin/globulin ratio | Unintended immunostimulation or immunosuppression |
| Changes of immunoglobulin isotype levels ^a | Unintended immunostimulation or immunosuppression |
| Reduction of hemolytic complement activity ^a | Unintended immunostimulation type III hypersensitivity |
| Antinuclear or anticytoplasmic antibodies ^a | Unintended immunostimulation or autoimmunity |
| <i>Histopathology</i> | |
| Changes of cellularity and/or microanatomy of lymphatic organs | Unintended immunostimulation or immunosuppression |
| Vasculitis, glomerulonephritis | Type III hypersensitivity |

^a This parameter is normally not measured during standard toxicity studies but may be integrated when a focus is drawn on immunotoxicity assessment.

colonies (PFC response) or determination of surface marker expression in combination with almost any other parameter significantly increased the ability to predict immunotoxicity when compared to the predictivity of any assay alone.

The FDA guidelines for immunotoxicity testing of food additives start with a type I battery of tests. Type I tests can be derived from the routine measurements and examinations performed in short-term and subchronic rodent toxicity studies, since they do not require any perturbation of the test animals (immunization or challenge with infectious agents). These measurements include hematology and serum chemistry profiles, routine histopathologic examinations of immune-associated organs and tissues, and organ and body weight measurements including thymus and spleen. If a compound produces any primary indicators of immunotoxicity from these measurements, more definitive immunotoxicity tests, such as those indicated in the preceding paragraph, may be recommended on a case-by-case basis.

The following is a brief explanation of some of the indicators that may be used to trigger additional definitive testing and a description of some of the most commonly used assays to assess humoral, cell-mediated, or nonspecific immune dysfunction, which are common to most immunotoxicology test strategies.

11.6.1 CDER Guidance for Investigational New Drugs

CDER's 2001 promulgated draft guidance for pre-INDA immunotoxicity clearly established the framework for FDA's approach, which was finalized under ICH S8 (2006). It begins by characterizing five adverse event categories:

1. Immunosuppression
2. Antigenicity
3. Hypersensitivity
4. Autoimmunity
5. Adverse immunostimulation

Specific tests are proposed for each of these categories. It notes that immune system effects in nonclinical toxicology studies are often attributed and written off as due to stress (Aden and Cohen, 1993). Such effects are frequently reversible with repeat dosing and tend not to be dose related. It is also proposed that when possible dose extrapolations to those in clinical use be based on relative body area. Specific recommendations are made for when to conduct specific testing (as opposed to the broader general evaluations integrated into existing repeat-dose testing) (Figure 11.2) and for follow-up studies for exploring mechanisms (Figure 11.3).

TABLE 11.12 Evaluation of Clinical Pathology

| Observation | Possible Considerations |
|----------------------|---|
| Hyperglobulinemia | Clinical chemistry Inflammation Infection Certain lymphoid tumors |
| Hypoalbuminemia | Immune-mediated glomerulonephritis (when present with hypercholesterolemia) |
| Hypercholesterolemia | Immune-mediated glomerulonephritis (when present with hypoalbuminemia) Hematology |
| Pancytopenia | Direct toxicity to bone marrow architecture or stem cells Antimitotic effect Effects on differentiation and/or maturation Immune-mediated (antibody- or cell-mediated) destruction of stem cells |
| Monocytes | |
| Morphology | Activated appearance (vacuolated, “fluffy”) Infection/sepsis (intracellular microorganisms) |
| Monocytosis | Stress (dogs and primates) Necrosis (immune-mediated or nonimmune) Inflammation (immune-related or nonimmune) Immune-mediated hemolytic anemia or thrombocytopenia Altered production or turnover |
| Lymphocytes | |
| Lymphopenia | Stress Direct toxicity (peripheral or bone marrow) Antiproliferative effect or signal Immunosuppression Infection/acute inflammation |
| Lymphocytosis | Excitability/fear (fight or flight syndrome) Immunostimulation Lymphoid neoplasia Chronic inflammation |
| Neutrophils | |
| Morphology | Bone marrow (asynchronous maturation) Infection/sepsis (intracellular microorganisms) Trafficking/prolonged presence in circulation (hypersegmentation) Inflammation (toxic change—vacuolation, basophilic stippling, Dohle bodies) |
| Neutropenia | Direct bone marrow suppression (including cytokine alterations) Tissue demand exceeds production capacity (inflammation, infection) Increased margination/trafficking to tissues Immune-mediated destruction Inhibition of release from bone marrow |
| Neutrophilia | Adrenaline-induced (excitation; fear) Corticosteroid-induced (stress) Inflammation (immune-relation or nonimmune) Immune-mediated hemolytic anemia or thrombocytopenia Effects on trafficking (defect in adhesion molecules/tissue migration) |
| Eosinophils | |
| Eosinophilia | Hypersensitivity Fungal infection, parasitemia Tumors (e.g., mast cell, basophilic or eosinophilic) |
| Eosinopenia | Stress General bone marrow suppression Suppression/alteration of differentiation Excitation/fear Acute infection/inflammation |
| Basophils | |
| Basophilia | Persistent lipemia Parasitemia Allergy |

(Continued)

TABLE 11.12 (Continued)

| Observation | Possible Considerations |
|-------------------------------|---|
| Erythrocytes Anemia | Altered lipid content, oxidative damage (nonimmune effects) Bone marrow suppression (nonregenerative) Monocyte activation Immune-mediated destruction Intra- or extravascular hemolysis (immune or nonimmune) Infection/inflammation (chronic disease) |
| Platelets Thrombocytopenia | Immune-mediated (often regenerative) Bone marrow suppression (nonregenerative) Disseminated intravascular coagulation (DIC) |

Source: Data from Evans (2008) and Stockholm and Scott (2008).
Considerations are not intended to be inclusive of all possibilities.

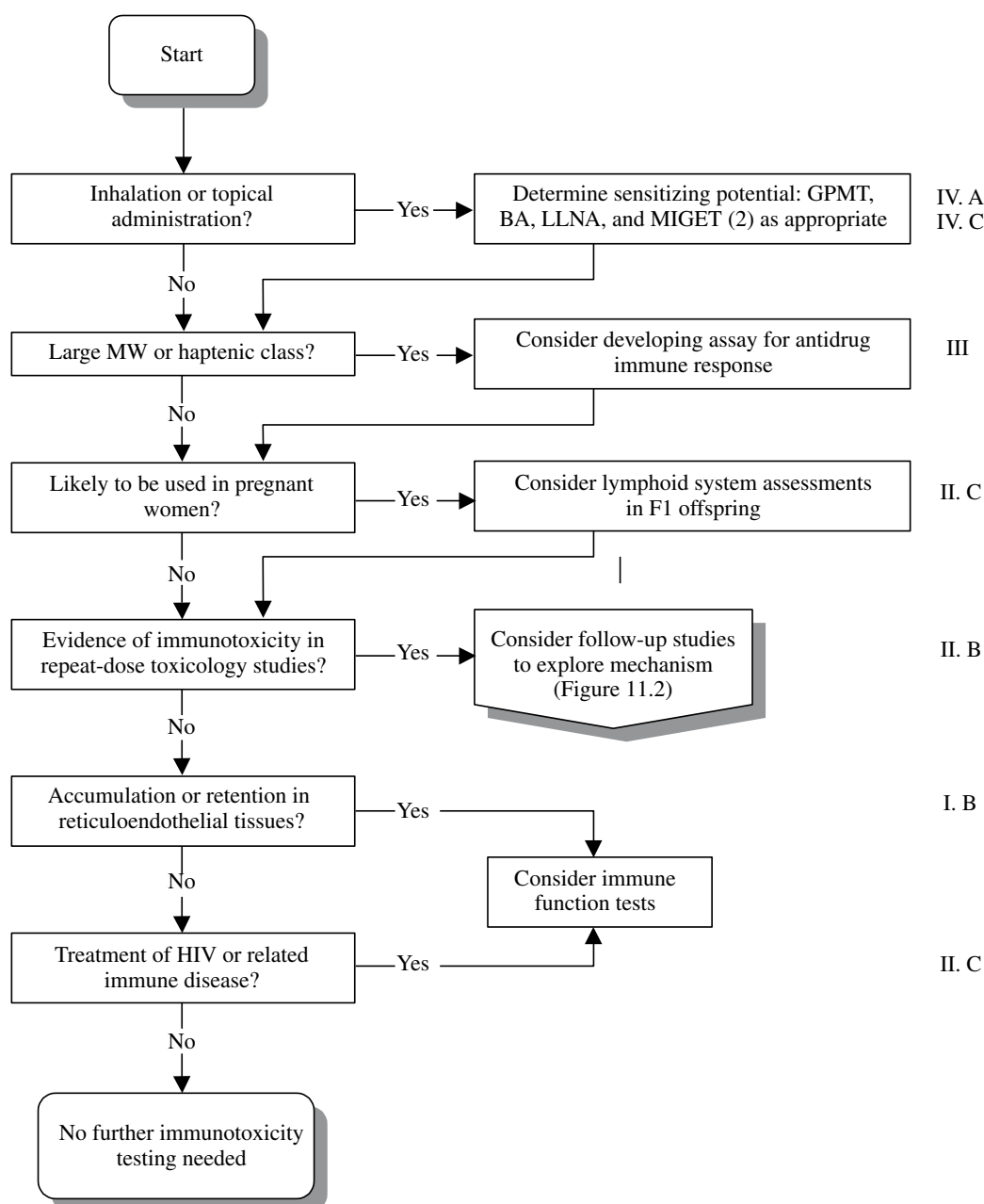


FIGURE 11.2 CDER flowchart for determining when to conduct specific immunotoxicity testing. Annotations in right margin indicate location of text describing specific advice. GPMT, guinea pig maximization test; BA, Buehler assay (Buehler patch test); LLNA, local lymph node assay; MIGET, mouse IgE test. (There is only a relatively small database available for assessing the usefulness of the MIGET for drug regulatory purposes.)

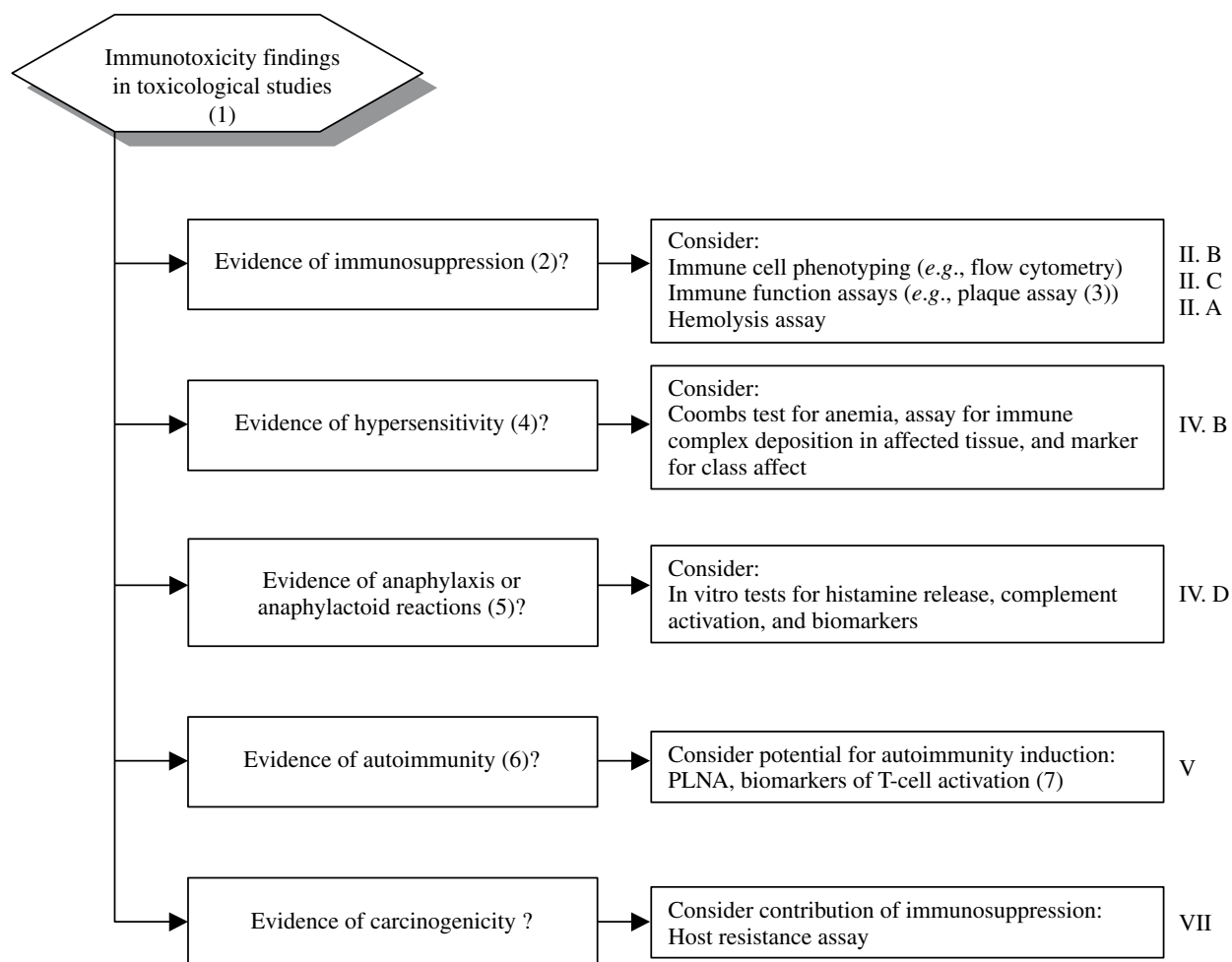


FIGURE 11.3 Follow-up studies to consider for exploring mechanisms of immunotoxicity. Annotations in right margin indicate location of text describing specific advice. (1) Examples include myelosuppression, histopathology in immune-associated tissues, increased infection, tumors, decreased serum Ig, and phenotypic changes in immune cells. (2) Other acceptable assays include drug effect on NK-cell function *in vitro* blastogenesis, cytotoxic T-cell function cytokine production, DTH, and host resistance to infections or implanted tumors. (3) Examples include anemia, leukopenia, thrombocytopenia, pneumonitis, vasculitis, lupus-like reactions, and glomerulonephritis. (4) Examples include cardiopulmonary distress, rashes, flushed skin, and swelling of face or limbs. (5) Examples include vasculitis, lupus-like reactions, glomerulonephritis, and hemolytic anemia. (6) There are no established assays that reliably assess potential for autoimmunity and acute systemic hypersensitivity. (7) The popliteal lymph node assay (PLNA) has only a relatively small database available for assessing its usefulness for drug regulatory purposes.

11.7 EVALUATION OF THE IMMUNE SYSTEM

The ICH/FDA CDER guidelines for immunotoxicity testing of small molecule pharmaceuticals (ICH S8) start with evaluation of parameters evaluated in repeat-dose (typically 28 days) systemic toxicity studies. These tests can generally be derived from the routine measurements and examinations performed in short-term and subchronic rodent and nonrodent toxicity studies, since they do not require any perturbation of the test animals (immunization or challenge with infectious agents). These measurements include hematology and serum chemistry profiles, routine histopathologic examinations of immune-associated organs and tissues, and organ and body weight measurements including

thymus and spleen. If a compound produces any primary indicators of immunotoxicity from these measurements, more definitive immunotoxicity tests, such as those indicated in the preceding paragraph, may be recommended on a case-by-case basis. While detection of up-modulated immune responses is believed to be effective in current nonclinical testing approaches, detection of immune-suppressive responses (leading to greater susceptibility to infection) is not felt to be as well served (Germolec, 2004). This has certainly proved to be the case of induction of progressive multifocal leukoencephalopathy (PM2, a usually fatal viral infection of the white matter of the brain) in some cases associated with the use of several immune-suppressive monoclonal antibodies.

The following is a brief explanation of some of the indicators that may be used to trigger additional definitive testing and a description of some of the most commonly used assays to assess humoral, cell-mediated, or nonspecific immune dysfunction, which are common to most immunotoxicology test strategies.

11.7.1 Immunopathologic Assessments

Various general toxicological and histopathologic evaluations of the immune system can be made as part of routine preclinical safety testing to obtain a preliminary assessment of potential drug-related effects on the immune system. At necropsy, various immunological organs of the immune system such as thymus, spleen, and lymph nodes are typically observed for gross abnormalities and weighed in order to detect decreased or increased cellularity. Bone marrow and peripheral blood samples are also taken to evaluate abnormal types and/or frequencies of the various cellular components. Tables 11.13 and 11.14 summarize the observations and interpretations.

11.7.1.1 Organ and Body Weights Changes in absolute weight, organ-to-body weight ratios, and organ-to-brain weight ratios of tissues such as thymus and spleen are useful general indicators of potential immunotoxicity. However, these measures are nonspecific for immunotoxicity since they may also reflect general toxicity and effects on endocrine function that can indirectly affect the immune system.

11.7.2 Humoral (Innate) Immune Response and Possible Entry Points for Immunotoxic Actions

11.7.2.1 Hematology Hemacytometers or electronic cell counters can be used to assess the numbers of lymphocytes, neutrophils, monocytes, basophils, and eosinophils in the peripheral blood, while changes in relative ratios of the various cell types can be assessed by microscopic differential evaluation. Similar evaluations can be performed with bone marrow aspirates, where changes may reflect immunotoxicity to the pluripotent stem cells and newly developing lymphoid precursor cells. Potential hematological indicators of

TABLE 11.13 Examples of Antemortem and Postmortem Findings That May Include Potential Immunotoxicity if Treatment Related

| Parameter | Possible Observation (Cause) | Possible State of Immune Competence |
|-------------------------|---|-------------------------------------|
| <i>Antemortem</i> | | |
| Mortality | Increased (infection) | Depressed |
| Body weight | Decreased (infection) | Depressed |
| Clinical signs | Rales, nasal discharge (respiratory infection) | Depressed |
| | Swollen cervical area (sialodacryoadenitis virus) | Depressed |
| Physical examinations | Enlarged tonsils (infection) | Depressed |
| Hematology | Leukopenia/lymphopenia | Depressed |
| | Leukocytosis (infection/cancer) | Enhanced/depressed |
| | Thrombocytopenia | Hypersensitivity |
| | Neutropenia | Hypersensitivity |
| Protein electrophoresis | Hypogammaglobulinemia | Depressed |
| | Hypergammaglobulinemia (ongoing immune response or infection) | Enhanced/activated |
| <i>Postmortem</i> | | |
| Organ weights | | |
| Thymus | Decreased | Depressed |
| Histopathology | | |
| Adrenal glands | Cortical hypertrophy (stress) | Depressed (secondary) |
| Bone marrow | Hypoplasia | Depressed |
| Kidney | Amyloidosis | Autoimmunity |
| | Glomerulonephritis (immune complex) | |
| Lung | Pneumonitis (infection) | Depressed |
| Lymph node | Atrophy | Depressed |
| Spleen | Hypertrophy/hyperplasia | Enhanced/activated |
| | Depletion of follicles | Depressed B cells |
| | Hypocellularity of periarteriolar sheath | Depressed T cells |
| | Active germinal centers | Enhanced/activated |
| Thymus | Atrophy | Depressed |
| Thyroid | Inflammation | Autoimmunity |

TABLE 11.14 Immune System Components in Organ Sites

| Organ | Antigen-Presenting Cell | B Cell | CD4 ⁺ T Cell | Possible Immunotoxic Entry Point | Outcome |
|-------------------|---|--|---|---|--|
| Bone marrow | Differentiation from pluripotent stem cells | Differentiation from pluripotent stem cells: rearrangement of antigen receptors, central repertoire selection, homing, antibody production | Differentiation from pluripotent stem cells: rearrangement of antigen receptors | Inhibition of cell proliferation, interference with signal transduction, genotoxicity, interference with gene recombination, interference with DNA repair, interference with signal transduction, interference with apoptosis, interference with signal transduction, inhibition of protein synthesis | Leukopenia, malignancies, lymphopenia, selective lymphopenia, autoimmunity, impaired adaptive immune response |
| Thymus | — | — | Proliferation and positive/negative repertoire selection | Interference with signal transduction, interference with apoptosis | Selective lymphopenia, autoimmunity |
| Peripheral tissue | Antigen uptake and processing, maturation | — | — | Inhibition of cell motility, interference with phagolysosomal degradation, interference with intracellular transport, inhibition of cell motility, interference with signal transduction | Lack of antigen presentation, impaired adaptive immune response, lack of antigen presentation, impaired adaptive immune response |
| Lymphatic organ | Antigen presentation | Antigen uptake and presentation, cross-communication with T-cell, activation, proliferation, differentiation | Antigen recognition (APC) and activation; proliferation; differentiation, cross-communication with B-cell | Inhibition of cell motility, interference with phagolysosomal degradation or intracellular transport, interference with signal transduction, inhibition of cytokine release, interference with signal transduction, inhibition of cell proliferation, interference with signal transduction | Impaired adaptive immune response, impaired adaptive immune response, autoimmunity; allergy, impaired adaptive immune response |

Source: Adapted from Haley et al. (2005, pp. 404–407).

immunotoxicity include altered white blood cell counts or differential ratios, lymphocytosis, lymphopenia, or eosinophilia. Changes in any of these parameters can be followed up with more sophisticated flow cytometric analyses or immunostaining techniques that are useful for phenotyping the various types of lymphocytes (B cell, T cell) and the T-cell subsets (CD4⁺ and CD8⁺) on the basis of unique surface markers. Decreases or increases in the percentages of any of the cell populations relative to controls or in the ratios of B cells/T cells or CD4⁺/CD8⁺ cells may be indicators of immunotoxicity.

11.7.2.2 Clinical Chemistry Nonspecific clinical chemistry indicators of potential immune dysfunction include changes in serum protein levels in conjunction with changes in the albumin-to-globulin (A/G) ratio. Immunoelectrophoretic analysis of serum proteins can then be performed to quantify the relative percentages of albumin and the α -, β -, and γ -globulin fractions. To perform these assays, a drop of serum (antigen) is placed into a well cut in a gel, and then the gel is subjected to electrophoresis so that each molecule in the serum moves in the electric field according to its charge. This separation is then exposed to specific antiserum, which is placed in a trough cut parallel to the direction in which the components have moved. By passive diffusion, the antibody reaches the electrophoretically separated antigen and reacts to form Ag-Ab complexes. The γ -globulin fractions can be separated and further quantified for the relative proportions of IgG, IgM, IgA, and IgE using similar techniques.

Serum concentrations of immunoglobulin classes and subclasses can also be measured using various techniques such as radioimmunoassays (RIAs) or enzyme-linked immunosorbent assays (ELISAs). In the ELISA, antigens specific for each class of immunoglobulin can be adsorbed onto the surfaces for microtiter plates. To determine the quantity of each antibody in a test sample, an aliquot of antiserum is allowed to react with the adsorbed antigens. Unreacted molecules are rinsed off and an enzyme-linked anti-Ig is then added to each well. Next, substrate is added and the amount of color that develops is quantified using a spectrophotometric device. The amount of antibody can then be extrapolated from standard curves since the amount of color is proportional to the amount of enzyme-linked antibody that reacts. Variations in levels of a given antibody may indicate the decreased ability of B cells or decreased numbers of B cells producing that antibody. In addition, serum autoantibodies to DNA, mitochondria, and parietal cells can be used to assess autoimmunity. Serum cytokines (IL-1, IL-2, and γ -interferon) can also be evaluated using immunochemical assays to evaluate macrophage, lymphocyte, and lymphokine activity; prostaglandin E₂ can also be measured to evaluate macrophage function.

CH50 determinations can be used to analyze the total serum complement and are useful for monitoring immune

complex diseases (Sullivan, 1989); activation of complement (Table 11.14) in the presence of autoantibodies is indicative of immune complex diseases and autoimmunity. The various components of the complement system (C3, C4) can also be measured to assess the integrity of the system. For instance, low serum concentrations of C3 and C4, with a concomitant decrease in CH50, may indicate activation of complement, while a low C4 alone is a sensitive indicator of reduced activation of the complement system. Since C3 is used as an alternate complement pathway, it usually measures high. Therefore, a low C3 with a normal C4 may indicate an alternate pathway of activation.

11.7.2.3 Histopathology Histopathologic abnormalities can be found in lymphoid tissues during gross and routine microscopic evaluations of the spleen, lymph nodes, thymus, bone marrow, and gut-associated lymphoid tissues such as Peyer's patches and mesenteric lymph nodes. Microscopic evaluations should include descriptive qualitative changes such as types of cells, densities of cell populations, proliferation in known T- and B-cell areas (e.g., germinal centers), relative numbers of follicles and germinal centers (immune activation), and the appearance of atrophy or necrosis. In addition, unusual findings such as granulomas and scattered, focal mononuclear cell infiltrates in nonlymphoid tissues may be observed as indicators of chronic hypersensitivity or autoimmunity. A complete histopathologic evaluation should also include a quantitative assessment of cellularity through direct counts of each cell type in the various lymphoid tissues. In addition, changes in cellularity of the spleen can be more precisely quantitated from routine hematoxylin and eosin (H&E) sections using morphometric analysis of the germinal centers (B cells) and periarteriolar lymphocyte sheath (T cells). Similar morphometric measurements can be made of the relative areas of the cortex and medulla of the thymus. If changes in cellularity are apparent from routinely stained histopathology sections, special immunostaining (immunoperoxidase or immunofluorescence) of B cells in the spleen and lymph nodes using polyclonal antibodies to IgG, or immunostaining of the T cells and their subsets in the spleen using mono- or polyclonal antibodies to their specific surface markers, can be used to further characterize changes in cellularity.

Numerous physiological and environmental factors such as age, stress, nutritional deficiency, and infections may affect the immune system (Sullivan, 1989). Thus, adverse findings in animal studies may reflect these indirect immunotoxic effects rather than the direct immunotoxic potential of a chemical or drug. Indirect immunotoxic effects may be assessed through histopathologic evaluations of endocrine organs such as the adrenals and pituitary.

It is also well known that the functional reserves of the immune system can allow biologically significant, immunotoxic insults to occur without the appearance of morphological

changes. In addition, there is some built-in redundancy in the system in that several mechanisms may produce the same outcome. For instance, cytotoxic T cells may alone be sufficient to protect the organism against a bacterial infection; however, the body will also produce antibodies for future protection. Thus, if one mechanism is insufficient to fight off infection, the second mechanism can serve as a backup. Because of this functional reserve, adverse effects may remain subclinical until the organism is subjected to undue stress or subsequent challenge (Bloom et al., 1987). Therefore, routine immunopathologic assessments as part of standard preclinical toxicity tests may not be sufficient to detect all immunotoxins. Although changes detected in routine toxicological and pathological evaluations are nonspecific and of undetermined biological significance to the test animal, they can be invaluable as flags for triggering additional testing.

As described previously, the humoral immune response results in the proliferation, activation, and subsequent production of antibodies by B cells following antigenic exposure and stimulation. The functionality and interplay between the three primary types of immune cells (macrophage, B cells, and T cells) required to elicit a humoral response can be assessed through various *in vitro* assays using cells from the peripheral blood or lymphoid tissues.

11.7.2.4 Antibody Plaque-Forming Cell (PFC) Assay

The number of B cells producing antibody (PFC) to a T-dependent antigen such as sheep red blood cells (SRBCs) can be assessed *in vitro* following *in vivo* exposure to the test article and antigen (*ex vivo* tests). The PFC response to a T-dependent antigen is included as a tier I test by the NTP since it appears to be the most commonly affected functional parameter of exposure to immunosuppressants. However, this test is designated as a type II test in the FDA Redbook since it requires an *in vivo* immunization of the animals with antigen and, thus, cannot be evaluated as part of an initial toxicity screen.

Although this assay requires that B cells be fully competent in secreting antibodies, T cells and macrophage cells are also essential for the proper functioning of humoral immunity. However, this assay is nonspecific in that it cannot determine which cell type(s) is responsible for dysfunction. Macrophage cells are needed to process antigen and produce IL-1. T cells are needed for several functions including antigen recognition of surface membrane proteins and B-cell maturation through the production of various lymphokines that stimulate growth and differentiation. SRBCs are most commonly used as the T-dependent antigen, although T-cell-independent antigens may also be useful to rule out T helper dysfunction as a cause of immunodysfunction.

The PFC assay has evolved from methodology originally developed as a hemolytic plaque assay (or Jerne plaque assay) by Nils Jerne to quantitate the number of antibody-forming cells in a cell suspension plated with RBCs onto

agar plates (Jerne and Nordin, 1963). In its present form, animals are treated *in vivo* with the test compound, immunized with approximately 5×10^8 SRBCs administered intravenously within 2–3 days posttreatment, and then sacrificed 4 days (IgM) or 6 days (IgG) later. Antibody-producing spleen cell suspensions are then mixed *in vitro* with SRBCs, placed onto covered slides, and incubated for a few hours in the presence of complement. During incubation, antibody diffuses from the anti-SRBC-producing cells and forms Ag–Ab complexes on the surfaces of nearby SRBCs. In the presence of complement, the Ag–Ab complexes cause lysis of the SRBCs, resulting in the formation of small, clear plaques on the slide. Plaques are then counted and expressed as PFCs/ 10^6 spleen cells. A dose-related reduction in PFCs is indicative of immunosuppression.

11.7.2.5 B-Cell Lymphoproliferation Response

The NTP has classified this assay as a tier I test since mitogenesis can be performed easily in tandem with other tests to provide an assessment of the proliferative capacity of the cells (Luster et al., 1988). Since this assay is performed *ex vivo* with peripheral blood (or spleen) and is well characterized for use in various animal species, it has also been included as an expanded type I test in the revised Redbook.

The proliferation of peripheral blood or splenic B cells following stimulation with lipopolysaccharide (LPS) or other mitogens (pokeweed mitogen extract) is another measurement of humoral immunity. LPS (a bacterial lipopolysaccharide) is a B-cell-specific mitogen that stimulates polyclonal proliferation (mitosis) as part of the natural sequence of antigen recognition, activation, and clonal expansion. The mitogen does not interact with just one particular antigen-specific clone but with all cells bearing the carbohydrate surface marker for which it is specific. Since mitogens are both polyclonal and polyfunctional, they can stimulate a wider spectrum of antigenic determinants than antigens, which can only stimulate a low number (10^{-6}) of specific cells.

In this assay, lymphocytes from animals are treated *in vivo* and cultured *in vitro* in microtiter plates in the presence of tritiated [^3H]thymidine (or uridine) using a range of at least three concentrations of mitogen to optimize the response. Lymphocytes can be obtained aseptically from peripheral blood or from single cell suspensions of spleen cells that are prepared by pushing the tissue through sterile gauze or 60-mesh wire screens. A decrease in DNA synthesis (incorporation of ^3H) as compared to the unexposed cells of control animals may indicate that the B cells were unable to respond to antigenic stimulation. Alternative methodology employs a 18–20h incubation with ^{125}I -labeled iododeoxyuridine ([^{125}I]IudR) and fluorodeoxyuridine (FudR) (White et al., 1985). After incubation, the cells are collected onto filter disks and then counted with a gamma counter.

Assays such as this that use polyclonal mitogens for activation may not be as sensitive as specific antigen-driven systems (Luster et al., 1988). In addition, suppression of the mitogen response does not always correlate with the PFC response. Since mitogenesis represents only a small aspect of B-cell function and maturation, this end point is not sensitive to early events that may affect activation or later events that may affect differentiation of B cells into antibody-secreting cells (Klaus and Hawrylowicz, 1984).

11.7.3 Cell-Mediated Immunity

11.7.3.1 T-Cell Lymphoproliferation Response This assay is analogous to the B-cell lymphoproliferative response assay described previously. Thus, this assay is also classified as a tier I test by the NTP and as an expanded type I test in the revised draft of the Redbook.

T cells from the peripheral blood or spleen undergo blastogenesis and proliferation in response to specific antigens that evoke a cell-mediated immune response. T-cell proliferation is assessed using T-cell-specific mitogens such as the plant lectins, concanavalin A (Con A), and phytohemagglutinin (PHA) or T-cell-specific antigens (i.e., tubercin, *Listeria*). Uptake of ^3H as an indicator of DNA synthesis is used as described previously for evaluating B-cell proliferation. T-cell mitogens do not just stimulate synthesis of DNA but, in fact, they also stimulate the expression of cell-specific function. For instance, Con A can trigger the expression of T helper, suppressor, and cytotoxic effector cells, and either mitogen may induce the expression (or reexpression of memory cells) of differentiated function (Clark, 1983). Since cell populations responsive to Con A are thought to be relatively immature compared to those that are stimulated with PHA, the parallel usage of both mitogens may be useful for distinguishing the affected subset (Tabo and Paul, 1973). A secondary response to T-cell antigens such as purified protein derivative of tuberculin (PPD) or tetanus toxoid can also be assessed.

11.7.3.2 Mixed Lymphocyte Response (MLR) Assay This assay has been shown to be sensitive for the detection of chemical-induced immunosuppression and is a recommended tier I assay by the NTP (Luster et al., 1988). In addition, it has been shown to be predictive of host response to transplantation and of general immunocompetence (Harmon et al., 1982).

The mixed lymphocyte response (MLR) assay assesses the ability of T cells to recognize foreign antigens on allogeneic lymphocytes and, thus, is an indirect measure of the cell-mediated ability to recognize graft or tumor cells as foreign. Responder lymphocytes from animals treated *in vivo* with the test compound are mixed with allogeneic stimulator lymphocytes that have been treated *in vitro* with mitomycin C or irradiated to render them unable to respond

(Bach and Voynow, 1966). Both cell types are cultured *in vitro* for 3–5 days and then incubated with ^3H for an additional 6 h. Once the radiolabel is incorporated into the DNA of the responding cells, the DNA is extracted and the amount of radioactive label is measured to quantitate proliferation of the responder cells of drug-treated animals compared to those of the controls.

11.7.3.3 Cytotoxic T Lymphocyte (CTL)-Mediated Assay

This assay is similar to the MLR assay and can be performed in parallel or as a tier II follow-up to the MLR assay.

The cytotoxic T Lymphocyte (CTL) assay ascertains the ability of cytotoxic T cells to lyse an allogeneic target cell or the specific target cell type with which they were immunized. In general, the cytolytic response of activated effector cells is assessed by measuring the amount of radioactivity (^{51}Cr) that is released from the target cell. When performed in conjunction with the MLR assay, lymphoid cells of the two strains are cultured together *in vitro* as described previously; however, ^{51}Cr is added to the culture after 4–5 days (instead of ^3H). Both responder and target cells are labeled with the ^{51}Cr , which is taken up rapidly by the cells through passive diffusion but is released slowly as long as the cell membrane is intact. Furthermore, since chromium is reduced from Cr^{6+} to Cr^{3+} and since Cr^{3+} enters the cells at a much slower rate than Cr^{6+} , the ^{51}Cr released from the damaged target cells is not significantly reincorporated into undamaged cells (Clark, 1983), which would reduce the sensitivity of the assay. Thus, the amount of chromium released into the medium and recovered in the supernatant of the mixture of the cells is directly proportionate to the extent of lysis of the target cells by the sensitized responder cells.

In a capillary tube assay developed in 1962 by George and Vaughan, the inhibition of migration of macrophage cells can be used to assess normal T-cell function (see Clark, 1983). T cells are obtained from the peripheral blood of animals treated *in vivo* with a test article and injected with antigen (e.g., tuberculin). These T cells are functioning normally, and they should release migration inhibition factor (MIF). As a consequence, the macrophages, which generally show a propensity for migration upon stimulation with the antigen, should show a MIF-induced reduction in migratory behavior.

11.7.3.4 Delayed-Type Hypersensitivity (DTH) Response

The DTH response assay is considered to be a comprehensive tier II assay for cell-mediated immunity by the NTP.

To express a DTH inflammatory response, the immune system must be capable of recognizing and processing antigen, blastogenesis and proliferation of T cells, migration of memory T cells to the challenge site of exposure to antigen, and subsequent production of inflammatory mediators and lymphokines that elicit the inflammatory response. Thus, by measuring a DTH response to an antigen, these assays assess

the functional status of both the afferent (antigen recognition and processing) and efferent (lymphokine production) arms of cellular immunity. Various antigens have been used for assessing DTH, including keyhole limpet hemocyanin (KLH), oxazolone, dinitrochlorobenzene, and SRBCs (Vos, 1977; Godfrey and Gell, 1978; Luster et al., 1988).

In one such assay described by White et al. (1985), mice previously treated with the test article are sensitized to SRBCs by inoculation of SRBCs into the hind footpad and 4 days later challenged in the same footpad. Seventeen hours following challenge, they are injected intravenously with ^{125}I -labeled human serum albumin (HSA) and then sacrificed 2 h later. Both hind feet are then radioassayed in a gamma counter (the second foot serves as a control for background infiltration of the label). With a normal functioning cell-mediated response, ^{125}I -labeled HSA will extravasate into the edematous area produced by the DTH response (Paranjpe and Boone, 1972). In general, a decrease in the extravasation of ^{125}I -labeled HSA is indicative of immunosuppression of the efferent arm of the cell-mediated immune system.

To assay specifically the afferent arm of the DTH response, the proliferation of the popliteal lymph node cells to SRBCs can also be measured (White et al., 1985). As described previously, mice treated with the test article are sensitized to SRBCs by inoculation of SRBCs into the hind footpad. However, 1.5 h later they are challenged intraperitoneally with FUdR and 2 h later they are administered [^{125}I]IUdR intravenously (instead of ^{125}I -labeled HSA). Mice are sacrificed 24 h after challenge and both popliteal lymph nodes are removed and counted in a gamma counter.

Similar assays for DTH have been traditionally performed with the antigen *Mycobacterium tuberculosis*, which preferentially elicits a cell-mediated response. In this assay a small amount of antigen contained in the supernatant fluid from the medium in which the pathogen was grown is injected into the footpad. Upon challenge, a visible and palpable lump should appear by 48 h. The amount of swelling is then measured and compared with the footpad that did not receive the challenge. Alternatively, methods used by the NTP employ a modified ^{125}I -labeled uridine (UdR) technique to measure the monocyte influx at the challenge site (ear) injected with KLH antigen. This assay has been shown to correlate well with decreased resistance to infectious disease (Luster et al., 1988). However, one should note that regardless of which technique is used, anti-inflammatory drugs may produce false-positive results in this type of assay.

Anaphylaxis is an acute allergic response currently seen in treatment with some biotherapeutics as well as to latex proteins (from gloves and other medical devices), bee stings, peanuts, and some surfactants. It requires exposure to a large molecule (3000 or greater MW), which can be either innately large molecules or small molecules bound to macromolecule carriers.

Unlike a type IV delayed constant hypersensitivity response, this type I response can affect the entire body, rapidly becoming more severe and even life threatening. There is rapid elevation of serum histamine and/or tryptase levels or agent specific IgE antibodies. While this is much more dangerous than a type IV topical response, there are currently no regulatorily mandated or accepted methods for predicting the potential effect from a drug or agent, though the Japanese previously had three such test protocols (Maki, 1997; Verdier et al., 1994). Passive cutaneous anaphylaxis (PCA) (Brocklehurst et al, 1960), active cutaneous anaphylaxis (ACA), and active systemic anaphylaxis (ASA) (Chazal et al, 1994; Choquet-Kastylevsky and Descotes, 2001). But while effective in predicting protein (large molecule) induced responses, they have been deemed not reliable for predicting such responses to small molecules (CDER, 2002).

11.8 NONSPECIFIC IMMUNITY FUNCTION ASSAY

11.8.1 Natural Killer Cell Assays

This assay is a tier I test for nonspecific immunity in the NTP testing scheme (Luster et al., 1988) and is proposed as an additional type I test in the draft Redbook.

NK cells, like cytotoxic T cells, have the ability to attack and destroy tumor cells or virus-infected cells. However, unlike T cells, they are not antigen specific, do not have unique, clonally distributed receptors, and do not undergo clonal selection. In *in vitro* or *ex vivo* tests, target cells (e.g., YAC-1 tumor cells) are radiolabeled *in vitro* or *in vivo* with ^{51}Cr and incubated *in vitro* with effector NK cells from the spleens of animals that had been treated with a xenobiotic. This assay can be run in microtiter plates over a range of various ratios of effector/target cells. Cytotoxic activity is then measured by the amount of radioactivity released from the damaged tumor cells as was previously described for cytotoxic T cells. This assay can also be performed *in vivo*, where YAC-1 cells labeled with [^{125}I]IUdR are injected directly into mice and NK-cell activity is correlated with its level of radioactivity (Riccardi et al., 1979). Immunotoxicity observed as reduced NK-cell activity is correlated with increased tumorigenesis and infectivity.

11.8.2 Macrophage Function

Several assays are available to measure various aspects of macrophage function, including quantitation of resident peritoneal cells, antigen presentation, cytokine production, phagocytosis, intracellular production of oxygen free radicals (used to kill foreign bodies), and direct tumor-killing potential. Techniques for quantitation of peritoneal cells and functional assays for phagocytic ability are classified as

comprehensive tier II tests by the NTP and as additional type I tests in the draft Redbook.

Macrophage cells and other polymorphonuclear cells (PMNs) contribute to the first-line defense of nonspecific immunity through their ability to phagocytize foreign materials, including pathogens, tumor cells, and fibers (e.g., silica, asbestos). Xenobiotics can affect macrophage function by direct toxicity to macrophages or by modulating their ability to become activated. Differential counts of resident peritoneal cells can be made as a rapid, preliminary assessment of macrophage function for xenobiotics that are not administered parenterally.

Numerous *in vitro* assays can be employed to assess common functions of macrophages and PMNs including adherence to glass, migration inhibition, phagocytosis, respiratory activity (chemiluminescent assays or nitroblue tetrazolium), and target cell killing. In one such assay, the chemotactic response to soluble attractants is evaluated using a Boyden chamber with two compartments that are separated by a filter. Macrophage cells or PMNs from treated animals are placed in one side and a chemotactic agent in the other. Chemotaxis is then quantified by counting the number of cells that pass through the filter. In another assay, the ability of the macrophages to phagocytize foreign materials can be evaluated by adding fluorescent latex beads to cultures containing macrophage cells and then determining the proportion of cells that have phagocytized the beads using a fluorescent microscope or by flow cytometry (Duke et al., 1985). Similar functions can be evaluated by incubating the cells with known amounts of bacteria. The cells are then removed by filtration or centrifugation, the remaining fluid is plated onto bacterial nutrient agar, and, after a few days of incubation, the bacterial colonies are counted. Furthermore, the efficiency of the cells to kill the bacteria once phagocytized can be assayed by lysing the cells and plating the lysate onto bacterial agar.

Various *in vivo* assessments of macrophage function have also been used. For example, peritoneal exudate cell (PEC) recruitment can be assessed using eliciting agents such as *Corynebacterium parvum*, MVE-2, or thioglycolate (Dean et al., 1984). In one such assay (White et al., 1985), mice are injected intraperitoneally with thioglycolate, sacrificed 4 days later, and the peritoneal cavity is flushed with culture medium. The cell suspension is then counted, the cell concentration is adjusted to a known density ($2 \times 10^5 \text{ mL}^{-1}$), and the cells are cultured for 1 h in 24-well culture dishes. Adherent cells are then washed with medium and aliquots of ^{51}Cr -labeled SRBCs that were opsonized with mouse IgG are added to each well and incubated for various times. This same system can be used to assess adherence and chemotaxis of the PECs (Laskin et al., 1981). Phagocytosis can also be evaluated *in vivo* by measuring the clearance of injected particles from the circulation and the accumulation of the particles in lymphatic tissues such as the spleen.

11.8.3 Mast Cell/Basophil Function

The function of mast cells and basophils to degranulate can be evaluated using a PCA test (Cromwell et al., 1986). Serum containing specific anaphylactic (IgE) antibodies from donor animals previously exposed to a known antigen is first administered by intradermal (or subcutaneous) injection into unexposed host animals. After a sufficient latency period to allow binding of the donor IgE to the host tissue mast cells, the animals are administered a second intravenous injection of the antigen. The anaphylactic antibodies present in the serum will stimulate normally functioning mast cells to degranulate (release histamines) and produce a marked inflammatory response. Using similar *in vitro* assays with mast cells and basophils, the quantities of histamines that are released from the cells can be measured directly in the culture medium.

11.8.3.1 Host-Resistance Assays Host-resistance assays can be used to assess the overall immunocompetence of the humoral or cell-mediated immune systems of the test animal (host) to fend off infection with pathogenic microbes or to resist tumorigenesis and metastasis. These assays are performed entirely *in vivo* and are dependent on all of the various components of the immune system to be functioning properly. Thus, these assays may be considered to be more biologically relevant than *in vitro* tests that only assess the function of cells from one source and of one type. Since these assays require that the animal be inoculated with a pathogen or exogenous tumor cell, they cannot be performed as part of a general preclinical toxicity assessment and are thus classified as type II tests in the revised Redbook. These assays are also included as tier II tests by the NTP.

Several host-resistance assays have been developed using various infectious agents, including bacteria (*Listeria monocytogenes*, *Streptococcus*, and *Escherichia coli*), viruses (influenza, cytomegalovirus, and herpes), yeast (*Candida albicans*), and parasites (*Trichinella spiralis* and *Plasmodium berghei*). These assays have been described in the NTP guidelines (Luster et al., 1988). In general, animals previously treated with a xenobiotic are injected with the pathogen at a target dose that is estimated to kill 10–30% of control animals (LD_{10-30}). After a period of time, the animals receive a challenge dose at a much higher concentration (LD_{60-80}) and by a different route to determine if animals are resistant to reinfection. Although these assays are similar in their mechanisms of resistance to different pathogens, they have been shown to differ with regard to varying degrees of susceptibility by the same drug (Morahan et al., 1979). Thus, for screening purposes, it is recommended that at least two tests be used (Descotes and Mazue, 1987). Although these tests are relatively easy to perform, those involving the use of pathogens require special handling, containment, and decontamination procedures to prevent infection to man and spread throughout the animal colony.

Similar host-resistance assays are used to evaluate the immunosurveillance of spontaneous tumors, which is assessed as the capacity of the organism to reject grafted syngeneic tumors. Various animal-bearing tumor models (Pastan et al., 1986) and host-resistance models have been used to assess immunotoxicity. Several of the host-resistance assays utilize cultured tumor cell lines such as PYB6 sarcoma and B16F10 melanoma cells that are used with C57/BL/6 mice or the MADB106 lung tumor cell lines that are used with Fischer 344 rats. For example, the PYB6 sarcoma model uses death as an end point. In this assay, syngeneic mice are injected with the PYB6 sarcoma cells and death due to tumor is recorded daily. In another routinely used assay, animals that have been treated with a xenobiotic are injected with either B16F10 melanoma cells or Lewis lung carcinoma cells, and then approximately 20 days later, they are sacrificed and pulmonary tumors are measured and counted.

11.9 T-CELL-DEPENDENT ANTIBODY RESPONSE (TDAR)

The TDAR should be performed using a recognized T-cell-dependent antigen like SRBCs, bovine serum albumin (BSA), or KLH that results in a robust antibody response. For the SRBC assay, IgM measurement is considered the most appropriate end point, whereas IgG measurement is considered to be most appropriate for BSA or KLH. Antibody can be measured by using an ELISA or other immunoassay method. One advantage of ELISA over the traditional plaque-forming cell (PFC) assay (Ladics, 2005) is that samples can be collected serially during the study, if necessary. Since immunization is likely to have effects on hematology, clinical chemistry, and histology of lymphatic organs, TDAR studies should always be performed as separate studies or at least in satellite groups of repeat-dose toxicity studies. Brief details of the assay are as follows.

11.9.1 Treatment

1. Use a suitable SPF mouse strain like BALB/c or C57BL/6 × C3H F1 (B6C3F1).
2. Allocate 120 animals in six groups of 10 males and 10 females each. The group sizes may be reduced to five males and five female animals per group when a substantial immunosuppressive effect can be expected. The larger group size should be chosen when immunosuppressives are excluded, since otherwise the statistical power of the assay might be insufficient to prove a lack of immunosuppression.
3. Allow acclimatization for 7 days before sampling of pretest serum (day -7) from the test groups, and allow an additional 7 days of rest before first dosing. Pretest

serum of recovery groups may be taken on day 35, which is 1 week postdosing.

4. Administer the test substance and vehicle daily over a period of 28 days (days 1–28) to all animals using an appropriate route of administration.
5. Use a low, intermediate, and high dose level, whereby the high dose level should be above the NOAEL and below a dose level that causes stress, if possible. The intermediate dose (or low dose) level should ideally represent the intended clinical dose level.
6. Immunize all mice of the test groups on day 14 and mice of the recovery groups on day 42 by intraperitoneal injection of 100 µg KLH per mouse without the use of adjuvant.
7. Sample immune serum from all animals of the test groups on day 29 and from all of the recovery groups on day 57.
8. Store serum at -20°C until ELISA testing.

11.9.2 Hypersensitivity

11.9.2.1 Type I Hypersensitivity Although there are acceptable systems for evaluating type I (immediate) reactions following systemic exposure, there are no reliable animal models for predicting type I reactions following dermal applications or oral administrations of drug. Repeated exposure of a xenobiotic is required to produce a type I response. A drug in the form of a hapten must covalently bind to macromolecules (proteins, nucleic acids) before it can initiate a primary antibody response. Once sensitized, even the smallest exposure to the xenobiotic can elicit a rapid, intensive IgE antibody-mediated inflammatory response. With the exception of antivirals and chemotherapeutic drugs, most drugs should not be reactive with biological nucleophiles since these drugs are usually screened out as mutagens or carcinogens in preclinical safety studies. However, type I hypersensitivity is a particular problem with biotechnology products themselves (e.g., insulin, growth hormones, interleukins), trace impurities from the producing organisms (e.g., *E. coli* proteins, mycelium), or the vehicles used to form emulsions (Matori et al., 1985).

The production of neutralizing antibodies to recombinant DNA protein products or their contaminants may be assayed using ELISAs or RIAs. A suitable animal model used to evaluate the potential for a type I response to protein hydrolysates is detailed in the US Pharmacopoeia. This test is very sensitive for testing proteins administered by the parenteral route but is of little value for low molecular weight drugs and those that are administered orally (Descotes and Mazue, 1987). Active systemic anaphylaxis assay (ASA) can be assessed in guinea pigs following systemic exposure to the test compound. For dermal exposures, however, rabbits or

guinea pigs must be exposed to the test article by intradermal injections and then evaluated for their ability to mount a systemic anaphylactic response. The PCA test (as described previously for mast cells) can also be used to assess a potential anaphylactic response to a test compound. The serum containing potential anaphylactic (IgE) antibodies from donor animals previously exposed to the test compound is first administered by intradermal (or subcutaneous) injection into unexposed host animals. After a latency period, the animals are administered an intravenous injection of the test compound together with a dye. If anaphylactic antibodies are present in the serum, the subsequent exposure to the test compound will cause a release of vasoactive amines (degranulation of mast cells), ultimately resulting in the migration of the dye to the sites of the intradermal serum injections.

11.9.2.2 Types II and III Hypersensitivity No simple animal models are currently available to assess type II (antibody-mediated cytotoxicity) hypersensitivity reactions. IgE antibodies and immune complexes in the sera of exposed animals can be assayed using ELISA or RIA (Radio immune assay) techniques that require the use of specific antibodies to the drug.

Type III (immune complex-related disease) reactions have been demonstrated by the presence of proteinuria and immune complex deposits in the kidneys of the Brown-Norway, Lewis, and PVG/C rat strains. However, susceptibility to the deposition and the subsequent lesions (glomerulonephritis) are often variable and dependent on the strain (Bigazzi, 1988). For example, despite the appearance of clinical signs and proteinuria, after 2-month administration of mercuric chloride, detectable levels of circulating antinuclear autoantibodies can no longer be observed in the Brown-Norway strain (Bellon et al., 1982). By contrast, in PVG/C rats administered mercuric chloride, immune complex deposition and antinuclear autoantibodies are present for longer periods of time; however, proteinuria is not observed (Weening et al., 1978).

11.9.2.3 Type IV Hypersensitivity There are several well-established preclinical models for assessing type IV (delayed-type) hypersensitivity reactions following dermal exposure but not for predicting this response after systemic exposure.

Type IV hypersensitivity responses are elicited by T lymphocytes and are controlled by accessory cells and suppressor T cells. Macrophages are also involved in that they secrete several monokines, which results in proliferation and differentiation of T cells. Thus, there are numerous points along this intricate pathway in which drugs may modulate the final response. To achieve a type IV response, an initial high-dose exposure or repeated lower-dose exposures are applied to the skin; the antigen is carried from the skin by Langerhans cells and presented to cells in the thymus to initiate T-cell proliferation and sensitization. Once sensitized, a

second "challenge" dose will elicit an inflammatory response. Thus, before sensitivity can be assessed, each of the models used to evaluate dermal hypersensitivity requires as a minimum:

- An initial induction exposure
- A latency period for expression
- A challenge exposure

A preliminary test for acute irritancy is also required to ensure that the initial dose is sufficient to stimulate sensitization and that the challenge dose is sufficient to ensure expression of the response without producing irritation, which would confound the response. To confirm suspected sensitization or determine a threshold dose, each assay may also include a second challenge dose 1–2 weeks after the first challenge, at the same or lower concentrations. To increase penetration of the test article, various methods of abrasion (e.g., tape stripping) and occlusive coverings may also be used.

Several systems are used routinely to test compounds for dermal hypersensitivity (Magnusson and Kligman, 1969). The two most commonly used, the modified Buehler test and the guinea pig maximization test (GPMT), are briefly reviewed. More detailed methodology and a description of alternative test systems can be found in Gad and Chengelis (1998). Although either rabbits or guinea pigs are sensitive test species, guinea pigs have traditionally been the animal of choice. Guinea pig models of skin sensitization have been (and remain) widely used and have been valuable in assessing human risk (Andersen and Maibach, 1985).

11.9.2.4 Modified Buehler Buehler (1964) developed the first test system to use an occlusive patch to maximize dermal exposure and to increase the test sensitivity (Buehler, 1964). Although this assay is still insensitive for some xenobiotics that may not sufficiently traverse the epidermis, it is particularly useful for compounds that are either highly irritating by intradermal injection or cannot be dissolved or suspended in a form that is conducive to injection. Other advantages are that the test produces few false positives, rarely overpredicts the potency of sensitizers, and is less likely to produce limiting system toxicity or ulceration at the induction sites. Figure 11.4 shows the test design in its current (OECD) form. The assay is no longer accepted in Europe due to a belief that it has an unacceptable rate of false-negative outcomes and is appropriate only for true topical (dermal) exposures.

During the induction phase, the test compound is applied to a cotton patch (1 in. × 1 in.) or placed in a Hilltop-style occlusive chamber. The patch is then placed onto a shaven area of epidermis on the left flank of a guinea pig and secured firmly in place for 24 h, after which time the patch is removed and the area is observed and scored for irritation

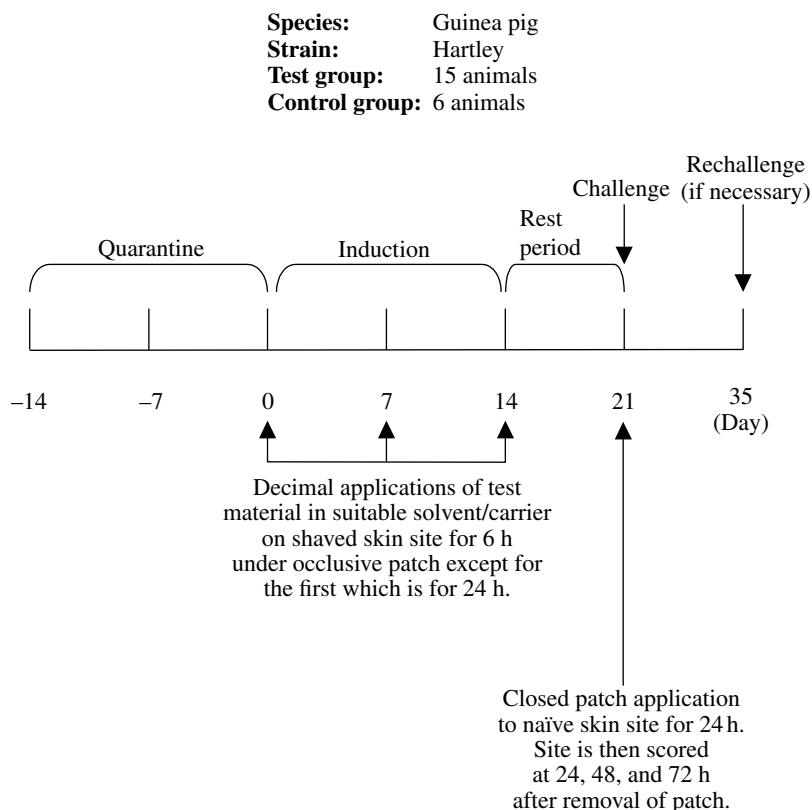


FIGURE 11.4 Line chart for modified Buehler test for delayed contact dermal sensitization in guinea pig.

(i.e., edema, erythema). A fresh patch is then reapplied for 6 h every other day during the induction period for a total of 10 treatments while continuing to score the application site at 24 and 48 h from the start of each treatment. Two weeks after the last induction exposure, the animals receive a challenge exposure for 24 h in the form of a patch applied to a shaven area of epidermis on the other flank (opposite the one used for induction). The challenge dose should be the highest concentration that does not produce dermal irritation after a single, 24 h exposure. The challenge site is observed for evidence of inflammation 24, 48, and 72 h after the patch is removed. Both the intensity and duration of a response to the test article compared to that of the vehicle are used to determine the potential and severity of sensitization.

11.9.2.5 Guinea Pig Maximization Test This assay, as developed by Magnusson and Kligman (1969), differs from the Buehler test in that the compound is administered by intradermal injection during the first stage of induction and coadministered with an adjuvant (Freund's complete adjuvant (FCA)) during the induction phase to further stimulate the immune system. This test system is more sensitive (fewer false negatives) than the Buehler test; however, it may overpredict the potency for many sensitizers. Figures 11.5 and 11.6 illustrate the study design.

Prior to induction, a 4×6 cm area of fur is clipped from the shoulder region of each guinea pig. On day 0, three pairs of intradermal injections are made along opposite sides of the dorsal midline of the animal. The first pair (closest to the head) are administered as test substance in vehicle, the second pair are administered proximal to the first pair and consist only of FCA, and the third pair (spaced most posteriorly) are administered as the test substance in FCA. Seven days later (day 7), a mild to moderately irritating dose of the test article is spread onto a 1×2 in. filter paper patch, secured, and occluded for 48 h on the epidermal site that received the initial injections. On day 21, an area of fur on each flank is shaved and a 1×1 in. patch containing a nonirritating concentration of the test article is applied to one flank and a patch containing vehicle alone is applied to the other flank. The patches are secured and occluded for 24 h, and the challenge sites are scored for inflammation 24 and 48 h after removal of the patches. The incidence of animals that respond and the intensity and duration of a response to the test article are used to determine the potential and severity of sensitization.

11.9.3 Local Lymph Node Assay (LLNA)

This method has developed out of the work of Ian Kimber and associates (Kimber et al., 1986, 1991; Kimber et al.,

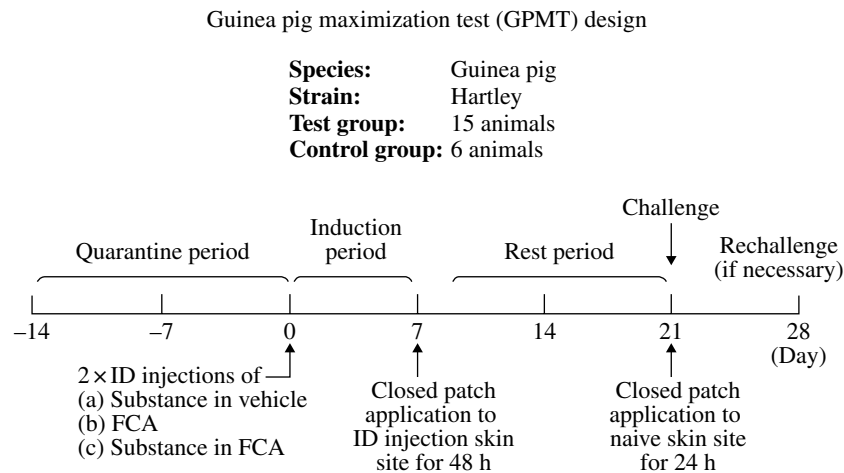


FIGURE 11.5 Line chart for guinea pig maximization test for dermal sensitization.

| Outline of guinea pig maximization test | | | | |
|---|--|--|---|---------------------------------|
| Stage | Induction | | Challenge | Rechallenge |
| Day | 0 | 7 | 21 | 28 |
| Test group (15) | A. 0.1 mL substance ID B. 0.1 mL FCA ID C. 0.1 mL substance + FCA ID | Closed patch—48 h application of substance | Closed patch—24 h substance vehicle | Closed patch—24 h vehicle |
| | A. 0.1 mL vehicle ID B. 0.1 mL FCA ID C. 0.1 mL vehicle + FCA ID | Closed patch—48 h application of vehicle | Closed patch—24 h substance vehicle | Closed patch—24 h substance |

FIGURE 11.6 Illustrative figures for injection and patching of animals in GPMT.

1989; Basketter and Scholes, 1991). It has the advantage over the other methods discussed in this chapter in that it provides an objective and quantifiable end point. The method is based on the fact that dermal sensitization requires the elicitation of an immune response. This immune response requires proliferation of a lymphocyte subpopulation. The local lymph node assay (LLNA) relies on the detection of increased DNA synthesis via tritiated thymidine incorporation. Sensitization is measured as a function of lymph node cell proliferative responses induced in a draining lymph node following repeated topical exposure of

the test animal to the test article. Unlike the other tests discussed in this chapter, this assay looks only at the induction phase, as there is neither a challenge (elicitation) phase or sufficient period of evaluation for development of the underlying clonal expansion in response. Additionally, acute clinical formulations cannot be evaluated in this test system—meaning the all-important question of whether there is actually the potential for the drug to react with the immune system in clinical use remains unaddressed. Rather, in the case of a positive outcome, such interaction is presumed to occur.

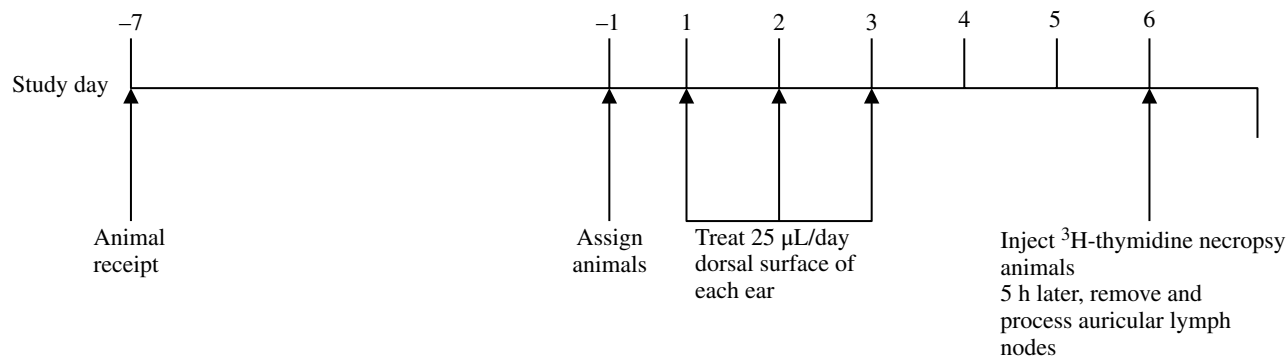


FIGURE 11.7 Mouse LLNA.

The typical test (illustrated in Figure 11.7) is performed using mice—normally female CBA mice 6–10 weeks of age. Female BALB/c and ICR mice have also been used. After animal receipt, they are typically acclimated to standard laboratory husbandry conditions for 7–10 days. The usual protocol will consist of at least two groups (vehicle control and test article treated) of five mice each. They are treated on the dorsal surface of both ears with 25 µL (on each ear) of test article solution for three consecutive days. Twenty-four to forty-eight hours after the last test article exposure, the animals are given a bolus (0.25 mL) dose of [³H]thymidine (20 µCi with a specific activity of 5.0–7.0 Ci mmol⁻¹) in phosphate-buffered saline via a tail vein. Five hours after the injection, the animals are euthanized by CO₂ asphyxiation and the auricular lymph nodes removed.

After removal, the lymph nodes can either be pooled by group or processed individually. Single cell suspensions are prepared by gentle mechanical disaggregation through a nylon (100 µm) mesh. Cells are washed twice by centrifugation in an excess of PBS. After the final supernatant wash is removed, the cells are precipitated with cold 5% trichloroacetic acid (TCA) and kept at 4°C for 12–18 h. The precipitate is then pelleted by centrifugation and resuspended in 1 mL 5% TCA, and the amount of radioactivity is determined by liquid scintillation counting, using established techniques for tritium.

The data are reduced to the stimulation index (SI):

$$SI = \frac{H(\text{dpm}) \text{ treated group}}{H(\text{dpm}) \text{ control group}}$$

An SI of 3 or greater is considered a positive response, that is, the data support the hypothesis that the test material is a sensitizer.

The test article concentration is normally the highest non-irritating concentration. Several concentrations could be tested at the same time should one wish to establish a dose-response curve for induction. The test is easiest to perform if the vehicle is a standard nonirritating organic, such as

acetone, ethanol, or dimethylformamide, or a solvent-olive oil blend. Until a laboratory develops its own historical control base, it is also preferable to include a positive control group. Either 0.25% dinitrochlorobenzene or 0.05% oxazolone is recommended for positive controls. If the vehicle for the positive control is different than the vehicle for the test material, then two vehicle control groups may be necessary.

This method has been extensively validated in two international laboratory exercises (Basketter et al., 1991; Loveless et al., 1996). In the earlier work (Basketter et al., 1991), there was good correlation between the results obtained with guinea pig tests and those obtained with the LLNA. In the 1996 report, for example, five laboratories correctly identified dinitrochlorobenzene and oxazolone as sensitizers and the fact that *p*-aminobenzoic acid was not (Loveless et al., 1996). Arts et al. (1996) demonstrated that rats could be used as well as mice. Interestingly, they validated their assay (for both rats and mice) using BrdU uptake and immunohistochemical staining (rather than [³H]thymidine) to quantitated lymph node cell proliferation.

This method is relatively quick and inexpensive because it uses relatively few mice (which are much less expensive than guinea pigs) and takes considerably less time than traditional guinea pig assays. It has an advantage over other methods in that it does not depend on an arguably subjective scoring system and produces a quantifiable end point. It does require a radiochemistry laboratory and license. Unless one already has an appropriately equipped laboratory used for other purposes (most likely metabolism studies), setting one up for the sole purpose of running the LLNA does not make economic sense. The standard version of the test has been adopted by OECD (OECD429), ICVAM, EMEA, and FDA (see Figure 11.8) but also has been shown to have a modest false-positive rate (misidentifying strong irritants as sensitizers).

Indeed, it has become clear that certain classes of structures (e.g., surfactants, fatty acids, fatty alcohols, siloxanes, and polyols) yield high incidences of false-positive outcomes. Though initially thought to be associated with (and due to) strong irritant responses, this is now clearly not the case (Kreiling et al., 2008; Mehling et al., 2008).

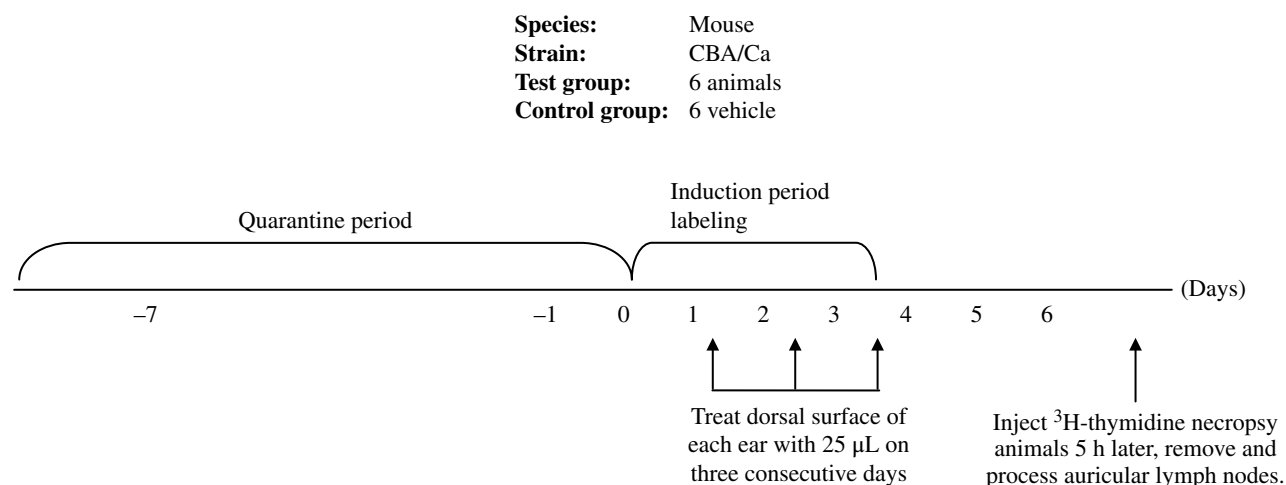


FIGURE 11.8 Mouse local lymph node assay (LLNA) (ICVAM protocol). Modification using flow cytometry instead of radiolabeling is preferable.

The other possibility, of course, is clinical evaluation using the human repeat insult path test (HRIPT). This 5-week hazard test is rarely taken due to cost, though actual clinical formulations are effectively evaluated clinically in the normal course of clinical development of a drug.

11.9.4 Photosensitization

Some compounds can act as photoantigens that require exposure to UV light to become photoactive haptens. The physiochemical characteristics of compounds can sometimes reveal them as potentially photoactive, particularly if they are photounstable to light in the UV range. There are several *in vivo* tests that are used for determining photosensitization. The two assays described here are similar to those previously described for DTH with the primary exception that the dermal test sites are exposed to a light source during the induction and challenge phases. Like the DTH assays, these assays may also include a second challenge dose or the use of various methods of abrasion and occlusion to increase dermal penetration of the test article. The methods outlined in the following are more thoroughly described in Gad and Chengelis (1998).

11.9.4.1 Harber and Shalita Method This method (Harber and Shalita, 1975) is similar to the Buehler test in that the compound is applied topically to guinea pigs without the use of adjuvants; however, the test site is not occluded during exposure. During the induction phase, the compound is applied on alternate days during a 12-day period for a total of six applications. Thirty minutes after each application, the test sites are exposed to a sunlamp for 30 min and then to a black light for 30 min. The challenge dose is applied 21 days after the last induction exposure. Thirty minutes after application, the challenge sites are shielded with a 3 mm thick

piece of glass, while the site is exposed to the black light for an additional 30 min. The glass filters out erythrogenic (causing redness) radiation of less than 320 nm that may confound scoring the reaction. The challenge sites are observed and scored for inflammatory reactions 24 h later.

11.9.4.2 Armstrong Method This test (Ichikawa et al., 1981) resembles the GPMT in that the FCA is injected intradermally at the test sites; however, covered Hilltop chambers are used to apply and occlude the test article at the test site as was described for the Buehler test. During induction, four intradermal injections of FCA are administered at the test site; the test article is applied to the Hilltop chamber, which is then applied over the test site and occluded. After 4 h, the patches are removed and the test site is exposed to UV-A light (320–400 nm) for 30 min. Five additional applications of the test article (without FCA) with subsequent exposure to light are made on alternate days throughout the 11-day induction period. Nine to thirteen days after the last induction exposure, the animals are challenged for 2 h with a nonirritating concentration of the test article on an occluded Hilltop chamber. The patches are then removed, and the sites are exposed to the UV-A light. Each site is graded for inflammation 24 and 48 h after challenge.

Despite its extreme clinical importance, the evaluation of small molecule pharmaceutical for allergenic potential is extremely unreliable in nonclinical toxicology studies. First, a drug that has been proved to be immunogenic in a laboratory animal species may not necessarily be immunogenic in humans, and vice versa. Second, the subtle factors that determine whether an individual responds to an antigen with an IgG or an IgE response can hardly be extrapolated from one species to another. Thus, an immune response that manifests as an allergic response in a laboratory animal may not necessarily manifest as an allergic response in humans, and vice versa.

With regard to antibody-mediated hypersensitivity, three methods have been used extensively to assess the induction of drug-specific (type I) anaphylactic reactions:

- The PCA assay
- The ACA assay
- The ASA assay

All three assays are normally conducted in guinea pigs, which is the only rodent species that actually develops symptoms of severe anaphylactic reactions and even fatal allergic shock. However, the usefulness of these assays for the safety assessment of drugs is considered limited. Since IgE as well as IgG antibodies can cause anaphylactic reactions in guinea pigs, a positive result in any of the three assays can only be weighted as proof of immunogenicity but not allergenicity of a drug. The PCA, ACA, and ASA assays are therefore not requested or recommended for the routine evaluation of allergenicity of investigational new drugs by any regulatory agency.

The situation is as unsatisfactory for the prediction of a type II and type III allergenic potential of drugs. Although there are examples of drugs that are associated with type II and type III hypersensitivity reactions, there are no standard nonclinical methods for predicting these effects. Manifestations of both kinds of immunopathies are often indistinguishable from direct, nonimmune-mediated drug toxicity. Thus, in some instances of hemolytic anemia, vasculitis, or glomerulonephritis, which may be observed during standard toxicity studies, follow-up studies should be considered to determine if antibody-mediated immune mechanisms are involved.

Since all available nonclinical assays to assess the antibody-based allergenic potential of drugs have a limited predictivity for the human situation, detection of drug immunogenicity should already be considered to be a potential safety alert. Whether proven immunogenicity in a nonclinical test does in fact lead to allergic implications in patients can only be convincingly demonstrated (or excluded) in clinical trials or even later still during postmarketing surveillance of the approved drug.

The most robust and predictive procedures available for assessment of allergenicity are those measuring the skin sensitizing potential of topically administered drug substances. In these cases, a drug has to permeate the keratinized skin, bind to MHC molecules of dermal APCs, and stimulate CD4⁺ T cells for proliferation and T_H1 differentiation. Any drug that is able to induce the aforementioned sequence of events will inevitably induce an inflammatory reaction, a so-called DTH reaction, after subsequent challenge exposure to the skin. Thus, in this special situation of dermal sensitization, a proof of immunogenicity is also a proof of allergenicity.

When a drug is intended for topical administration (dermal, ocular, vaginal, rectal), the skin sensitizing potential

of the drug should be determined using an appropriate assay based on sensitization and challenge as part of nonclinical safety evaluation. The most common methods for evaluating the dermal sensitizing potential of drugs have been the Buehler assay (BA) and the GPMT. Both *in vivo* guinea pig-based methods are reliable and have demonstrated a high correlation with known human skin sensitizers. Techniques using mice, like the mouse ear swelling test, which uses an induction and challenge pattern similar to the traditional guinea pig tests, or the murine LLNA, correlate well with traditional guinea pig tests. Especially the LLNA, which is designed to detect lymphoproliferation in draining lymph nodes of the exposition area instead of inflammation following challenge, gives quantitative results. Furthermore, the assay is now accepted by most regulatory agencies with regard to reduction, refinement, and replacement of animal experimentation.

Pseudoallergic (anaphylactoid) reactions, which are independent of antigen-specific immune responses, result from direct drug-mediated histamine release or complement activation. Anaphylactoid reactions can be differentiated from true IgE-mediated anaphylaxis by *in vitro* testing of drug-induced histamine release from mast cell lines or by the detection of activated complement products in serum of animals showing signs of anaphylaxis.

11.10 APPROACHES TO COMPOUND EVALUATION

As outlined previously, there are numerous assays available to assess the various end points that are relevant to immunotoxicity. Early in the development process, a new compound should be evaluated with regard to various factors that may flag it as a potential immunotoxin, including chemical structural or physiochemical properties (e.g., photoallergen) and therapeutic class (i.e., immunomodulators, anti-inflammatory drugs, and antimetabolites) (Herzyk and Bussiere, 2008). Compounds from therapeutic or structural classes that are known to be potential immunotoxins or immunomodulators should be evaluated for the effects in question on a case-by-case basis. With the exception of immunomodulators, protein products, and products of biotechnology, the majority of pharmaceuticals can be assessed for most forms of immunotoxicity during routine preclinical toxicity tests. In general, a well-conducted preclinical toxicity study can detect most serious immunotoxins in the form of altered clinical, hematologic, or histological end points (Haley et al., 2005). For example, possible effects on humoral immunity may be indicated from clinical observation of gastrointestinal or respiratory pathology, changes in serum total protein and globulin, and by histological changes in lymphoid cellularity. Likewise, effects on the cell-mediated response may be observed as increases in

infections and tumor incidences and by changes in the T-cell compartments of lymphoid tissues. In the case of immunosuppressive drugs such as cyclophosphamide and cyclosporin A, the immune effects seen in rodents are similar to those observed in the clinic (Dean et al., 1987).

If perturbations are observed in any hematologic or histopathologic indicators of immunotoxicity, it is then prudent to follow up these findings with one or more of the following:

- Use of special immunochemical and cytological assays that can be performed retrospectively on samples taken from the animals in question
- Use of more specific *in vitro* assays to further assess effects on the pertinent target system and potential mechanism of activity
- Use of more specific *in vitro* and *ex vivo* assays to determine toxicological significance
- Inclusion of additional nonroutine parameters for immunotoxicity assessment in subsequent (longer-term) toxicity assays (can also include additional satellite groups for functional tests that may require coadministration of adjuvants, pathogens, or tumor cells)

11.10.1 Use of *In Vivo* Tests

In vivo tests are more relevant indicators than are *in vitro* tests of immunotoxicity since the dynamic interactions between the various immunocomponents, as well as the pertinent pharmacokinetic (absorption, distribution, plasma concentrations) and metabolic factors, are taken into consideration. However, it is important to select the appropriate animal model and to design the protocol such that it will accurately reflect drug (or relevant metabolite) exposure to humans. For example, one should consider species variability when selecting the animal model, since biological diversity may further obscure the ability to accurately predict human toxicity.

11.10.1.1 Species Selection When possible, the species selected should demonstrate similar pharmacology and toxicity profiles to those anticipated in the clinic. Thus, the test animals should metabolize the drug and express the same target organ responses and toxic effects as humans. Although the rat and dog are the most common species used in preclinical safety tests, they are not as well characterized and validated as the mouse for assessing effects on immune function. For most immunosuppressive drugs, rodent data on target organ toxicities and comparability of immunosuppressive doses have been reflective of what was later observed in the clinic. Immunosuppressive effects and the doses that produced them have been shown to be similar in the various species that are typically used in preclinical safety tests (Dean and Thurmond, 1987). An exception has been seen

with glucocorticosteroids, which are lympholytic in rodents but not in primates (Claman, 1972; Haynes and Murad, 1985). Although some compounds may show different pharmacokinetics and pharmacological effects in rodents than in humans, rodents still appear to be the most appropriate animals for assessing immunotoxicity of non-species-specific compounds (Dean and Thurmond, 1987).

The appropriate animal model is also important when performing follow-up testing or additional mechanistic tests to further investigate findings observed as part of the routine preclinical safety tests. When possible, these studies should employ the same animals or animal model in which the change was initially observed for several reasons, as outlined by Bloom et al. (1987), including:

- The incidence of adverse effect may be low and not easy to reproduce.
- Another species may not be genetically susceptible to the toxic effect.
- The biological significance of the change is well defined in that model.
- If the change follows long-term exposure to the drug, reproducing the effect in another model may be costly and impractical.

11.10.1.2 Route and Treatment Regimen When possible, it is important to administer the compound by the route and treatment regimen most appropriate for demonstrating the specific response and/or reflecting the intended clinical route of administration. It is also necessary for the compound to be in the same dosage form (i.e., salt form, excipients, solubilizers) that will be used clinically. With the exception of tests for contact hypersensitivity, most of the *in vivo* tests can be carried out with a minimum of three dose levels, which are needed to assess a dose-response relationship. Dose levels should range from the proposed clinical dose, or one that approximates the no-effect level, to a maximum-tolerated (or limit) dose that is lower than the LD₅₀ but that produces some evidence of general toxicity (e.g., reduced body weight). A wide dose interval may be necessary to detect immune changes that show a nonlinear dose response. Proper dose selection is crucial for a meaningful interpretation of test results since severe stress and malnutrition may produce indirect immunotoxic effects that would confound a clear interpretation of the data.

For compounds such as antibiotics, with a relatively short duration of therapeutic exposure, a short treatment period of 1–2 weeks in the animal model is generally appropriate. Longer treatments may not be suitable for these drugs since animals can adapt to toxic doses or develop a tolerance by inducing enzymes that increase metabolism of the drug. However, for compounds with intended chronic or prolonged

usage, animals should be treated at least daily for at least a month, to assess the cumulative effects of the drug.

11.10.2 Use of *In Vitro* Tests

In vitro tests are useful as sensitive follow-up tests to determine potential effects or mechanisms of effects on specific cell types at the cellular and molecular levels (Urbisch et al., 2015). In addition, most are relatively simple to perform and *ex vivo* tests can be performed in conjunction with preclinical *in vivo* tests. There are several advantages to using *in vitro* tests, which are listed as follows:

- Specific cell types of humoral components of the immune system can be isolated and studied.
- Cells can be stimulated with various mitogens to assess their proliferative functions *in vitro*.
- For mechanistic studies, cells and their secretory products can be systematically studied in isolation and in various combinations to assess their interactions and cell-to-cell communications.

However, for general preclinical assessments and screening purposes, *in vitro* tests should be well validated and used cautiously for several reasons:

- They may over- or underestimate an effect or give contradictory results compared to *in vivo* tests.
- Most immunotoxic responses express a clear dose-response relationship that can be used for human risk assessment. However, it is more difficult to extrapolate *in vitro* concentrations than *in vivo* animal doses (plasma concentrations) to the clinical dose.
- It is difficult to simulate *in vitro* the interaction of all of the various cell types and modulators of immune function that make up the *in vivo* system.
- Cells can be harvested from a variety of sources and each source may have a different sensitivity since they may be at various stages of maturation or activation.

11.10.3 Assessment of Immunotoxicity and Immunogenicity/Allergenicity of Biotechnology-Derived Drugs

This is an area outside of current ICH S8 guidance and yet of clinical importance to the safety of current new therapeutics. The decision on a suitable species for preclinical immunotoxicity assessment of biotechnology-derived drugs must be made on a case-by-case basis. In any case, the limits of predictivity should be clearly stated in the rationales for choosing a certain assay protocol. The biological activity together with species and/or tissue specificity of many biotechnology-derived pharmaceuticals (e.g., recombinant

cytokines, therapeutic antibodies) often precludes standard toxicity testing designs in commonly used species (e.g., rats and dogs). The same holds for immunotoxicity testing (Shankar et al., 2007). The design of immunotoxicity testing programs for biotechnology-derived drugs should include the use of relevant species. A relevant species is one in which the test material is pharmacologically active due to expression of the receptor or an epitope (in the case of monoclonal antibodies). A variety of techniques (e.g., immunochemical or functional *in vitro* tests) can be used to identify a relevant species. In some cases, nonhuman primates may be the only suitable species available. When no relevant species exists, the use of transgenic mice expressing the human receptor or epitope may be accepted by regulatory agencies. The information gained from use of a transgenic mouse model expressing the human receptor is optimized when the interaction of the product and the humanized receptor has similar physiological consequences to those expected in humans. The entrance of "Follow-on Biologicals" into the marketplace has raised levels of concern that these may well lead to undesirable immunogenicity as a complication (Wadhwa, 2007).

In other cases, the use of the homologous animal protein instead of the human counterpart may be considered. While useful information may be also gained from the use of homologous proteins, it should be noted that the production process, range of impurities/contaminants, pharmacokinetics, and exact pharmacological mechanism(s) may differ between the homologous form and the product intended for clinical use. Thus, from a formalistic point of view, the test item used in such protocols is not identical to the drug substance to be assessed. Results from such studies can therefore only be weighed as "supportive data." In such situations, it is highly recommended to discuss the testing strategy with the responsible regulatory agency for scientific advice. Where it is not possible to use transgenic animal models of homologous proteins, it may be advisable to assess certain aspects of potential immunotoxicity *in vitro* using human material like PBMCs, monocyte-derived macrophages, or long-term cultivated cell lines of hematopoietic origin.

Most biotechnology-derived pharmaceuticals intended for human use are *per se* immunogenic in animals. The induction of antibody formation in animals is therefore not predictive of a potential for antibody formation in humans. In this regard, the results of, for example, a guinea pig anaphylaxis test, which is usually positive for xenogeneic protein products, are not predictive for reactions in humans. Such studies are therefore considered of little value for the routine evaluation of these types of products. It must be kept in mind that even humanized proteins may be immunogenic in humans. In most cases, reliable information on immunogenicity of biotechnology-derived drugs can therefore only be obtained during clinical studies. However, immunogenicity studies in animals using biotechnology-derived drugs

may yield valuable information when comparing the immunogenic potential of a test compound with a biosimilar reference compound or between different production batches.

Even if immunogenicity assessment of biotechnology-derived pharmaceuticals has limited predictivity for the human situation, measurement of antibodies associated with the administration of biotechnology-derived drugs should always be included in the design of a repeat-dose toxicity study (Wierda et al., 2001). Antibody responses should be characterized with regard to titer, number of responding animals, and neutralizing or nonneutralizing antibodies. Furthermore, the detection and quantization of antibodies should be correlated with any pharmacological and/or toxicological changes. Specifically, the effects of antibody formation on pharmacokinetic/pharmacodynamic parameters, incidence and/or severity of adverse effects, complement activation, or the emergence of new toxic effects should be considered when interpreting the data. Attention should also be paid to the evaluation of possible pathological changes related to immune complex formation and deposition.

11.10.4 Suggested Approaches to Evaluation of Results

Several rodent toxicity studies have shown impaired host resistance to infectious agents or tumor cells at exposure levels of drugs that did not cause overt signs of toxicity (Vos, 1977; Dean et al., 1982). One serious limitation to the incorporation of specific immunotoxicological evaluations into general use in safety assessment for pharmaceuticals is a lack of clarity in how to evaluate and use such findings. This problem is true for all new diagnostic techniques in medicine and for all the new and more sensitive tools designed to evaluate specific target organ toxicities. Ultimately, as we have more experience and a reliable database that allows us to correlate laboratory findings with clinical experience, the required course of action will become clearer. However, some general suggestions and guidance can be offered:

- First, it is generally agreed that adverse effects observed above a certain level of severity should be given the same importance as any other life-threatening events when assessing biological significance. These are effects that are so severe that they are detected as part of the routine evaluations made in safety assessment studies. Such findings may include death, severe weight loss, early appearance of tumors, and the like. Findings such as significantly increased mortalities in a host-resistance assay would also fit into this category.
- Second, there are specific end-point assays for which an adverse outcome clearly dictates the action to be taken. These end points include either immediate or delayed hypersensitivity reactions, because once the individual is sensitized, a dose–response relationship may not apply.
- Third, as with most toxicological effects, toxic effects to the immune system are dependent upon dose to the target site. The dose–response curve can be used to determine no-effect and low-effect levels for immunotoxicity. These levels can then be compared to the therapeutic levels to assess whether there is an adequate margin of safety for humans.

If we consider both the specific immunotoxicity assays surveyed earlier in the chapter and the arrays of end points evaluated in traditional toxicology studies, which may be indicative of an immune system effect, these guidelines leave many potential questions unanswered. As additional data on individual end points indicative of immune system responses are collected, the pharmaceutical toxicologist is challenged with various issues regarding assay interpretation and relevance to proposed (or future) clinical trials. For example, what do significant, but non-life-threatening, decreases in antibody response, lymphocyte numbers, macrophage functions, or host resistance in an animal mean about the clinical use of a drug in a patient? The intended patient population is clearly relevant here—if the disease is one in which the immune system is already challenged or incorrectly modulated, any immune system effect other than an intended one should be avoided. There are several additional considerations and questions that should be answered when evaluating the biological and clinical significance of a statistically significant immune response:

1. *Is there a dose response?* The dose response should be evaluated as a dose-related trend in both incidence and severity of the response. If there is a dose-related response, is the lowest dose (preferably plasma level) at which the effect is seen near or below the target clinical dose (plasma level), and is there an adequate therapeutic margin of safety?
2. *Does the finding stand alone?* Is a change observed in only one parameter, or are there correlated findings that suggest a generalized, biologically significant effect? For example, are there changes in lymph node and spleen weights and morphological changes in these tissues to accompany changes in lymphocyte numbers?
3. *Is the effect a measure of function or a single end-point measurement?* Functional measures such as host resistance of phagocytosis involve multiple cells and immunocomponents and, therefore, are considered to be more biologically relevant than a significant change in a single end-point measurement (e.g., T-cell number).
4. *Is the effect reversible?* Reversibility of a response is dependent on the drug itself, exposure levels/duration, and factors related to the test animal (metabolic

capability, genetic susceptibility, etc.). Most effects produced by immunosuppressive drugs have been shown to be reversible after cessation of therapy, such as those produced during cancer chemotherapy. However, if a tumor develops before the immune system is restored, the effect is not reversible, as is the case of secondary tumors related to chemotherapy.

5. *Is there sufficient systematic toxicity data available at levels that demonstrate adequate exposure?* If a study was designed such that there was insufficient exposure or duration of exposure to potential lymphoid target tissues, the test protocol may not be adequate to demonstrate an adverse effect.

In general, a well-conducted long-term study in two species, with no indication of immunotoxicity, based on the considerations outlined previously, should be adequate to evaluate the potential for drug-induced immunotoxicity. If the results from these studies do not produce evidence of immune-specific toxicity after examination of standard and/or additional hematologic, serum chemical, and histopathologic parameters, then additional testing should not be indicated. However, if there are structure–activity considerations that may indicate a potential for concern, or if significant abnormalities are observed that cannot be clearly attributed to other toxicities, then it is important to perform additional tests to fully assess the biological significance of the findings.

11.11 PROBLEMS AND FUTURE DIRECTIONS

There are some very pressing problems for immunotoxicology, particularly in the context of pharmaceuticals and biologic therapeutics and the assessment of their safety. Unlike industrial chemicals, environmental agents, or agricultural chemicals, pharmaceutical products are intended for human exposure, are usually systemically absorbed, and have intentional biological effects on man—some of which are intentionally immunomodulating (interleukins, growth factors) or immunotoxic (cyclosporin, cyclophosphamide).

11.11.1 Data Interpretation

The first major issue was presented and explored in the preceding section. This is how to evaluate and utilize the entire range of data that current immunotoxicological methodologies provide to determine the potential for immunotoxicity and how to interpret the biological significance of minor findings.

11.11.2 Appropriate Animal Models

As previously addressed, most routine preclinical toxicology tests are performed with rats and dogs; therefore, toxicity, pharmacokinetic, and pharmacology data are most abundant

for these species. However, most immunological parameters are best characterized and validated with mice. In addition, the NTP test battery was developed for the mouse, and some of these assays cannot be readily transferred to the rat. Over the last few years, several laboratories have begun adapting tests to both the rat and the dog (Bloom et al., 1987; Thiem et al., 1988); however, efforts need to continue along these lines to further our understanding of the immune responses in these species and how they correlate with other animal models and man.

11.11.3 Indirect Immunotoxic Effects

A problem related to data interpretation is how to distinguish secondary effects that may indirectly result in immunotoxicity from the primary effects of immunotoxicity in preclinical toxicity studies. Various factors may produce pathology similar to that of an immunotoxin, including:

- Stress in a chronically ill animal as related to general toxicity, such as lung or liver damage, can result in immune suppression.
- Malnutrition in animals with drug-induced anorexia or malabsorption can trigger immune suppression.
- Infections and/or parasites may also modulate immune parameters.

These indirect factors must be systematically ruled out, and additional mechanistic studies may be necessary to address this problem. The potential for some indirect effects may be assessed through histopathologic evaluation of endocrine organs such as the adrenals and pituitary.

11.11.4 Hypersensitivity Tests

Probably the largest immunotoxicity concern in clinical studies is unexpected hypersensitivity reactions. While the available guinea pig- and mouse-based tests for delayed contact hypersensitivity resulting from dermal exposure are generally good predictors, there are currently no well-validated models for either immediate or (dermal) delayed hypersensitivity responses resulting from either oral ingestion or parenteral administration. Yet these two situations are the largest single cause for discontinuing clinical trials.

One assay that may hold some promise for delayed hypersensitivity is an adoptive transfer–popliteal lymph node assay (Klinkhammer et al., 1988). This assay, based on the techniques previously described for the popliteal lymph node assay, allows assessment of hypersensitivity following systemic exposure of the drug. Donor mice are first injected with drug for five consecutive days. After a 4-week latency period, potentially sensitized T cells obtained from the spleen are injected into the footpad of a syngeneic mouse together with a subcutaneous challenge dose of the drug.

Two to five days after the cell transfer, the popliteal lymph nodes are measured and observed for evidence of a response (enlargement). Once this assay is validated, it should allow for a more relevant assessment of hypersensitivity for drugs that are administered systemically (Gleichmann et al., 1989).

11.11.5 Autoimmunity

Traditional methods for assessing immunotoxicity as part of routine preclinical toxicity tests are primarily geared toward the detection of immunosuppressive effects. Although it is possible to incorporate clinical methods for detecting immune complexes and autoantibodies into the preclinical test protocols, the significance of adverse findings is ambiguous. Since these effects have a genetic component to their expression, the relevance of findings in animals is of questionable significance, particularly since these findings in the clinic do not always correlate with pathological effects.

11.11.6 Functional Reserve Capacity

As previously discussed, the immune system has a tremendous reserve capacity that offers several levels of protection and backups to the primary response. As a consequence, this functional reserve can allow biologically significant, immunotoxic insults to occur without the appearance of morphological changes. Furthermore, adverse effects may remain subclinical until the organism is subjected to undue stress or subsequent challenge. Thus, there is some concern that routine immunopathologic assessments by themselves may not be sufficiently sensitive to detect all immunotoxins, particularly when testing is conducted in a relatively pathogen-free, stress-free laboratory environment.

11.11.7 Significance of Minor Perturbations

Although the immune system has a well-developed reserve capacity, some of these systems may act synergistically rather than independently. For instance, a macrophage can recognize and kill bacteria coated with antibodies more effectively than can either the macrophage or antibodies alone. Thus, even minor deficiencies and impairments may have some impact on the organism's ability to fend off infection or tumors, particularly if the organism is very young, old, ill, stressed, genetically predisposed to certain cancers, or otherwise immunocompromised. These considerations lead to some additional questions that must be addressed:

- What level of immunosuppression will predispose healthy or immunocompromised individuals to increased risk of infections or tumors?
- Will slight disturbances or immunosuppression lead to a prolonged recovery from viral or bacterial infections?

- Will slight up-modulation for extended periods result in autoimmune diseases or increased susceptibility to allergy?
- Are individuals that are slightly immunosuppressed at higher risk of developing AIDS after exposure to HIV?

11.11.8 Biotechnology Products

Immunotherapeutics such as interferons and interleukins hold tremendous promise for those diseases where malfunctioning of the immune system is not the root of pathogenesis. Likewise, many of the new approaches to therapy of yet untreatable diseases are aimed at modulating the body's own immune system. Many of the new therapeutics coming from biotechnology are proteins of human origin. As such, they can evoke antibody responses in nonhuman species that are not indicative of what will be seen in patients. Meaningful evaluations must allow the toxicologist to discriminate between those responses that are relevant to clinical development/utilization and those that are not. With "follow-on" biologic therapeutics now entering the marketplace, there is a greater risk of potential unwanted immunogenicity (Wadhwa, 2007).

In summary, it is the role of preclinical immunotoxicology testing to allow us to identify potential immune hazards early in development, before they are found in the clinic, and to provide us with a mechanistic understanding for the basis of these effects so that we may direct the development of alternative agents and/or treatment regimens to avoid them. The challenge for the toxicologist is to determine the appropriate course of action for evaluating each unique drug and to differentiate the desired therapeutic effects from the undesired and potentially adverse effects.

REFERENCES

- Aden R, Cohen N. (1993) Psychoneuroimmunology: conditioning and stress. *Annu. Rev. Psychol.* 44:53–85.
- Alper JC, Wieman MC, Rueckl FS, McDonald CJ, Calabresi P. (1985) Rationally designed combination chemotherapy for the treatment of patients with recalcitrant psoriasis. *J. Am. Acad. Dermatol.* 13:567–577.
- Andersen KE, Maibach HI. (Eds.) (1985) Contact allergy predictive test in guinea pigs. In: *Current Problems in Dermatology*, Vol. 14. Karger, Basel.
- Arts JHE, Droge SCM, Bloksma N, Kuper CF. (1996) Local lymph node activation in rats after dermal application of the sensitizers 2,4-dinitrochlorobenzene and trimellitic anhydride. *Food Chem. Toxicol.* 34:55–62.
- Bach FH, Voynow NK. (1966) One-way stimulation in mixed leukocyte cultures. *Science.* 153:545–547.
- Bakke OM, Wardell WM, Lasagna L. (1984) Drug discontinuations in the United Kingdom and United States, 1964–1983: issues of safety. *Clin. Pharmacol. Ther.* 35:559–567.

- Barnett MJ, Waxman JH, Richards MA, Ganesan TS, Bragman KS, Rohatiner AZS, Lister TA. (1985) High-dose cytosine arabinoside in the initial treatment of acute leukemia. *Semin. Oncol.* 12:133–138.
- Basketter DA, Scholes EW. (1991) Comparison of the local lymph node assay with the guinea pig maximization test for the detection of a range of contact allergens. *Food Chem. Toxicol.* 30:65–69.
- Basketter DA, Scholes EW, Kimber I, Botham PA, Hilton J, Miller K, Robbins MC, Harrison PTC, Waite SJ. (1991) Interlaboratory evaluation of the local lymph node assay with 25 chemicals and comparison with guinea pig test data. *Toxicol. Mech. Methods.* 1:30–43.
- Batchelor FR, Dewdney JM, Cazzard D. (1965) Penicillin allergy: the formation of penicilloyl determinant. *Nature (London).* 206:362–364.
- Bellon B, Capron M, Druet E, Verroust P, Wial MC, Sapin C, Girard JF, Foidart JM, Mahieu P, Druet P. (1982) Mercuric chloride induced autoimmune disease in Brown-Norway rats: sequential search for anti-basement membrane antibodies and circulating immune complexes. *Eur. J. Clin. Investig.* 12:127–133.
- Bigazzi PE. (1988) Autoimmunity induced by chemicals. *Clin. Toxicol.* 26:125–126.
- Bloom JC, Thiem PA, Sellers TS, Deldar A, Lewis HB. (1985a) Cephalosporin-induced immune cytopenia in the dog—demonstration of cell-associated antibodies. *Blood.* 66:1232.
- Bloom JC, Blackmer SA, Bugelski PJ, Sowinski JM, Saunders LZ. (1985b) Gold-induced immune thrombocytopenia in the dog. *Vet. Pathol.* 22:492–499.
- Bloom JC, Thiem PA, Morgan DG. (1987) The role of conventional pathology and toxicology in evaluating the immunotoxic potential of xenobiotics. *Toxicol. Pathol.* 15:283–293.
- Brocklehurst WE, Humphrey JH, Perry WLM. (1960) Cutaneous antigen-antibody reactions in the rat. *J. Physiol.* 150:489–500.
- Brunton LL, Lazo JS, Parker KL. (2006) *Goodman & Gilman's The Pharmacologic Basis of Therapeutics*, 11th Ed. McGraw-Hill, New York.
- Buehler EV. (1964) A new method for detecting potential sensitizers using the guinea pig. *Toxicol. Appl. Pharmacol.* 6:341.
- Burleson GR, Dean JH, Munson AE. (1995) *Methods in Immunotoxicology*. Wiley-Liss, New York.
- Burnet FM. (1970) The concept of immunological surveillance. *Prog. Exp. Tumor Res.* 13:1–27.
- Burns-Naas LA, Pallardy MJ. (2013) Immunotoxicology. In: *Nonclinical Safety Assessment, a Guide to International Pharmaceutical Regulations* (Brock WJ, Hastings KL and McGown KM, Eds.). John Wiley & Sons, Inc, Hoboken, NJ.
- Calabresi P, Chabner BA. (1990) Antineoplastic agents. In: *The Pharmacological Basis of Therapeutics* (Goodman AG, Rall TW, Nies AS and Taylor P, Eds.). Pergamon Press, New York, pp. 1209–1263.
- Calabresi P, Parks R. (1985) Antiproliferative agents and drugs used for immunosuppression. In: *Goodman and Gilman's The Pharmacological Basis of Therapeutics* (Gilman AG, Goodman LS, Rall TW and Murad F, Eds.), 7th Ed. Macmillan Publishing, New York, pp. 1247–1306.
- Caspritz G, Hadden J. (1987) The immunopharmacology of immunotoxicology and immunorestitution. *Toxicol. Pathol.* 15:320–322.
- CDER (2002) *Guidance for Industry: Immunotoxicology Evaluation of Investigational New Drugs*. U.S. Department of Health and Human Services, FDA, Rockville, MD.
- CDER (2006) *Guidance for Industry: S8 Immunotoxicity Studies for Human Pharmaceuticals*. U.S. Department of Health and Human Services, FDA, Rockville, MD.
- Center for Devices and Radiological Health (CDRH) (1999). Immunotoxicity Testing Guidance, Draft Document. US Department of Health and Human Services, Washington, DC.
- Chazal I, Verdier F, Virat M, Descotes J. (1994) Prediction of drug-induced immediate hypersensitivity in the guinea pig. *Toxicol. In Vitro.* 8:1045–1047.
- Choquet-Kastylevsky G, Vial T, Descotes J. (2001) Drug allergy diagnosis in humans: possibilities and pitfalls. *Toxicology.* 158:1–10.
- Claman H. (1972) Corticosteroids and lymphoid cells. *N. Engl. J. Med.* 287:388–397.
- Clark WR. (1983) *The Experimental Foundations of Modern Immunology*, 2nd Ed. John Wiley & sons, New York, pp. 1–453.
- Colvin M. (1982) The alkylating agents. In: *Pharmacologic Principles of Cancer Treatment* (Chabner BA, Ed.). W. B. Saunders, Philadelphia, PA, pp. 276–308.
- Coombs RRA, Gell PGH. (1975) Classification of allergic reactions responsible for clinical hypersensitivity and disease. In: *Clinical Aspects of Immunology* (Gell PGH, Coombs RRA and Lachman DJ, Eds.). Blackwell Scientific Publications, Oxford, p. 761.
- Cromwell O, Durham SR, Shaw RJ, Mackay JA, Kay AB. (1986) Provocation tests and measurements of mediators from mast cells and basophils in asthma and allergic rhinitis. In: *Handbook of Experimental Immunology* (Weir DM, Herzenberg LA and Blackwell C, Eds.), 4th Ed. Blackwell Scientific Publications, Oxford, pp. 127.1–127.51.
- Dean JH, Thurmond LM. (1987) Immunotoxicology: an overview. *Toxicol. Pathol.* 12:265–271.
- Dean JH, Luster MI, Boorman GA, Luebke RW, Lauer LD. (1980) The effect of adult exposure to diethylstilbestrol in the mouse: alterations in tumor susceptibility and host resistance parameters. *J. Reticuloendothel. Soc.* 28:571–583.
- Dean JH, Luster MI, Boorman GA. (1982) Immunotoxicology. In: *Immunopharmacology* (Sirois P and Rola-Pleszczynski M, Eds.). Elsevier Biomedical Press, Amsterdam, pp. 349–397.
- Dean JH, Boorman GA, Luster MI, Adkins BJ, Lauer LD, Adams DO. (1984) Effect of agents of environmental concern on macrophage functions. In: *Mononuclear Phagocyte Biology* (Volkman A, Ed.). Dekker, New York, pp. 473–485.
- Dean JH, Thurmond LM, Lauer LD, House RV. (1987) Comparative toxicology and correlative immunotoxicology in rodents. In: *Environmental Chemical Exposure and Immune System Integrity* (Burger EJ, Tardiff RG and Bellanti JA, Eds.). Princeton Scientific Publishing Company, Princeton, NJ, pp. 265–271.
- Deldar A, Lewis HB, Bloom JC, Apostoli A, Weiss L. (1985) Residual stem cell defects associated with cephalosporin therapy in dogs. *Blood.* 66:1202.

- Descotes G, Mazue G. (1987) Immunotoxicology. *Adv. Vet. Sci. Comp. Med.* 31:95–119.
- DeSwarte RD. (1986) Drug allergy: an overview. *Clin. Rev. Allergy.* 4:143–169.
- Dighiero G, Lymberi J, Marie JC, Rouyse S, Butler-Browne GS, Whalen RG, Avrameas S. (1983) Murine hybridomas secreting natural monoclonal antibodies reacting with self antigens. *J. Immunol.* 135:2267–2271.
- Druet P, Bernard A, Hirsch F, Weening JJ, Gengoux P, Mahieu P, Brikenland S. (1982) Immunologically mediated glomerulonephritis induced by heavy metals. *Arch. Toxicol.* 50:187–194.
- Duke SS, Schook LB, Holsapple MP. (1985) Effects of *N*-nitrosodimethylamine on tumor susceptibility. *J. Leukoc. Biol.* 37:383–394.
- Elion GB, Hitchings GH. (1975) Azathioprine. In: *Antineoplastic and Immunosuppressive Agents* (Sartorelli AC and Johns DG, Eds.). Springer-Verlag, Berlin, pp. 403–425.
- Elliot JF, Lin Y, Mizel SB, Bleackley RC, Harnish DG, Peaettian V. (1984) Induction of interleukin 2 messenger RNA inhibited by cyclosporin A. *Science.* 226:1439–1441.
- EMA (2012) *Immunogenicity Assessment of Biotechnology Derived Therapeutic Proteins*. European Medicines Agency, London, UK.
- Evans E. (2008) Clinical pathology as a crucial insight into immunotoxicology testing. In: *Immunotoxicology Strategies for Pharmaceutical Safety Assessment* (Herzyk DJ and Bussiere JL, Eds.). John Wiley & Sons, Inc, Hoboken, NJ.
- Exon JH, Koller LD, Talcott PA, O'Reilly CA, Henningsen GM. (1986) Immunotoxicology testing: an economical multiple assay approach. *Fundam. Appl. Toxicol.* 7:387–397.
- FDA (Food and Drug Administration), Center for Safety and Applied Nutrition. (1993) Draft: toxicological principles for the safety assessment of direct food additives and color additives used in food. *Fed. Regist.* 58:10536.
- Gad SC, Chengelis CP. (1998) *Acute Toxicology Testing*, 2nd Ed. Academic Press, San Diego, CA.
- Garratty G, Petz LD. (1975) Drug-induced immune hemolytic anemia. *Am. J. Med.* 58:398–407.
- Germolec DR. (2004) Sensitivity and predictivity in immunotoxicity testing; immune endpoints and disease resistance. *Toxicol. Lett.* 149:109–114.
- Gleichmann E, Kimber I, Purchase IFH. (1989) Immunotoxicology: suppressive and stimulatory effects of drugs and environmental chemical on the immune system. *Arch. Toxicol.* 63:257–273.
- Godfrey HP, Gell PGH. (1978) Cellular and molecular events in the delayed-onset hypersensitivities. *Rev. Physiol. Biochem. Pharmacol.* 84:2.
- Golub ES, Green DR. (1991) *Immunology: A Synthesis*. Sinauer, Sunderland, MA.
- Guillaume JC, Roujeau JC, Touraine R. (1984) Lupus systémique après prothèses mammaires. *Ann. Dermatol. Venerol.* 111:703–704.
- Hadden JW, Cornaglia-Ferraris P, Coffey RG. (1984) Purine analogs as immunomodulators. In: *Progress in Immunology IV* (Yamamura Y and Tada T, Eds.). Academic Press, London, pp. 1393–1407.
- Haley P, Perry R, Ennulat D, Frame S, Johnson C, Lapointe JM, Nyska A, Snyder P, Walker D, Walter G. (2005) STP position paper: best practice guideline for the routine pathology evaluation of the immune system. *Toxicol. Pathol.* 33(3):404–407.
- Harber LC, Shalita AR. (1975) The guinea pig as an effective model for the demonstration of immunologically-mediated contact photosensitivity. In: *Animal Models in Dermatology* (Maibach H, Ed.). Churchill Livingstone, New York, pp. 90–102.
- Harmon WE, Parkman R, Gavin PT, Grupe WE, Ingelfunger JR, Yunis EJ, Levey RH. (1982) Comparison of cell-mediated lympholysis and mixed lymphocyte culture in the immunologic evaluation for renal transplantation. *J. Immunol.* 129:1573–1577.
- Hassman RA, Lazarus JH, Dieguez C, Weetman AP, Hall R, McGregor AM. (1985) The influence of lithium chloride on experimental autoimmune thyroid disease. *Clin. Exp. Immunol.* 61:49–57.
- Hastings KL. (2001) Pre-clinical methods for detecting the hypersensitivity potential of pharmaceuticals: regulatory consideration. *Toxicology.* 158:85–89.
- Haukaas SA, Hoisater PA, Kalland T. (1982) *In vitro* and *in vivo* effects of diethylstilbestrol and estramustine phosphate (Estracyte) on the mutagen responsiveness of human peripheral blood lymphocytes. *Prostate.* 3:405–414.
- Haynes RC, Murad F. (1985) Adrenocorticotrophic hormone; adrenocortical steroids and their synthetic analogs; inhibitors of adrenocortical steroid biosynthesis. In: *Goodman and Gilman's The Pharmacological Basis of Therapeutics* (Goodman AG, Goodman LS, Rell TW and Murad F, Eds.). Macmillan Publishing Company, New York, pp. 1459–1489.
- Herold KC, Lancki DW, Moldwin RL, Fitch FW. (1986) Immunosuppressive effects of cyclosporin A on cloned T cells. *J. Immunol.* 136:1315–1321.
- Herzyk DJ, Bussiere JL. (2008) *Immunotoxicology Strategies for Pharmaceutical Safety Assessment*. John Wiley & Sons, Inc., Hoboken, NJ.
- Hinton DM. (1992) Testing guidelines for evaluation of the immunotoxic potential of direct food additives. *Crit. Rev. Food Sci. Nutr.* 32:173–190.
- Hollenberg SM, Giguere V, Segui P, Evans RM. (1987) Colocalization of DNA-binding and transcriptional activation functions in the human glucocorticoid receptor. *Cell.* 49:39–46.
- Hunter T, Urowitz MB, Gordon DA, Smythe HA, Ogryzlo MA. (1975) Azathioprine in rheumatoid arthritis. A long-term follow-up study. *Arthritis Rheum.* 8:15–20.
- Hutchings P, Nador D, Cooke A. (1985) Effects of low doses of cyclophosphamide and low doses of irradiation on the regulation of induced erythrocyte autoantibodies in mice. *Immunology.* 54:97–104.
- ICH S8 (2006) *Guidance for Industry-Immunotoxicity Studies for Human Pharmaceuticals*, CBER, Rockville, MD.
- Ichikawa H, Armstrong RB, Harber LC. (1981) Photoallergic contact dermatitis in guinea pigs: improved induction technique using Freund's complete adjuvant. *J. Invest. Dermatol.* 76:498–501.
- Idsoe O, Guthe T, Willcox RR, DeWeck AL. (1968) Nature and extent of penicillin side-reactions, with particular reference to fatalities from anaphylactic shock. *Bull. World Health Organ.* 38:159–188.

- Jerne NK, Nordin AA. (1963) Plaque formation in agar by single antibody-producing cells. *Science*. 140:405.
- Kahan BD and Bach JF. (1988) Proceedings of the second international congress on cyclosporine. *Transplant. Proc.*, 20(Suppl.):11131.
- Kay JE, Benzie CR. (1984) Rapid loss of sensitivity of mitogen-induced lymphocyte activation to inhibition by cyclosporin A. *Cell. Immunol.* 87:217–224.
- Kimber I, Mitchell JA, Griffin AC. (1986) Development of a murine local lymph node assay for the determination of sensitizing potential. *Food Chem. Toxicol.* 24:481–494.
- Kimber I, Hilton J, Weisenberger C. (1989) The murine local lymph node assay for identification of contact allergens: a preliminary evaluation of in situ measurement of lymphocyte proliferation. *Contact Dermatitis*. 21:215–220.
- Kimber I, Hilton J, Botham PA, Basketter DA, Scholes EW, Miller K, Robbins MC, Harrison PTC, Gray TJB, Waite SJ. (1991) The murine local lymph node assay: results of an inter-laboratory trial. *Toxicol. Lett.* 55:203–213.
- Klaus GGB, Hawrylowicz CM. (1984) Cell-cycle control in lymphocyte stimulation. *Immunol. Today*. 5:15–19.
- Kleinman S, Nelson R, Smith L, Goldfinger D. (1984) Positive direct antiglobulin tests and immune hemolytic anemia in patients receiving procainamide. *N. Engl. J. Med.* 311:809–812.
- Klinkhammer C, Popowa P, Gleichmann H. (1988) Specific immunity to the diabetogen streptozotocin: cellular requirements for induction of lymphoproliferation. *Diabetes*. 37:74–80.
- Koller LD. (1987) Immunotoxicology today. *Toxicol. Pathol.* 15:346–351.
- Kreiling R, Hollnagel HM, Hareng L, Eigler D, Lee MS, Griem P, Dreessen B, Kleber M, Albrecht A, Garcia C, Wendel A. (2008) Comparison of the skin sensitizing potential of unsaturated compounds as assessed by the murine local lymph node assay (LLNA) and the guinea pig maximization test (GPMT). *Food Chem. Toxicol.* 46(6):1896–1904.
- Kumagai Y, Shiokawa Y, Medsger TAJ, Rodnan GP. (1984) Clinical spectrum of connective tissue disease after cosmetic surgery. Observations of eighteen patients and a review of the Japanese literature. *Arthritis Rheum.* 27:1–12.
- Ladics G. (2005) Assays for antibody production. In: *Encyclopedia Reference of Immunotoxicity* (Voher H-W, Ed.). Springer, New York, p. SC-53.
- Laskin DL, Laskin JD, Weinstein IB, Carchman RA. (1981) Induction of chemotaxis in mouse peritoneal macrophages by phorbol ester tumor promoters. *Cancer Res.* 41:1923.
- Lawrence DA. (1985) Immunotoxicity of heavy metals. In: *Immunotoxicology and Immunopharmacology* (Dean JH, Luster MI, Munson AE and Amos HA, Eds.). Raven, New York, pp. 341–353.
- Logue GL, Boyd AE, Rosse WF. (1970) Chlorpropamide-induced immune hemolytic anemia. *N. Engl. J. Med.* 283:900–904.
- Loveless EE, Ladics GS, Gerberick GF, Ryan CA, Basketter DA, Scholes EW, House RV, Hilton J, Dearman RJ, Kimber I. (1996) Further evaluation of the local lymph node assay in the final phase of an international collaborative trial. *Toxicology*. 108:141–152.
- Luebke R, House R, Kimber I. (2007) *Immunotoxicology and Immunopharmacology*, 3rd Ed. CRC Press, Boca Raton, FL.
- Luster ML, Boorman GA, Dean JH, Luebke RW, Lawson LD. (1980) The effect of adult exposure to diethylstilbestrol in the mouse. Alterations in immunological function. *J. Reticuloendothel. Soc.* 28:561–569.
- Luster ML, Hayes HT, Korach K, Tucker AN, Dean JH. (1984) Estrogen immunosuppression is regulated through estrogenic responses in the thymus. *J. Immunol.* 133:110–116.
- Luster ML, Pfeifer RW, Tucher AN. (1985) Influence of sex hormones on immunoregulation with specific reference to natural and synthetic estrogens. In: *Endocrine Toxicology* (McLachlin JA, Korach K and Thomas J, Eds.). Raven, New York, pp. 67–83.
- Luster ML, Blank JA, Dean JH. (1987) Molecular and cellular basis of chemically induced immunotoxicity. *Annu. Rev. Pharmacol. Toxicol.* 27:23–49.
- Luster ML, Munson AE, Thomas PT, Holsapple MP, Fenters JD, White KL Jr, Lauer LD, Germolec DR, Rosenthal GJ, Dean JH. (1988) Development of a testing battery to assess chemical-induced immunotoxicity: national toxicology program's guidelines for immunotoxicity evaluation in mice. *Fundam. Appl. Toxicol.* 10:2–19.
- Luster ML, Pait DG, Portier C, Rosenthal GJ, Dermolec DR, Comment CE, Munson AE, White K, Pollock P. (1992a) Qualitative and quantitative experimental models to aid in risk assessment for immunotoxicology. *Toxicol. Lett.*, 64/65:71–78.
- Luster ML, Portier C, Pait DG, White KL Jr, Gennings C, Munson AE, Rosenthal GJ. (1992b) Risk assessment in Immunotoxicology. I. Sensitivity and predictability of immune tests. *Fundam. Appl. Toxicol.* 18:200–210.
- Mach PS, Brouilhet H, Amor B. (1986) d-penicillamine: a modulator of anti-DNA antibody production. *Clin. Exp. Immunol.* 63:414–418.
- Magnusson B, Kligman AM. (1969) The identification of contact allergens by animal assay. *J. Invest. Dermatol.* 52:268–276.
- Maki E. (1997) The practice of preclinical immune toxicology testing in Japan. *Ther. Innov. Regul. Sci.* 31:1325–1329.
- Male D, Brostoff J, Roth DB, Roitt I. (2012) *Immunology*, 8th Ed. Elsevier, Philadelphia, PA.
- Matori YL, Chang AE, Lipford EH, Brazier R, Hyatt CL, McDonald HD, Rosenberg SA. (1985) Toxicity of recombinant human interleukin-2 in rats following intravenous infusion. *J. Biol. Response Mod.* 4:377–390.
- Mazzone SB, Canning BJ. (2002) Guinea pig models of asthma. In: *Current Protocols in Pharmacology*. John Wiley & Sons, New York, pp. 5.26.1–5.26.30.
- Mehling A, Ball N, Cagen S, Carrillo JC, Eigler D, Certa H, Cooman CD, Esch H, Graham C, Haux C, Kreiling R, Skirda K, Veenstra G. (2008) *Discordant results between the local lymph node assay and the guinea pig tests: specificity versus selectivity*. Presented at SOT Annual Meeting, March 17, 2008, Seattle, Washington.
- Merluzzi VJ. (1985) Comparison of murine lymphokine, activated killer cells, natural killer cells, and cytotoxic T lymphocytes. *Cell. Immunol.* 95:95–104.
- Morahan PS, Klykken PC, Smith SH, Harris LS, Munson AE. (1979) Effects of cannabinoids on host resistance to *Listeria monocytogenes* and herpes simplex virus. *Infect. Immun.* 23(3):670–674.
- Ntoso KA, Tomaszewski JE, Jimenez SA, Neilson EG. (1986) Penicillamine-induced rapidly progressive glomerulonephritis in patients with progressive systemic sclerosis: successful

- treatment of two patients and a review of the literature. *Am. J. Kidney Dis.* 8:159–163.
- Paranjpe MS, Boone CW. (1972) Delayed hypersensitivity to simian virus 40 tumor cells in BALB-c mice demonstrated by a radioisotopic footpad assay. *J. Natl. Cancer Inst.* 48:563.
- Pastan L, Willingham MC, FitzGerald DJ. (1986) Immunotoxins. *Cell.* 47:64148.
- Pattels H-G, Taylor J. (2008) Immunotoxicity testing: ICH guidelines S8 and related aspects. In: *Preclinical Development Handbook: Toxicology* (Gad SC, Ed.). John Wiley & Sons, Inc, Hoboken, NJ, pp. 323–362.
- Patterson R, De Ewarte RD, Greenberger PA, Grammer LC. (1986) Drug allergies and protocols for management of drug allergies. *NER Allergy Proc.* 7:325–242.
- Penn I. (1977) Development of cancer as a complication of clinical transplantation. *Transplant. Proc.* 9:1121–1127.
- Perry HM, Tane M, Camody S. (1970) Relationship of acetyl transferase activity to antinuclear antibodies and toxic symptoms in hypertensive patients treated with hydralazine. *J. Lab. Clin. Med.* 76:114–125.
- Pieters R. (2001) The popliteal lymph node assay: a tool for predicting drug allergies. *Toxicology.* 158:65–69.
- Pung OJ, Luster ML, Hayes HT, Rader J. (1984) Influence of steroid and nonsteroid sex hormones on host resistance in the mouse: increased susceptibility to *Listeria monocytogenes* following exposure to estrogenic hormones. *Infect. Immun.* 46:301–307.
- Pung OJ, Tucker AN, Yore SJ, Luster MI. (1985) Influence of estrogen on host resistance: increased susceptibility of mice to *Listeria monocytogenes* correlates with depressed production of interleukin 2. *Infect. Immun.* 50:91–96.
- Riccardi C, Puccetti P, Santoni A, Herberman RB. (1979) Rapid *in vivo* assay of mouse natural killer cell activity. *J. Natl. Cancer Inst.* 63:1041–1045.
- Roitt IM, Brostoff J, Male DK. (1985) *Immunology*. CV Mosby, St. Louis, MO.
- Russel AS. (1981) Drug-induced autoimmune disease. *J. Allergy Clin. Immunol.* 1:57.
- Shand FL. (1979) Review/commentary: the immunopharmacology of cyclophosphamide. *Int. J. Immunopharmacol.* 1:165–171.
- Shankar G, Pendley C, Stein KE. (2007) A risk-based bioanalytical strategy for the assessment of antibody immune responses against biologic drugs. *Nat. Biotechnol.* 25:555–561.
- Snyder PW. (2012) Immunology for the toxicologic pathologist. *Toxicol. Pathol.* 40:143–147.
- Spears CJ, Batchelor JR. (1987) Drug-induced autoimmune disease. *Adv. Nephrol.* 16:219–230.
- Spreafico F. (1988) Immunotoxicology in 1987: problems and challenges. *Fundam. Clin. Pharmacol.* 2:353–367.
- Spreafico F, Anaclerio A. (1977) Immunosuppressive agents. In: *Immunopharmacology 3* (Hadden J, Coffey R and Spreafico R, Eds.). Plenum Medical Book Company, New York, pp. 245–278.
- Stockholm SL, Scott MA. (Eds.) (2008) *Fundamental of Veterinary Clinical Pathology*, 2nd Ed. Blackwell Publishing, Ames, IA.
- Sullivan J. (1989) Immunological alterations and chemical exposure. *Clin. Toxicol.* 27:311–343.
- Tabo JD, Paul WE. (1973) Functional heterogeneity of murine lymphoid cells. III. Differential responsiveness of T cells to phytohemagglutinin and concanavalin A for T cell subsets. *J. Immunol.* 110:362–369.
- Talal N. (1987) Autoimmune mechanisms in patients and animal models. *Toxicol. Pathol.* 15:272–275.
- Thiem PA, Halper LK, Bloom JC. (1988) Techniques for assessing canine mononuclear phagocyte function as part of an immunotoxicologic evaluation. *Int. J. Immunopharmacol.* 10:765–771.
- Ueda S, Wakahim Y, Takei L, Mori T, Lesato K. (1980) Autologous immune complex nephritis in gold injected guinea pigs. *Nippon Jinzo Gakkai Shi.* 22:1221–1230.
- Unkovic J, Combes M, Mazue G, Roncucci R. (1984) *Poster*. Annual Meeting of the American Society of Dermatology, December 1–6, 1984, Washington, DC.
- Urbisch D, Mehling A, Guth K, Ramirez T, Honarvar N, Kolle S, Landsiedel R, Jaworska J, Kern PS, Gerberick F, Natsch A, Emter R, Ashikaga T, Miyazawa M, Sakaguchi H. (2015) Assessing skin sensitization hazard in mice and men using non-animal test methods. *Regul. Toxicol. Pharmacol.* 71(2):337–351.
- Verdier F, Chazai I, Descotes J. (1994) Anaphylaxis models in the guinea pig. *Toxicology.* 93(1):55–61.
- Volkman A. (1984) *Mononuclear Phagocyte Function*. Marcel Dekker, New York.
- Vos JG. (1977) Immune suppression as related to toxicology. *CRC Crit. Rev. Toxicol.* 5:67.
- Wadhwa M. (2007) Unwanted immunogenicity implications for follow-on biologics. *Drug Inf. J.* 41:1–10.
- Wallner BP, Mattaliano RJ, Hession C, Cate RL, Tizard R, Sinclair LK, Foeller C, Chow EP, Browning JL, Ramachandran KL, Pepinsky RB. (1986) Cloning and expression of human lipocortin, a phospholipase A2 inhibitor with potential anti-inflammatory activity. *Nature.* 320:77–80.
- Ways SC, Mortola JF, Zvaifler NJ, Weiss RJ, Yen SSC. (1987) Alterations in immune responsiveness in women exposed to diethylstilbestrol in utero. *Fertil. Steril.* 48:193–197.
- Webb DR, Winkelstein A. (1982) Immunosuppression, immunopotential and anti-inflammatory drugs. In: *Basic and Clinical Immunology* (Stites DP, Stobo JD, Fudenberg HH and Wells JV, Eds.), 4th Ed. Lange Medical, Los Altos, CA, pp. 277–292.
- deWeck AL. (1983) Immunopathological mechanisms and clinical aspects of allergic reactions to drugs. In: *Handbook of Experimental Pharmacology: Allergic Reactions to Drugs* (deWeck AL and Bundgaard H, Eds.). Springer-Verlag, New York, pp. 75–133.
- Weening JJ, Fleuren GJ, Hoedemaeker J. (1978) Demonstration of antinuclear antibodies in mercuric chloride-induced glomerulopathy in the rat. *Lab. Invest.* 39:405–411.
- Weigle WO. (1980) Analysis of autoimmunity through experimental models of thyroiditis and allergic encephalomyelitis. *Adv. Immunol.* 30:159–275.
- White KL Jr, Sanders VM, Barnes DW, Shopp JGM, Munson AE. (1985) Immunotoxicological investigations in the mouse: general approach and methods. *Drug Chem. Toxicol.* 8:299–331.
- Wierda D, Smith HW, Zevieki CM. (2001) Immunogenicity of biopharmaceuticals in laboratory animals. *Toxicology.* 58:71–74.
- Yoshida S, Golub MS, Gershwin ME. (1989) Immunological aspects of toxicology: premises not promises. *Regul. Toxicol. Pharmacol.* 9:56–80.

12

NONRODENT ANIMAL STUDIES

12.1 INTRODUCTION

Most safety assessment studies are conducted in rodents (rats, mice, and hamsters) or their close “cousins” rabbits and guinea pigs. Outside of the pharmaceutical, medical device, and veterinary product industries, it has become rare for the practicing toxicologist to have close familiarity with the nonrodent animal species addressed in this chapter. Indeed, it is unlikely that a toxicologist has received any significant academic experience or training with these species. Additionally, use of these species in the EU even for therapeutics has become very difficult at best. However, the proper use of nonrodent species is essential in the evaluation of potential new therapeutic entities on both scientific and regulatory grounds. Indeed, there are numerous studies showing significantly better concordance as to pharmacokinetics, target organs, and tolerance between man and nonrodents with man than with rodents for detection of adverse responses to pharmaceuticals (Olson et al., 2000). This is tacitly recognized in regulatory practice for those new product development plans where a single species is deemed appropriate and sufficient (medical devices, protein therapeutics, and 505(b)(2) approval candidates) and the single species is overwhelmingly a nonrodent. This critical aspect of nonrodents is reflected by recent books focused on such use in single species (the nonhuman primate in Bluemel et al., 2015 and the minipig in McAnulty et al., 2010).

In addition to rodent studies, regulatory guidelines for pharmaceuticals require that repeated-dose safety studies of up to 9 months (in the United States, where 12 months is also infrequently a regulatory requirement—6 months elsewhere) in duration be conducted in a nonrodent species. These nonrodent species also tend to be closer to man than do rodents

in many aspects including gastrointestinal absorption and transit times (DeSesso and Williams, 2008). The most commonly used nonrodent species is the dog, followed by the nonhuman primate (NHP), and then the pig (though the minipig is gaining in frequency of use). Another nonrodent model used to a limited extent in systemic safety evaluation is the ferret. The major objectives of this chapter are (i) to discuss differences in rodent and nonrodent experimental design; (ii) to examine the feasibility of using the dog, NHP, pig, and ferret in safety assessment testing; and (iii) to identify the advantages and limitations associated with each species.

12.2 COMPARISON BETWEEN RODENT AND NONRODENT EXPERIMENTAL DESIGN

12.2.1 Number of Animals

One of the main differences in experimental design between rodent and nonrodent safety studies is the number of animals used (Table 12.1). In general, approximately nine times as many rodents are used in toxicity studies, as compared to nonrodents. This difference is reflected in the 2012 estimates of overall usage of animals in the United States (891 161 nonrodents total) as published by the US Department of Agriculture, which showed that only about 10% of the animals used in general research for that year were either nonhuman primates (7.2%), pigs (6.3%), or dogs (7.6%). It is estimated that 93% of the animals used in research are not included in this count. The smaller number of nonrodents used is related in part to the higher costs associated with their purchase, housing, and maintenance and in part to their limited use in other areas of research.

TABLE 12.1 Comparison of Rodent and Nonrodent Experimental Design

| Duration of Study | The Total Number of Animals on Study (No./Group/Sex) | | | |
|-------------------|---|---------------------|---------------------|---------------------|
| | Rat | Dog | Pig | Monkey |
| 4 weeks | 360 (20) ^a | 40 (4) | 40 (4) | 32 (4) |
| 13 weeks | 280 (20) ^{a,b} | 48 (6) ^b | 48 (6) ^b | 48 (6) ^b |
| 52 weeks/9 months | 360 (10) ^{a,b} | 64 (8) ^b | 64 (8) ^b | 48 (6) ^b |

^aIncludes satellite animals for pharmacokinetic evaluation.^bThe number of animals/group includes several animals.

12.3 DIFFERENCES IN STUDY ACTIVITIES

12.3.1 Blood Collection

In rodent studies, a large number of satellite animals (often close to the number used in the main study phase) are usually needed for pharmacokinetic blood sampling, whereas with most nonrodent species adequate, blood samples can be collected from the main study animals without compromising their health status or study results (Diehl et al., 2001).

12.3.2 Dosing

Capsule dosing is probably the most appropriate route of oral administration for dogs and pigs and gavage for monkeys and ferrets. Nasogastric gavage is also occasionally used for dogs and primates.

12.3.3 Handling of Animals

Once rodents are acclimated to handling, they are generally relatively easy to work with. In contrast, some nonrodent species, such as nonhuman primates and pigs, are often difficult to handle because of their size, strength, emotionality, and aggressiveness. This can make the conduct of routine study activities (such as dosing, blood collection, and recording electrocardiograms (ECGs)) relatively time consuming, as well as stressful to the animals (Gad, 2015).

12.3.4 Behavioral Evaluation

Behavioral assessment of nonrodents is generally more difficult than evaluation of rodents because of their larger size, difficulties associated with handling and manipulation, and their greater awareness of and reactivity to the experimenter. Such factors can confound detection and/or interpretation of more subtle test compound-related behavioral changes. However, there is now a published and validated functional observational battery (FOB) for neurological exams of dogs, and it is widely used (Gad et al., 2003).

12.4 NONRODENT MODELS

This section is devoted to the definition and comparison of the four nonrodent animal models (dog, NHP, pig, and ferret) in terms of experimental procedures, environmental and dietary requirements, and advantages and disadvantages of use in safety assessment testing. The dogs (beagles in toxicology) and pigs and ferrets are purpose bred for use in research. Indeed, contrary to PETA, it is no longer legal to use dogs sourced otherwise from “class B” sources. NHPs are generally wild caught.

12.5 DOG

12.5.1 Environmental and Dietary Requirements

Typical housing for laboratory dogs consists of stainless steel (of dimensions appropriate to the dog’s size) or indoor pens (typical dimensions are 3 ft, 8 in. wide, 8 ft high, and 10 ft long). The two important aspects of the laboratory dog’s environment are the need for exercise and socialization. Recent amendments to the US Department of Agriculture’s Animal Welfare Act require that an exercise and socialization program be established for dogs maintained in a laboratory environment (APHIS, 1989). Difficulty often arises in establishing a program that will be truly beneficial to the animals. One important consideration is whether dogs should be exercised as a group or individually. Studies have demonstrated that dogs exercised alone tend to spend most of their time walking or investigating the area rather than jumping or running (Campbell et al., 1988), which suggests that group exercise is more beneficial.

The need for a certain degree of socialization is also important in terms of both dog–dog and dog–human contact. If at all possible, dogs should share a cage or pen with another animal. One on-study approach undertaken by some laboratories has been to allow dogs of the same sex and treatment group to have daily contact with each other, usually from early evening to early morning. If study dictates do not make this approach feasible, efforts should be made to ensure the animals are housed in such a way that they have visual, auditory, and olfactory access to each other.

The recommended dry-bulb temperatures and relative humidity ranges for dogs are 64.4–84.2°F (16–27°C) and 30–70%, respectively (National Research Council (NRC), 2011). Increases in temperature and high humidity are of particular concern because of the dog’s limited capacity to dissipate heat (primarily through panting and, to a lesser extent, through radiation and conduction). Dogs would likely not survive exposure for extended periods of time to environments where the temperature is in excess of 40°C and there is 40% relative humidity (Norris et al., 1968).

While the dog is a carnivore, it is able to adapt to an omnivorous diet. Requirements for dietary sources of energy, amino acids, glucose precursors, fatty acids, minerals, vitamins, and water have been established based on recommendations by the National Research Council (NRC, 2006). Adult beagles maintained in a laboratory environment function well with one feeding of standard laboratory chow per day. In safety assessment testing, however, some compounds may induce serious dietary deficiencies through induced loss of appetite, malabsorption, or vomiting, and, in these cases, it may be advisable to provide a dietary supplement.

The dog's requirement for water appears to be self-regulated and depends on factors such as the type of feed consumed, ambient temperature, amount of exercise, and physiological state; therefore, in most cases, dogs should have free access to water.

12.5.2 Common Study Protocols

The dog is the most commonly used nonrodent species in safety assessment testing (i.e., acute, subchronic, and chronic studies and cardiovascular safety pharmacology studies). The exception to this is its use in developmental toxicity and reproductive studies. For developmental toxicity studies, the dog does not appear to be as sensitive an indicator of teratogens as other nonrodent species such as the monkey (Earl et al., 1973) and the ferret (Gulamhusein et al., 1980), and, for reproductive studies, the dog is not the species of choice because fertility testing is difficult to conduct (due to prolonged anestrus and the inability to predict the onset of proestrus) and there is no reliable procedure for induction of estrus or ovulation.

Examples of experimental designs and suggested timing of various study activities for 4- and 13-week dog studies are shown in Tables 12.2 and 12.3, respectively. Beagles are generally in the age range of 6–9 months at study start, and the number of animals per sex per treatment group (N) will depend on the duration of the study. For a 2- to 4-week study without a reversal phase, N will likely be 4, whereas for a 26-week or 1-year study, N will be larger (e.g., N may be 9 including 3/sex/group for the reversal phase); it should be noted that it has become an FDA expectation for high-dose and control groups to include animals to allow the assessment of recovery.

Dogs should be selected for study use on the basis of acceptable body weights, urinalysis, and clinical pathology findings, as well as physical, ophthalmic, and electrocardiographic evaluations. To minimize familial effects, efforts should be made to ensure that no two littermates of the same sex are assigned to the same treatment group.

Because most, if not all, study-related activities are conducted in the same dogs, the stress induced by repeated manipulation of dogs for activities such as blood collection,

TABLE 12.2 Four-Week Dog or Primate Toxicity Study

Experimental Design

Four to five groups (including a control group)—4/sex/group main groups, plus extra in high-dose and control groups to evaluate recovery

Repeated daily dosing for 28 or 29 days

Necropsy starting on day 29

Study Activities

Daily observations: pretreatment and twice daily during the study period

Physical examinations: pretreatment and after dosing during weeks 2 and 4

ECG: pretreatment and after dosing during weeks 2 and 4

Ophthalmic examinations: pretreatment and during week 4

Body weight: pretreatment, weekly, and prior to scheduled necropsy

Feed consumption: pretreatment and weekly

Clinical lab: twice before the first dosing day, before dosing on day 2, during week 2, and prior to scheduled necropsy

Urine collection: pretreatment and during weeks 2 and 4

Pharmacokinetic: blood collected at specified times after dosing on days 1 and 28

Source: Adapted from Gad (2015).

TABLE 12.3 Thirteen-Week Dog Toxicity Study

Experimental Design

Four groups (including a control group)—6/sex/group (plus 3/sex extra in high-dose and control groups for assessing potential recovery)

Repeated daily dosing for 91–93 days

Necropsy of main group (4/sex/group) on week 14

Necropsy of reversal group (2/sex/group) on week 18

Study Activities

Daily observations: pretreatment, twice daily during treatment, and once daily during reversal

Physical examinations: pretreatment; after dosing during weeks 4, 8, and 13 of treatment; and during week 4 of reversal

ECG: pretreatment; after dosing during weeks 4, 8, and 13 of treatment; and during week 4 of reversal

Ophthalmic examinations: pretreatment, during weeks 6 and 13 of treatment, and during week 4 of reversal

Body weight: pretreatment (three times), weekly during the treatment and reversal periods, and prior to scheduled necropsy

Feed consumption: pretreatment, weekly through first month, bimonthly during the remainder of treatment period, and weekly during reversal

Clinical lab: pretreatment, during weeks 4 and 8 of treatment, prior to scheduled necropsy, and during weeks 1 and 4 of reversal

Urine collection: pretreatment, monthly during treatment, and during week 4 of reversal

Pharmacokinetic samples: blood collected at specified times after dosing on day 1 and during weeks 6 and 12

Source: Adapted from Gad (2015).

ECG, and physical examinations needs to be taken into consideration. Efforts should be made wherever possible to separate study activities by several days.

12.5.3 General Study Activities

12.5.3.1 Dosing Techniques The most frequently used route of administration in dog safety assessment studies is oral. Dosing by capsule is usually the preferred oral route in the dog. Gavage is also used, but is a more labor-intensive technique, and there is always the possibility of gavage error or aspiration. Since dogs have a natural tendency to vomit, it is recommended that they be sham dosed with empty capsules or gavaged with a water solution for several days prior to starting a study so that they can become acclimated to the dosing procedure. If the test substance is irritating (and therefore proemetic), the use of enterically coated capsules should be considered. Attention should be paid, however, to the time course of emesis. If it follows dose administration by more than 4–6 h, the cause will not have been due to the gastric irritation, as the material will have already exited the stomach.

Next to oral dosing, the most common dosing route for dogs is intravenous. For bolus or limited infusion intravenous dosing, the femoral, cephalic, and saphenous veins are commonly used. For continuous infusion, the jugular is often the vein of choice, and the procedure will require surgical preparation for either a direct line catheterization or subcutaneous insertion of a vascular access port (a rigid, multipuncturable reservoir equipped with an indwelling catheter).

Other routes of administration used less commonly in dog safety studies are subcutaneous, intramuscular, intraperitoneal, rectal, and vaginal and dermal.

12.5.3.2 Clinical Observations and Physical Examinations Daily clinical observations in dog safety studies, usually conducted pretreatment (prior to cage cleaning) and at a specified time(s) after dosing, consist of a home cage observation with notation of clinical signs indicative of poor health (such as salivation, weight loss, abnormal feces, and vomitus) or abnormal behavior (such as reduced activity or increased aggression).

Physical examinations are conducted less frequently and generally involve the evaluation of gait, mobility, demeanor, and reflexes (pupillary, corneal, patellar, wheelbarrowing, hopping, etc.), as well as an examination of the head (eyes, ears, mouth, teeth, gums, and tongue), body (palpation for signs of masses and nodal swellings), and urogenital and anal regions.

12.5.3.3 Feed Consumption Feed consumption is relatively easy to measure in the dog, since dogs do not usually spill much of their feed. Generally, the full feed bowl is weighed at the beginning and the empty bowl at the end of the feeding period (usually a 4 h period). This is repeated over two or

three consecutive days and the average daily feed consumption calculated from the numbers.

12.5.3.4 Electrocardiograms The collection of recorded ECGs in test dogs has become increasingly of interest due to the concurrence to the QT prolongation by drugs and ICH/FDA/EMA-mandated evaluation studies. Traditionally, a 10-lead system, consisting of the bipolar leads (I, II, and III), the augmented unipolar leads (aVR, aVL, and aVF), and the unipolar precordial leads (V10, CV6LL (V2), CV6LU (V4), and CV5RL (rV2)), has been recommended for dogs in the conscious state (Detweiler et al., 1979; Detweiler, 1981). For toxicity studies, dog ECGs are usually recorded by technical personnel and read at a later time by a veterinary cardiologist. Depending on the length of the study and the pharmacological-toxicological profile of the test compound, ECGs may be recorded as frequently as every day or as infrequently as every 3 months. Dog ECGs can also be highly variable. Factors that can affect the quality of the tracing include the positioning of the electrodes, the positioning of the dog, and the degree of nervousness and excitability of the animal. Conditioning the dogs to the electrode clips and the recording position (usually sphinx or right lateral recumbency) during the pretreatment period will help improve the quality of the recording.

The use of surgically implanted sensors to be able to remotely monitor electrophysiology, heart rate, blood pressure, and blood gases has become very common for cardiovascular safety pharmacology studies. It should always be the case using such dogs that a concurrent evaluation of baseline and vehicle effects be performed prior to that of acute drug grounds. Such sensors allow us to effectively monitor an increasing number of physiological and biochemical biomarkers in a continuous manner.

12.5.3.5 Blood and Urine Collection As mentioned previously, serial blood samples can be fairly easily collected from the dog. The jugular vein is probably the most commonly used vein because of its size and accessibility. Other veins used less frequently are the cephalic, femoral, brachial, and saphenous.

Due to the difficulty in obtaining sufficient volumes of urine in dogs over short collection periods, urine is usually collected overnight (~16–17 h period) in stainless steel metabolism cages. It is recommended that a sample for urinalysis be taken early in the collection process and that all samples be collected in light-resistant containers to help avoid problems such as dissolution of urine casts, increased bacterial activity, and breakdown of bilirubin with exposure of the sample to light.

12.5.4 Advantages and Disadvantages

Some of the advantages and disadvantages of using the dog in safety assessment studies are listed in Table 12.4. With

TABLE 12.4 Use of the Beagle in Safety Assessment Studies*Advantages*

Medium size
 Moderate length of hair coat
 Adaptability to living in group housing
 Ease of handling (e.g., dosing, blood collection, and ECG)

Disadvantages

Variation in size and body weight
 Loud, penetrating bark
 Greater test compound requirements than smaller nonrodent species
 Availability
 Exercise and housing requirements

Source: Adapted from Gad (2015).

respect to its medium size and even temperament, the beagle is certainly a desirable nonrodent model. The relative ease in handling beagles makes them suitable for activities such as serial collection of blood samples and recording of ECGs.

Disadvantages include an often wide variation in size and body weight and a loud, penetrating bark. The large amount of space required to house dogs and the current emphasis on regular exercise may also be disadvantages. Test compound requirements are generally higher for the dog than for either the NHP or ferret when these are alternatives to nonrodent species used in safety testing. This may be a problem in the early period of drug development when compound availability is often limited. Other problems center around the dog's tendency to vomit, which can be a disadvantage when compounds are orally administered, and the fact that, unlike rodents, studies requiring a large number of dogs need careful advance planning to ensure that a sufficient number of animals of the appropriate age can be obtained in a timely manner.

12.6 THE FERRET

The ferret, *Mustela putorius furo*, is a small carnivore that has become an increasingly popular species in various areas of research including anatomy, virology, bacteriology, physiology (gastrointestinal, pulmonary, and cardiovascular), pharmacology, neurology, teratology, and, to some extent, toxicology. The reader is referred to the excellent review by Fox (1988) on the biology and diseases of the ferret and to the chapter in Gad (2015) or the recent paper by Gad (2000) on the ferret as an animal model in toxicology. Since 1990, the literature has reported on work done by Pfizer, Hoffman-La Roche, Gilead, Bristol-Myers Squibb, Merck, Yamanouchi, Procter & Gamble, Abbott, and Glaxo Wellcome using ferrets in pharmaceutical development. The ferret is also the species of choice for respiratory virus (especially influenza) efficacy studies. Most recently, Fox and Marini (2014) published a comprehensive volume on biology and disease of the ferret.

12.6.1 Environmental and Dietary Requirements

For reasons of environmental control, ferrets used in safety assessment studies should be housed indoors. It has been suggested that an optimal temperature range for the ferret is 40–65°F (4–18°C), while relative humidity should be maintained in the range of 40–65% (Fox, 1988). The ferret does not tolerate heat well due to its lack of well-developed sweat glands; the primary method of regulating heat loss appears to be through panting (Moody et al., 1985).

Since ferrets are seasonal breeders, the female being monestrous and an induced ovulator, the breeding cycle can be controlled by varying the length of exposure to artificial light. For safety studies, it is desirable to prevent both estrus in females and increased sexual activity in males; thus it has been recommended that the light period be kept short (Fox, 1988). In this laboratory, a 9 h light/15 h dark cycle has been used successfully for this purpose.

Ferrets should be housed in well-ventilated rooms that provide at least 10–15 air changes per hour. Good ventilation is important, since ferrets are susceptible to respiratory viral infections. Additionally, there is a need to dissipate the musky odor of the animals. While housing standards for ferrets are not specified in the NRC (2011), space requirements of 49 × 46 × 46 cm have been defined by other groups (Wilson and Donnoghue, 1982). Stainless steel cat or rabbit cages equipped with a drop pan to catch feces and urine are a suitable form of primary housing for ferrets. Ferrets are more content when they have access to a small secluded nesting area within their cage in which they can sleep, such as hammocks. The use of paper to line the cage or drop pan is not recommended, since the ferrets are likely to eat it. For socialization purposes, ferrets should be housed as a group or have visual access to neighboring ferrets if housed individually.

Since ferrets eat only their caloric requirements, and since their gastrointestinal transit time is short (3–5 h), it is recommended that they receive diet ad libitum. Only one of the available standardized ferret chows commercially available should be fed. The most important dietary variable is the quality of the protein, and ferrets appear to do best with a high percentage of animal protein in their diet (Morton, 2010). Feed consumption will be higher in the fall and winter and lower in the spring and summer. Hairball laxative is essential during the spring and summer months when the animals experience considerable hair loss. Water should be available at all times.

12.6.2 Study Protocols

Historically, the ferret was used more often in teratology (Hoar, 1984), reproductive (Hoar, 1984), and acute safety studies than in repeated-dose studies (4–52 weeks in duration). This has changed since 2000, however, and the use of the ferret in pivotal repeated-dose safety assessment testing

TABLE 12.5 Four-Week Ferret Toxicity Study*Experimental Design*

Five groups (including a control group)—7/sex/group

Repeated daily dosing for 28 or 29 days

Necropsy starting on day 29

Study Activities

Daily observations: pretreatment and twice daily during the study period

Physical examinations: pretreatment and after dosing during weeks 1, 2, and 4

ECG: pretreatment and after dosing during weeks 2 and 4

Ophthalmic examinations: pretreatment and during week 4

Body weight: pretreatment, twice weekly, and prior to scheduled necropsy

Feed consumption: pretreatment and weekly during the study

Clinical lab: pretreatment, week 2, and prior to scheduled necropsy

Urine collection: pretreatment and during weeks 2 and 4

Pharmacokinetic samples: blood collected at specified times after dosing on days 1 and 28

(Thornton et al., 1979; Hart, 1986; Haggerty et al., 1989) has increased particularly for vaccines.

An example of the experimental design for a 4-week pivotal study in ferrets is shown in Table 12.5. Young adult ferrets are usually in the age range of 9–11 months at study start, and there should be a sufficient number of animals in each group for statistical confidence (generally in the range of six to eight animals per sex per group reflecting the high degree of intergroup variability in the species). For longer-term studies, the number of animals per group would be increased to include reversal group animals (three to four animals per sex per dose group).

Assignment of ferrets to a study should be based on evaluation of pretreatment clinical signs and body weights, as well as physical, electrocardiographic, and ophthalmologic findings.

For longer-term studies, females should be spayed to avoid the development of aplastic anemia, which will occur if the animals come into heat and are not bred (Morton, 2010).

As with dogs, efforts should be made to separate study activities as much as possible to minimize the stress of multiple activities being performed in the same animals.

12.6.3 General Study Activities

12.6.3.1 Dosing Techniques Oral dosing of ferrets is usually done by gavage. One method is to hold the ferret perpendicular to the floor and insert the appropriate size stainless steel gavage needle into the animal's mouth, back into the esophagus, and down toward the stomach. Confirmation of correct positioning of the tube can be determined by visual inspection of the aspirate. As with dogs, ferrets

have a tendency to retch or vomit, and daily gavage with a water solution for several days prior to starting a study (for adaptation purposes) is recommended.

Repeated daily intravenous dosing in the ferret is generally considered to be technically difficult and time consuming; the use of an indwelling catheter is recommended (Moody et al., 1985). There are, however, reports in the literature of subchronic intravenous dosing (three times weekly for 3 months) of the ferret via the caudal vein (McLain et al., 1987).

Dosing techniques, such as intramuscular, intradermal, subcutaneous, and intraperitoneal administration, can be used for the ferret. Care needs to be taken, however, when administering lipophilic compounds by the subcutaneous or intradermal routes to avoid inadvertently injecting compounds into the ferret's thick layer of subcutaneous fat, which can result in poor absorption (Moody et al., 1985).

12.6.3.2 Clinical Observations and Examinations Daily clinical observations will usually begin the week prior to study start and continue twice daily (pre- and postdosing) throughout the study. Ferrets are observed in their home cage for signs of physical debilitation (such as abnormal feces and vomitus), behavioral abnormalities, hair loss, swelling of the vulva (females), and testicular prominence (males). A physical examination should periodically be made and should include measurement of rectal temperature; observation of general demeanor and activity; palpation of the head, thorax, and abdomen; examination of the eyes, ears, and body orifices; and testing of the pupillary and patellar reflexes.

12.6.3.3 Feed Consumption Feed consumption can be measured over two to three consecutive days, and the average daily intake calculated. A problem with measuring feed intake in ferrets is their tendency to dig through their feed bowl, which often results in an unacceptable amount of spillage. Use of a feed follower may help reduce the spillage.

12.6.3.4 Electrocardiograms Most electrocardiographic evaluation in the ferret has been previously done in the anesthetized animal, though sensors may be surgically implanted. This allows ECGs to be recorded using the bipolar (I, II, and III) and augmented (aVR, aVL, and aVF) leads. It is quite possible to obtain fairly good-quality ECGs in the conscious ferret using leads I, II, and III. The standard position used for recording ECGs in the conscious or anesthetized ferret is right lateral incumbency. ECGs have also been measured in the ferret using surface electrodes placed between two points on the chest, with the signals being led off to an amplifier by a long, flexible cable and recorded on magnetic tape for later analysis (Andrews et al., 1979). The advantage of such a system is that the animals are allowed to move freely during the recording.

12.6.3.5 Blood and Urine Collection About 5–10 mL of blood can be collected from adult ferrets using retro-orbital blood collection techniques. Other methods of blood collection include cephalic and jugular veins and caudal tail venipuncture, as well as bleeding via the ventral tail artery. Cardiac puncture is also used, but in the opinion of this and other (Hart, 1986) laboratories, the procedure is traumatic and can cause myocardial scarring. Blood collection from the tail can be difficult, because the ferret tail is short and the tail veins and arteries cannot be seen. For all the aforementioned collection techniques, some form of pharmacological or mechanical restraint is required. To facilitate serial blood collection, methodology has been developed for a tethered restraint system with an implanted indwelling venous jugular catheter, which does not interfere with the normal activities of the ferrets and allows blood sampling to occur from outside the cage (Jackson et al., 1988).

For urine collection, glass or plastic rat metabolism cages work well for short-term or overnight collection. Care needs to be taken to avoid contamination of the urine with feces.

12.6.4 Advantages and Disadvantages

Two advantages to using the ferret are its cost and its size (Table 12.6). The cost of the ferret is approximately one-tenth that of the dog. The ferret's smaller size means that it is easier to maintain and more economically housed and fed than the dog (Hart, 1986). Smaller size also means that test compound requirements for the ferret will be considerably less than those for larger nonrodent species (e.g., on the order of one-tenth of that needed for the dog). Another advantage is that if exercise requirements are ever established for the ferret, it will be an easier species than the dog for which to design an acceptable exercise program.

Disadvantages associated with the ferret include its pervading musky odor and its background disease profile. While the scent glands can be removed, 90% of the animal's odor is derived from sebaceous secretions onto the skin. However, neutering the males and spaying the females, in addition to descenting, will markedly reduce the odor. Rats

should be housed as far away from ferrets as possible because of their inherent fear of ferrets (triggered by olfactory stimulation), which can interrupt breeding cycles or disturb other physiological functions (Fox, 1988). In this laboratory's experience, ferrets, which are generally less docile than dogs, can be difficult to handle and prone to bite, especially when restrained for activities such as ophthalmic and ECG examinations. The lack of easily accessible veins for intravenous dosing and serial blood collection is also a disadvantage.

The major disadvantage in the use of the ferret in safety studies is the profile of diseases associated with the species and the resulting variability in background clinical and anatomical pathology. Pneumonitis and hepatic lymphoid accumulation, associated with chronic parvovirus infection, have been observed in ferrets in this laboratory (Haggerty et al., 1989). Submucosal lymphoid nodules of the intestines are also a common finding (Hart, 1986). Additionally, a relatively high incidence of electrocardiographic anomalies (atrial or ventricular premature depolarization and atrial and ventricular extrasystoles) and ophthalmological anomalies (optic nerve hypoplasia and cataracts) has been found in ferrets in this laboratory. While it may be possible to work with the animal suppliers to reduce the chances of receiving animals with background ECG or ocular abnormalities, at the present time there is no supplier of a disease- and viral-free ferret.

12.7 THE PIG

12.7.1 Background

The use of pigs (*Sus scrofa*) in biomedical research is well established. In toxicology, whereas the use of pigs in the United States is largely limited to dermal studies, in Europe they have become very popular for pharmaceutical studies in place of dogs and primates. They have been extensively used for surgical (Swindle et al., 1988) and physiological (primarily cardiovascular, renal, and digestive) research (Khan, 1984; Clausing et al., 1986) for years. Prior to the last 10 years, the pig's use in toxicity testing was uncommon, except in the testing of veterinary or herd management drugs intended for use in swine or in dermal toxicity and absorption studies (McAnulty et al., 2010). Because of their well-accepted physiological similarities to humans, minipigs are becoming increasingly attractive toxicological models (Table 12.7). In fact, they long have been more frequently used in nutritional toxicology studies (Clausing et al., 1986). Among the more common experimental animals, pigs are the only ones whose use is on the increase (Khan, 1984). Their expense (in both procurement and maintenance) and their relatively large size have mitigated against their use in more general toxicity testing. The development of minipigs

TABLE 12.6 Use of the Ferret in Safety Assessment Studies

Advantages

Small size
Significantly lower cost than most other nonrodents
Lower test material requirements (relative to larger nonrodents)
Adaptability to an exercise program

Disadvantages

Pervading musky odor
Rodents' inherent fear of ferrets
Can be difficult to handle
Background disease profile with resulting "background noise" and increased variability in clinical and anatomical pathology

TABLE 12.7 The Minipig in Toxicity Testing

| |
|--|
| Due to the many advantages, mini- and micropigs are real alternatives to the use of nonrodents (dogs, ferrets, and primates) |
| Minnesota minipig introduced in 1949 |
| Body weights at age 2 years |
| Yucatan minipig: 70–90 kg |
| Yucatan micropig: 40–45 kg |
| Göttingen micropig: 35–40 kg |
| Use in general toxicity testing and reproduction and teratological and behavioral toxicity (aspects of public acceptance as a species for testing) |

TABLE 12.8 Main Advantages of the Minipig

| |
|---|
| Similarity to humans in |
| Cardiovascular anatomy and physiology |
| Ventricular performance |
| Electrophysiology |
| Coronary artery distribution |
| Human skin |
| Thickness and permeability |
| Pigmentation |
| Allergic reaction |
| Reaction to burning and distress |
| Gastrointestinal system and digestion |
| Renal system |
| Immune system (FDA: "...better than rodents") |
| P450 total enzyme activity (especially CYP2E1 and CYP3A4) |

has resulted in a strain of more manageable size. In addition, the increase of expense in the use of dogs and the perceived lay opposition to their use make minipigs (in Europe and Israel) much more attractive as a nonrodent species for general toxicity studies. The dog is a far more common companion animal, and many of the recent developments in animal care and use laws have made specific provisions about the care of dogs. Minipigs have been shown to be more sensitive to a wide variety of drugs and chemicals (e.g., carbaryl and methylmercury) than dogs (Khan, 1984). The Food and Drug Administration (FDA) has kept its own breeding colony of minipigs since the early 1960s. In short, there are scientific, economic, and sociological reasons that make minipigs good toxicological models. The reader is referred to an excellent short review by Phillips and Tumbleson (1986) that puts the issue of minipigs in biomedical research into the context of modeling in general. Table 12.8 presents the advantages of the minipig.

Several breeds of miniature swine have been developed. These include, in the United States, the Yucatan micro- and minipigs, the Hanford, the Sinclair, the Pitman-Moore, and the Hormel. The Yucatan and the Sinclair tend to be the most commonly used, though the Göttingen (widely used in Europe) is seeing increasing use (Ellegaard et al., 1995; Koch et al., 2001). Panepinto and Phillips (1986)

have discussed the characteristics, advantages, and disadvantages of the Yucatan minipig in some detail. In Europe, the Göttingen minipig is extensively used. At sexual maturity (4–6 months) the typical minipig weighs 20–40 kg, as compared to 102 kg for the more common pig and 8–15 kg for the dog. Micropigs weigh about 14–20 kg at sexual maturity. The minipig and the dog have comparable life spans; for example, Peggins et al. (1984) reported that the average life span for miniature swine is 15–17 years. The average beagle dog may have a life span of 8–12 years. Most of this discussion will focus on the purpose-bred minipigs, primarily the Yucatan and the Sinclair.

The greatest area of use of the pig in pharmaceutical safety assessment is for dermal agents. It should be noted that there are differences in skin thicknesses, and these differences are not well characterized.

12.7.1.1 Restraint and Dosing In general, minipigs are docile and easily socialized and trained. Barnett and Hensworth (1986) recommended a socialization regimen of 2 min of gentle interaction. Pigs, like most experimental animals, are simply kept and fed (Hunsaker et al., 1984) but have to be occasionally restrained so samples can be taken and other measurements made. Restrain methods designed for commercial swine should not be used for laboratory swine. Panepinto (1986) have described a sling method that provides restraint with minimal stress. The most frequently mentioned dosing routes in the literature are dietary admix, dermal (topical), gavage, and intravenous injections. Generally, minipigs are restrained in a sling while being dosed by the active route such as gavage. If the experiment requires the implantation of, for example, an indwelling catheter, minipigs can be anesthetized with ketamine (20 mg kg⁻¹ IM) as described by Swindle et al. (1988).

12.7.2 Clinical Laboratory

Clinical chemical and hematologic parameters for minipigs have been studied. Ranges for some of the more commonly examined parameters from Yucatan minipigs are summarized in Tables 12.9 and 12.10 (from Radin et al., 1986). Parsons and Wells (1986) have published similar data on the Yucatan minipig. Brechbuler et al. (1984), Oldigs (1986), Ellegaard et al. (1995), and Koch et al. (2001) have published on the Göttingen minipig. Middleton and coworkers have published extensive lists (organized by age and sex) on the hematologic parameters (Burks et al., 1977) and serum electrolytes (Hutcheson et al., 1979) for the Sinclair minipig. In general, the clinical laboratory picture of the various strains is quite similar. No real differences between sexes have been identified, but age can be very much a factor. For example, serum creatinine can be 33% higher in 3-month-old as compared to 18-month-old Sinclair minipigs (based on data reported by Burks et al., 1977). As with other species,

TABLE 12.9 Minipig Clinical Chemistry Parameters in Different Strains

| Parameter | Yucatan | Göttingen |
|---|-------------|--------------|
| Glucose (mmol L ⁻¹) | 3.75 ± 0.64 | 5.98 ± 1.01 |
| Urea (mmol L ⁻¹) | 7.84 ± 2.64 | 3.19 ± 1.15 |
| Creatinine (μmol L ⁻¹) | 115 ± 16 | 52.2 ± 11.1 |
| Total protein (g L ⁻¹) | 74 ± 9 | 54.0 ± 4.6 |
| Albumin (g L ⁻¹) | 50 ± 6 | 26.2 ± 6.0 |
| Bilirubin total (μmol L ⁻¹) | 3.42 ± 1.37 | — |
| Triglycerides (mg L ⁻¹) | 267 ± 134 | 565 ± 250 |
| Total cholesterol (mmol L ⁻¹) | 1.85 ± 0.38 | 1.65 ± 0.38 |
| γ-Glutamyl transpeptidase (UL ⁻¹) | 61.6 ± 11.2 | — |
| Alanine aminotransferase (UL ⁻¹) | 72.5 ± 13.6 | — |
| Aspartate aminotransferase (UL ⁻¹) | 40.3 ± 5.9 | — |
| Na ⁺ (mmol L ⁻¹) | 140.5 ± 4.2 | 142.3 ± 3.00 |
| K ⁺ (mmol L ⁻¹) | 4.1 ± 0.3 | 3.94 ± 0.32 |
| Cl ⁻ (mmol L ⁻¹) | 103.1 ± 4.3 | 101.3 ± 3.6 |
| Ca ²⁺ (mmol L ⁻¹) | 2.62 ± 0.18 | 2.58 ± 0.16 |
| PO ₄ ²⁻ (mmol L ⁻¹) | 2.41 ± 0.26 | 1.61 ± 0.30 |

Source: Data from Parsons and Wells (1986), Brechbuler et al. (1984), and Oldigs (1986).

Note: Data are mean ± SD.

TABLE 12.10 Minipig Hematologic Parameters in Different Strains

| Parameter | Yucatan | Göttingen |
|---|--------------|-------------|
| Red blood cell (10 ⁶ mm ⁻³) | 7.61 ± 0.15 | 7.0 ± 0.80 |
| Hemoglobin (g dL ⁻¹) | 14.87 ± 0.18 | 14.9 ± 1.20 |
| Hematocrit (%) | 44 ± 0.5 | 44.6 ± 4.1 |
| Mean corpuscular volume (fL) | 58.5 ± 0.8 | 64.4 ± 3.7 |
| Mean corpuscular hemoglobin (pg) | 19.8 ± 0.3 | 21.4 ± 1.3 |
| Mean corpuscular hemoglobin concentration (g dL ⁻¹) | 33.9 ± 0.3 | 33.2 ± 0.8 |
| White blood cell (10 ³ mm ⁻³) | 12.73 ± 0.41 | 12.6 ± 3.0 |
| Lymphocytes (10 ³ mm ⁻³) | 7.25 ± 0.24 | 5.75 ± 1.52 |
| Neutrophils (mm ⁻³) | 4.47 ± 0.24 | 5.27 ± 1.29 |
| Eosinophils (mm ⁻³) | 534 ± 57 | 517 ± 31 |
| Monocyte (mm ⁻³) | 422 ± 35 | 945 ± 71 |
| Basophils (mm ⁻³) | 89 ± 15 | 63 ± 1.3 |
| Platelets (10 ³ mm ⁻³) | — | 441 ± 119 |

Source: Data from Burks et al. (1977) (12 months old, sexes pooled); Radin et al. (1986).

health status, feed composition, feeding regimen, fasting state, season, time of day, and so on can affect clinical laboratory results in the minipig. Toxicological experiments should not be run without concurrent controls.

12.7.3 Xenobiotic Metabolism

Some critical parameters of hepatic microsomal drug metabolism in the minipig, common swine, and rats are given in Table 12.11. As most investigators tend to use younger minipigs, the values reported in this table are for

young (<4 years old) minipigs. Relatively few papers have examined the MMFO in a broad age range (10 months to 12 years) of Hanford minipigs. They identified definite age-related differences. The amounts of cytochrome P450, the mitochondrial mixed functional oxidase (MMFO) activity with aniline and *p*-chloro-*N*-methylaniline, and glucuronosyltransferase activity were all significantly higher in middle-aged (5–8 years) versus young (<4 years) minipigs. Freudenthal et al. (1976) examined Hanford minipigs in the 2- to 8-month age range and obtained somewhat different cytochrome P450 (~0.95 nmol mg⁻¹) values than did Peggins et al. (1984) (~0.50 nmol). The reported ranges for aniline hydroxylase (about 0.70 nmol min⁻¹ mg⁻¹) and UDP-glucuronosyltransferase (about 50 nmol min⁻¹ mg⁻¹) were similar in the two papers. Hence, the available data on the MMFO of young Hanford minipigs are fairly consistent. P450 (CYP) isoenzymes 1A, 2A, 2B, 2C, 2D, 2E, and 3A have all been well characterized in the pig (Gad, 2015).

The flavin adenine dinucleotide (FAD) containing monooxygenase (FMFO) has traditionally been studied in hog liver obtained from slaughterhouses (Tynes and Hodgson, 1984). Interestingly, when FMFO activity is compared between species, substrate specificities are found to be generally very similar (Tynes and Hodgson, 1984). Rettie et al. (1990) isolated and studied the FMFO from Yucatan minipig liver. As with the enzyme studied from other species, the hepatic enzyme exists as a single isozymic species, is active with both dimethylaniline (*N*-oxide formation) and alkyl *p*-tolyl sulfides (sulfoxidation), and is enantioselective in metabolite formation. It would thus appear that the minipig does not differ appreciably from regular swine in the presence or activity of FMFO.

Perhaps some aspects of minipig xenobiotic metabolism can be inferred from studies in regular swine. For example, Rendic et al. (1984) demonstrated that cimetidine and ranitidine are excellent inhibitors of the porcine MMFO in vitro and are probably also inhibitory in microsomal preparations from minipigs (Van Ryzin and Trapoid, 1980). Walker et al. (1978) reported on epoxide hydratase activity in various species, including the pig. Depending on the substrate, the pig had activities equivalent to or greater than those of the rat. This was confirmed by Smith et al. (1984) and Watkins and Klaassen (1986). The MMFO, epoxide hydrolase, UDP-glucuronosyltransferase, *N*-acetyltransferase, glutathione *S*-transferase, and sulfotransferase activities in regular swine may be used to help infer the expected activity in minipigs until more complete and specific information appears in the literature on minipigs.

There are relatively few papers that compare in vivo pharmacokinetic behavior of a specific chemical in the minipig versus another animal. Schneider et al. (1977) reported on the toxicology and pharmacokinetics of cyclotrimethylene-trinitramine in the rat and minipig. Rats convulsed within the first several hours after receiving this chemical, whereas

TABLE 12.11 Comparison of Xenobiotic Metabolism Systems in Rat and Pig

| Enzyme | Rat ^a | Minipig ^b | Common Swine ^a |
|---------------------------------|------------------|----------------------|---------------------------|
| Cytochrome P450 ^c | 0.59 ± 0.04 | 0.95 ± 0.02 | 0.30 ± 0.04 |
| MMFO activity ^d | | | |
| Ethylmorphine | 5.09 ± 0.34 | 8.53 ± 0.51 | 1.39 ± 0.16 |
| Ethoxyresorufin | 0.134 ± 0.022 | | 0.88 ± 0.02 |
| Epoxide hydrolase styrene oxide | 8.36 ± 2.48 | — | 11.4 ± 1.67 |
| UDP-glucuronosyltransferase | | | |
| 1-Naphthol | 6.43 ± 1.66 | — | 5.50 ± 0.89 |
| 4-Nitrophenol | 4.51 ± 0.50 | 5.5 ± 1.5 | 9.38 ± 1.07 |
| Glutathione S-transferase | | | |
| DNCB | 2659 ± 168 | — | 2746 ± 499 |
| DCNB | 118 ± 8.8 | — | 2.44 ± 0.23 |
| PAPS sulfotransferase | | | |
| 2-Naphthol | 0.785 ± 0.066 | — | 0.095 ± 0.025 |
| Acetyltransferase | | | |
| <i>p</i> -Aminobenzoate | 0.77 ± 0.23 | — | 0.621 ± 0.111 |

^a Source: Data from Mueller et al. (1980), Smith et al. (1984) and Watkins and Klaassen (1986).

^b Source: Data from Freudenthal et al. (1976) or Peggins et al. (1984).

^c nmol mg⁻¹ microsomal protein.

^d All enzyme activities; nmol min⁻¹ mg⁻¹ (either microsomal or cytosolic) protein.

minipigs convulsed 12–14 h later. This is consistent with the observation that at 24 h postdosing (100 mg kg⁻¹ po), the plasma levels were 3.0 µg mL⁻¹ in rats and 4.7 µg mL⁻¹ in minipigs. Other differences in pharmacokinetics and metabolism between the two species were described. The latent period for convulsion development was more similar between minipigs and humans than between rats and humans. The implication in this chapter is that the minipig is a more suitable model for the study of the toxicity and metabolism of the nitramines than rats.

12.7.4 Dermal Toxicity

Although rabbits are commonly used for the assessment of primary (and even cumulative) dermal irritation (and in some cases accepted by regulators as a nonrodent species), pigs are now generally being considered to be better models for the more sophisticated study of dermal permeability and toxicity. As reviewed by Sambuco (1985), human and porcine skin are similar with regard to sparsity of the pelage, thickness and general morphology, epidermal cell turnover time, size, orientation, and distribution of vessels in the skin. The particularly thin hair coat and lack of pigments in the Yucatan minipigs make it particularly ideal for dermal studies. The size of the animal also provides the additional practical advantage of abundant surface area for multiple site testing.

Sambuco (1985) has described the sunburn response of the Yucatan minipig to ultraviolet (UV) light, suggesting that this species would also make a good model in phototoxicity as well as photocontact dermatitis studies. Thirty 12 cm sites were demarcated, permitting the study of 15 different dermal dosages of UV radiation.

Mannisto and coworkers (1984) have published a series of articles on the dermal toxicity of the anthralins in the minipig. In one experiment, 24 sites per minipig were used to assess the acute dermal irritation of various concentrations to four different chemicals per site. The range of concentrations tested permitted them to calculate the median erythema concentration and median irritation concentrations with relatively few animals. They were able to show clear differences between anthralin congeners (antipsoriatic drugs) with regard to irritation. When compared to other species (mouse and guinea pig), the response of the minipig was the most similar to humans in that in both species these chemicals are delayed irritants, and several days postexposure may pass before the maximal irritant response is presented.

Likewise, Hanhijarvi et al. (1985) studied the chronic, cumulative dermal effects of anthralin chemicals in minipigs. Using only 12 animals, they were able, by having 32 sites per animal, to study the effects of two different chemicals (dithranol and butantrone; both anthralins) in three different formulations at three different concentrations each. The protocol also included observations for systemic toxicity, clinical laboratory measurements, plasma drug analyses, and gross and histopathological examinations.

In a third report (very similar to the second), Hanhijarvi et al. (1985) clearly demonstrated that the type of vehicle can greatly influence irritation in that dithranol was clearly more irritating when applied in paraffin than when applied in a gel. They were also able to demonstrate that although dithranol was less irritating than butantrone acutely, the cumulative irritations (mean scores at the end of 6 months of six times per week applications) were quite similar (Mannisto et al., 1984). There was neither evidence of

systemic toxicity nor evidence of test article in plasma with either species.

12.7.5 Cardiovascular Toxicity

In general, the published literature consistently maintains that the cardiovascular systems of swine and humans are very similar (McAnulty et al., 2010). For example, as reviewed by Lee (1986), swine, including minipigs, have a noticeable background incidence of atherosclerotic lesions, and swine fed with high lipid diets will develop even more extensive atherosclerotic lesions. High lipid diets will produce lesions similar to advanced atheromatous lesions seen in humans. Although few drugs or chemicals have been shown to cause atherosclerosis, this information has three general applications to toxicology and pharmacology. First, the feeding regimen of minipigs should be carefully controlled in general toxicity studies to minimize the incidence of arterial disease, especially in long-term studies. Second, the pathologist should be aware of the natural background of this disease when preparing a diagnosis. Third, the minipig could provide a convenient model for the study of atherosclerotic disease and the screening of potential therapies.

The minipig has been used to study cardiotoxicity, particularly with medical device and drug/device combination products. Van Vleet et al. (1984) reported that minipigs were the only other species studied other than dogs to develop cardiac damage in response to large doses of minoxidil. In both the pig and the dog, minoxidil cardiotoxicity is characterized by vascular damage (with hemorrhage in the arterial epicardium) and myocardial necrosis (mostly of the left ventricular papillary muscles). Interestingly, in the dog the atrial lesion is largely restricted to the right atrium, whereas in the pig it is restricted to the left atrium. These lesions can be produced in roughly 50% of the minipigs given 10 mg kg^{-1} of minoxidil for 2 days and sacrificed 48 h after the last dose (Herman et al., 1988, 1989). Herman and colleagues have published extensive descriptions of minoxidil-induced lesions in minipigs in comparison to those produced in dogs (Herman et al., 1988, 1989). The right versus left arterial difference is believed to be due to differences in the anatomical pattern of coronary circulation between two species (Herman et al., 1989).

Minipigs are also sensitive to the cardiotoxic effect of doxorubicin. When given six intravenous injections of either 1.6 or 2.4 mg kg^{-1} of doxorubicin at 3-week intervals, minipigs develop cardiac lesions similar to those seen in dogs, rabbits, and other experimental animals (Herman et al., 1989). The lesion is characterized by cytoplasmic vacuolation and varying degrees of myofibrillar degeneration and loss. Thus, the minipig is sensitive to the cardiotoxic effect of two well-known and extensively studied chemicals. Therefore, it is a suitable nonrodent species for the general assessment of cardiotoxicity.

12.7.6 Advantages and Disadvantages

There is one disadvantage to the use of minipigs. The first is their size. Although minipigs are smaller than regular swine, at maturity they are generally larger than beagle dogs. Among the advantages is the fact that they are long-lived, cooperative animals with well-defined physiological and metabolic characteristics. As they are not either popular companion animals (like dogs) or do not physically resemble humans (like monkeys), minipigs are not specifically discussed in animal “welfare” laws like the other two species. Depending on their final form, new animal welfare regulations could make the space and maintenance costs for dogs and monkeys very prohibitive. This may make minipigs increasingly more attractive as a nonrodent species for general toxicity testing.

12.8 NONHUMAN PRIMATES

Nonhuman primates are often the nonrodent species of choice for safety assessment studies. There are over 500 species of nonhuman primates that differ widely from each other in size and physical characteristics. Most of the monkeys used in experimental research belong to the suborder Anthropeidea and especially to the superfamilies of Ceboidea (marmoset and squirrel monkey) and Cercopithecoidea (macaque, *Papio* species, and rhesus). These have been popular because of (i) assumed better concordance of effects seen to those in man and (ii) smaller weights (and therefore reduced compound requirement). However, predominant factors leading to a decision whether or not to select primates as the nonrodent species for safety evaluation are summarized as follows (Hobson, 2000; Bluemel et al., 2015):

Primates are selected for safety studies because:

- They are the only species which exhibit the human response to the test article.
- Due to smaller body size, they conserve rare or expensive test articles.
- They don't form neutralizing antibodies to the test article.
- They are physiologically more similar to man.
- Regulatory agencies require their use.
- Prior development history dictates species choice.
- Known class effects have previously been seen in primates.

Primates may not be selected for safety studies because of:

- Perceived expense
- Facility and logistic concerns
- Limited supplies
- Biosafety concerns

- Perceived animal rights or animal welfare concerns or pressures
- Tradition and prior development history
- Regulatory agency direction
- Data suggesting that other animal models are the “most sensitive species”

Tradition and cost are the two most frequently quoted reasons for selecting dogs as the second toxicology species instead of nonhuman primates. Many pharmaceutical companies, especially those that primarily work with small molecules, have many years of background data in dogs and do not choose to venture into nonhuman primate research without a compelling reason to do so. Secondary concerns often center on perceived biosafety or animal rights issues. Contrary to conventional wisdom, nonhuman primate studies are often more cost efficient than studies in dogs. Although the purchase cost for nonhuman primates is approximately twice that of dogs, many other factors suggest that the total cost of a nonhuman primate safety study may be less than the cost for a similar-sized dog study. Husbandry costs are higher in dogs because of the US Department of Agriculture (USDA) requirement for exercise. Approximately fourfold more building space is required for a dog study due to the larger cages needed. Perhaps most importantly, the smaller body size of macaques means that the requirement for expensive or scarce test article in a primate study is approximately a third of that of a dog study.

Comparison of Costs for Typical 90-Day Studies Conducted with Nonhuman Primates OR DOGS

| | |
|---|--|
| Animal cost (assume 40 animals) | |
| Cost of dogs | 40 × \$900 = \$36 000 |
| Cost of primates | 40 × \$2 200 = \$88 000 |
| Per diem | |
| Dogs | 90 days × \$11 × 40 = \$39 600 |
| Primates | 90 days × \$7 × 40 = \$25 200 |
| Test article (at \$400 mg ⁻¹) | |
| Dogs | 10 kg × 90 days × 40 animals × 100 µg kg ⁻¹ day ⁻¹ = \$1 440 000 |
| Primates | 4 kg × 90 days × 40 animals × 100 µg kg ⁻¹ day ⁻¹ = \$576 000 |

Clearly, the amount and cost of test article and the length of the study determine which species is the most cost efficient. Because nonhuman primates are phylogenetically closer to man than other species, there is less chance that they will recognize human protein, peptide, or antibody-based biopharmaceuticals as foreign. Thus, they are often selected for safety studies of these materials. Although highly conserved proteins may not be immunogenic in lower species, clearly the formation of neutralizing antibodies to less conserved

proteins during a safety study can confound experimental results (Dean et al., 1990). The formation of neutralizing antibodies to human biopharmaceuticals almost never occurs in chimpanzees (which no longer may be used in pharmaceutical research) but occurs more and more frequently as the nonhuman primate phylogenetic tree is descended. It is generally believed that nonhuman primates’ phylogenetic difference from man is ranked as follows: great apes, baboons, other old-world primates (including macaques), and new-world primates. Clearly, as proteins are modified, they can become immunogenic in all nonhuman primate species, including man (Konrad, 1989).

The physiological similarity and phylogenetic proximity of nonhuman primates to man are often cited as rationale for nonhuman primate selection for safety studies, especially when mechanisms of toxicity or pharmacological action are expected to be closely related to potential physiological reactions in man. Likewise, species selection is often based on the demonstration of pharmacological activity of the test article. Many biopharmaceuticals do not exhibit their intended activity in nonprimate species, whereas small molecules may have activity across all species.

Regulatory agencies sometimes suggest (read dictate) the use of nonhuman primates for certain study designs or drug classes. These requirements are often a surprise to companies when they are first presented. Usually they are derived from confidential data that the regulatory agencies have previously reviewed. Often the regulatory bodies are privy to data that suggest that a class effect is seen in nonhuman primates and not in other species or that nonhuman primates are the most sensitive species. An example was regulatory agency encouragement to perform cardiovascular evaluations of oligonucleotide pharmaceutical candidates in nonhuman primates (Black et al., 1993, 1994). This was based on background information that suggested that oligonucleotides induced complement activation, and the attendant hemodynamic and cardiovascular changes were observed in nonhuman primates but not in other species (Galbraith et al., 1994).

Animal welfare and conservative issues have frequently led to decisions to avoid primate use. Through the mid-1980s many nonhuman primates used in medical research came from wild populations. This led to strong conservationist concerns with the use of monkeys in research. Now, however, almost all nonhuman primates used in research are purpose bred and the conservationist concern has abated. Although there is some animal rights pressure specifically directed against primate use, it is not as formidable as the well-financed and sophisticated efforts to prevent the use of cast-off dogs (pound or shelter dogs or dogs from other class B sources in research). As a consequence, a few pharmaceutical companies are considering switching to nonhuman primates for their second toxicology species.

12.8.1 Environmental and Dietary Requirements

For most nonhuman primate species, room temperatures should be maintained in the range of $75 \pm 5^\circ\text{F}$ with a relative humidity of 40% or greater. These temperature and humidity ranges have been found to be beneficial to the prevention of pneumonia and bloody nose syndrome. Rooms in which monkeys are housed should have 10–15 air changes per hour and be kept under negative pressure in relation to other parts of the building. Where there is significant risk of airborne infection, it is necessary to contain infected animals in units designed to remove the air away from personnel (Mazue and Richez, 1982).

Physical comfort should be an important consideration when determining the appropriate housing for nonhuman primates. For individually housed NHP, the floor area and height of cages should be about $0.28\text{ m}^2 \times 76.20\text{ cm}$ for animals in the weight range of 1–3 kg and $0.40\text{ m}^2 \times 76.20\text{ cm}$ for monkeys weighing 3–10 kg (NRC, 2011). Probably the most common form of commercially available housing is mobile stainless steel rack-mounted cages. Group housing of nonhuman primates used in safety studies is likely to become more common in the future as a result of the US Department of Agriculture's current animal welfare regulations, which require that nonhuman primates have the opportunity for socialization.

Another requirement of the new animal welfare regulations is that any cage or pen in which nonhuman primates are housed must also contain toys, food, or other objects that animals can manipulate, as they would objects in their natural environment. From experience, laboratories have found that toys in themselves are not sufficient since the animals quickly lose interest. Effective enrichment materials include foraging boards (fur-covered objects under which food is buried) and puzzle feeders for more complex foraging.

In many laboratories, monkeys are often fed with commercial pelleted chow ad libitum supplemented with fresh fruits and bread. Like man and the guinea pig, the nonhuman primate cannot synthesize vitamin C and, thus, has a dietary requirement for this vitamin. Powdered chow is an inefficient form for feeding nonhuman primates, because a high percentage of the diet is wasted. Also, dust associated with the chow can cause respiratory problems in some species (NRC, 2003). Even with pelleted or extruded food, monkeys will waste about 50% of the ration sorting through the pellets (Mazue and Richez, 1982). Monkeys should have ad libitum access to water, and it is important that the device (either a water bottle equipped with a sipper tube or an automatic watering system) be fixed securely to the cage to avoid detachment by the animal.

12.8.2 Common Study Protocols

Group sizes and numbers of animals per group for nonhuman primate toxicology studies vary slightly from country to country and from company to company; however, with the

TABLE 12.12 Permissible Dosing Volumes for Nonhuman Primates

| Route | Maximum Permissible Dose ^a |
|------------------|--|
| Intravenous | Varies with duration of administration and character of test article |
| Subcutaneous | 2 mL site ⁻¹ and 5 mL kg ⁻¹ |
| Intramuscular | 0.25 mL site ⁻¹ and 0.5 mL kg ⁻¹ |
| Oral/nasogastric | 5 mL kg ⁻¹ |

^aValues are given for single or infrequent administration. Smaller volumes are appropriate for repeated dosing.

movement for international harmonization, there is a trend toward less variation in study design. Selection of group size is a compromise among regulatory guidelines, cost, statistical power, and conservation of animals.

An example of a protocol for a 4-week safety study in cynomolgus monkeys is shown in Table 12.2. Cynomolgus monkeys are generally in the age range of 1–3 years at study start. A 2- or 4-week study will usually have about four animals per sex per group, not including the mandatory extra animal for evaluating recovery. For the longer-term studies, the number of animals per group will be larger in order to include recovery animals. As with dogs and ferrets, monkeys should be selected for study use based on acceptable pretreatment body weights, clinical laboratory profiles, and physical, ECG, and ophthalmic examinations.

One aspect of study design in nonhuman primates that is not well understood is caused by the variability in the age at which monkeys undergo puberty. Although age at the onset of puberty is highly variable within macaque species, there is a remarkable correlation between body weight and sexual maturity in macaques. Rhesus females undergo menarche at $3000 \pm 200\text{ g}$ irrespective of age, whereas males tend to become sexually mature around 4500 g. This means that a “typical” study is initiated with sexually mature females and sexually immature males. This practice is debatable and is certainly not universally adopted. A few pharmaceutical companies require mature animals of both sexes. Because sexual maturation in males occurs many months later than in females, rearing costs are higher for males and animal numbers may be limited because the younger males may have already been sold with their female birth year counterparts.

12.8.3 General Study Activities

12.8.3.1 Common Dosing Techniques Dosing routes and permissible volumes for nonhuman primates vary between laboratories. The volume limitations from our laboratory are presented in Table 12.12.

Primates offer all of the possible dosing routes available in man, but body size often limits dosing volumes. If volumes for subcutaneous or intramuscular injections exceed those suggested previously, enzyme elevations (particularly

alanine aminotransferase (ALT) and aspartate aminotransferase (AST)) are frequently observed (unpublished results). Continuous infusion techniques in alert animals are available in some laboratories through the use of either programmable backpack pumps or jacket-and-tether systems (Perkin and Stejskal, 1994).

Probably the most common oral route is gavage. This procedure usually requires some degree of physical restraint of the animal (by one or more persons) while a stomach tube for dosing is inserted either orally or intranasally. Other oral dosing methods include buccal, capsule, or addition of the test compound to the drinking water. It is also possible to prepare a modified diet admixture consisting of test compound, diet meal, water, agar, and a jelling agent. This type of preparation will reduce both the feed spillage and the dust normally associated with powdered chow; however, it is susceptible to microbial growth and must be kept frozen or refrigerated (NRC, 2003).

Bolus intravenous, intramuscular, or subcutaneous injections can be administered by a single person by securing the animal's arm through the cage bars (Mazue and Richez, 1982). For safety considerations, many investigators prefer to have the animal physically restrained by a second person before the injection is given. Arterial injections (via the femoral artery) as well as limited or continuous intravenous infusion (via catheterization of the femoral or jugular vein) are other less commonly used parenteral routes in the monkey.

Other routes of administration sometimes used in monkey safety assessment studies are intravaginal dosing, topical application, inhalation, and nasal administration.

12.8.3.2 Clinical Observations and Examinations As with other species, it is important to have a good understanding of the types of normative behaviors and clinical signs that can be seen in normal, untreated monkeys before attempting to make observations in drug-treated animals. Cage-side observations in the monkey should be conducted at least two times daily to monitor general health and behavior. The first observation should be made before cages are cleaned in the morning, and the floors of the cages should be critically examined for signs of blood, abnormal feces, or vomitus. Clinical signs to which investigators should pay particular attention include reduced activity and lethargy, excessive excitation, reduced feed consumption, vomiting, and abnormal feces. If at all possible, the same people should work on a study for its entirety. The behavior of more timorous monkeys can be affected by the presence of unfamiliar personnel, resulting in undesirable clinical signs such as a loss of appetite and lethargy (Evans, 1982). To circumvent these kinds of problems, isolated observation using a video camera system may be a preferable approach.

Physical examinations of monkeys are usually conducted no more than once a week and generally consist of the

measurement of rectal temperature; observation of general demeanor; palpation of the head, thorax, and abdomen; examination of the eyes, ears, and bodily orifices; and testing of the pupillary and patellar reflexes.

12.8.3.3 Feed Consumption As mentioned previously, monkeys tend to scatter their feed, which can make feed consumption difficult to measure. It may be possible to successfully monitor feed consumption in monkeys by using the larger chow biscuits and counting the number of biscuits (or fractions of biscuits) consumed over two consecutive 24 h periods.

12.8.3.4 ECGs and Cardiovascular Measurements The availability of excellent GLP-validated telemetry systems has led to recent increases in the number of cardiovascular safety pharmacology studies conducted in primates. In addition, telemetry is now sometimes included as a design element in standard safety studies. Because of the ability to collect large amounts of high-quality data over an extended time, the total number of animals can often be reduced by appropriate application of telemetry. Indeed, it is often difficult to avoid statistical and reporting problems caused by the temptation to collect too much data using telemetry. Implanted transmitters can function continuously for up to a year without battery replacement while providing data such as blood pressure, heart rate, ECG, body temperature, and activity.

For safety assessment studies, it is preferable to record monkey ECGs in the conscious animal, which, if using standard ECG techniques, requires chairing the animal. Electrocardiographic leads used in this laboratory include II, aVL, and V10. To help reduce emotional tachycardia, it is recommended that there be pretreatment habituation (no more than 10–30 min at least twice before study start) to the chairing and attachment of the surface electrodes. Probably the best and least stressful approach to monitoring ECG activity in conscious monkeys is automatic monitoring using a biotelemetry system. With this system, a transmitter, surgically implanted subcutaneously along the dorsal midline, broadcasts a radio signal encoding the ECG to a receiver mounted on top of the animal's cage, and a computer records the signal at 2 min intervals (Line et al., 1989).

12.8.3.5 Blood and Urine Collections For blood collection, the rhesus can be bled from the saphenous or femoral vein. For female rhesus monkeys, it may not be possible to use the saphenous vein because of the swelling of the sex skin (i.e., the edematous thickening and reddening of the skin over the external genital region, rump, and tail that often extends down the leg to the knee). For the cynomolgus and squirrel monkeys, the veins are very small, and the femoral vein is usually the one of choice. Depending on the species, 2 (marmoset) to 24 (cynomolgus) mL may be collected.

However, experimental designs in primates are often constrained by the limitations in the amount of blood that can be safely and humanely obtained during the course of a study (Walker et al., 2015). With increasing emphasis on obtaining toxicokinetic data during safety studies, these constraints have become more vexing. A guideline for maximum blood withdrawal is $10 \text{ mL kg}^{-1} \text{ day}^{-1}$ (Heiser, 1970). More blood can be collected, but hematocrit should be monitored (Schalm, 2010). These amounts do not approach maximum amounts allowable for humane considerations but do represent the maximum that can be collected without causing anything more than slight decreases in hematologic parameters (notably hematocrit, hemoglobin, and red cell count). Currently, modern catheter material and vascular access ports support long-term frequent blood collection without the catheter clotting and emboli problems experienced in the past. The new vascular access ports remain patent for over a year with routine maintenance. Vascular access ports are particularly useful in nonhuman primates where frequent samples are required because blood sample collection through ports appears to be far less stressful than collection by needle stick. We have also found them useful when evaluating anticoagulant test articles because venipuncture is contraindicated. Sample quality is also superior with ported collections.

Urine collection in nonhuman primates can be measured using either a metabolism cage or a collection pan (equipped with a screen to catch the feces), which is inserted under the floor grid of the home cage. The advantage of the latter system is that the animals do not need to be removed from their home cage; however, care needs to be taken to avoid contamination of the urine with drinking water.

12.8.4 Advantages and Disadvantages

Advantages of using NHP in safety assessment studies include their phylogenetic proximity, as well as their physiological, behavioral, and, often, metabolic similarities to humans (Table 12.13). An example is the similarity between the ovarian cycle of female monkeys and women (Mazue and Richez, 1982), which makes the monkey the

ideal animal model for reproductive studies. Another advantage associated with most species of monkeys used in safety assessment studies is that they are much smaller than nonrodents such as the dog and, thus like the ferret, require less test compound.

The most significant disadvantage to working with monkeys is the serious spontaneous diseases they can carry that are transmissible, and often life threatening, to humans. An example of one such disease is *Herpesvirus simiae* (B virus). B virus is widespread, especially among wild-caught, and to some extent laboratory-bred, rhesus monkeys, including cynomolgus monkeys. Human exposure to B virus occurs during handling of monkeys and monkey tissues (via contact with tears, blood, or saliva of infected animals) and is associated with a high incidence of human mortality (DiGiacomo and Shah, 1972). Other serious to very serious diseases that can be transmitted from monkey to man are Marburg disease, viral hepatitis, tuberculosis, and monkeypox.

12.9 STATISTICS IN LARGE ANIMAL STUDIES

Large animal toxicology studies, typically ranging from 14 days to generally a maximum of 52 weeks, pose different types of statistical problems and open up new possibilities in terms of statistical evaluations. Standard statistical methods used for chronic toxicology studies, such as one-way designs, often do not provide any meaningful insights because of small sample sizes used in large animal studies. The designs for such studies are, generally speaking, non-optimal. As a consequence, an investigator must attempt to use optimal statistical methods to evaluate such studies. Fortunately, for many of the relevant parameters for such studies, there are fewer to none dropouts (if one is careful), and there are repeated measurements on the same parameters of interest, both pre- and posttreatment intervals. Optimality of statistical methods for such studies is then achieved by making use of the longitudinal observations in the analysis. The optimality can be further enhanced by introducing sex as a factor in the evaluation of the data in many such studies.

Many of the standard assumptions in both parametric and distribution-free statistical methods cannot be meaningfully tested in large animal studies because of extremely small sample sizes (which are not necessarily dictated by scientific doctrine but by economic and minimum regulatory requirements). Fortunately, by making use of solid biological as well as statistical judgments, we seem to have made many discoveries in terms of human safety and efficacy in large animal toxicology.

Instead of a conventional textbook-type layout, this discussion will try to focus on various issues in large animal toxicology experiments with plausible examples. One word of caution before we get deeper into our discussion: like in

TABLE 12.13 Use of the Nonhuman Primate in Safety Assessment Studies

Advantages

Small size of many species
Less test material needed than for other nonrodent species
Physiological, behavioral, and, often, metabolic similarities to man

Disadvantages

Limited availability
Cost
Need to develop environmental enrichment programs
High potential for spontaneous diseases

most areas of applied statistics, there really is no gospel in what we will be discussing today. Many statisticians may have variations of the theme to be brought out here.

12.9.1 Reasons for Small Sample Sizes in Large Animal Toxicology

The following are some of the main reasons for having only three to five dogs or monkeys per sex in a typical large animal study:

1. These studies are very expensive. A typical full-fledged study may cost as much as \$30 000 to \$500 000 (for 26 weeks).
2. There is tremendous pressure from the animal rights groups to look for alternatives rather than using dogs and monkeys for investigating purposes.
3. Regulatory agencies throughout the world recognize these two facts and recommend such small sample sizes as minimum requirements. As a consequence, the pharmaceutical and chemical industries are reluctant to expand the scopes of such studies.

12.9.2 Cross-Sectional or Longitudinal Analysis?

Many of the studies we deal with have various parameters, such as body weight, food consumption, clinical chemistry, and hematology, that are collected repeatedly at various pretreatment and posttreatment intervals. Unfortunately, many investigators in the field do not take advantage of this important design feature of such studies. Instead, one finds the literature is full of simple parametric or distribution-free one-way techniques, such as Student's *t*-test, Wilcoxon–Mann–Whitney rank test, one-way analysis of variance (ANOVA) methods, and so on, widely being used, sometimes without satisfaction. The argument then is given that “although there is apparent biological effect (or lack of it), because of small sample sizes and poor statistics, no significant effects can be determined from these data” or something like that. If truth be known, the small sample size part of this argument may be correct; however, no attempts were made to optimize on the statistical methods above using the various pieces of the particular design. The repeated sampling part of the design (repeated measures) is very important for such studies and therefore should be incorporated in the analysis of the data. After all, design of experiment and analysis of data are inseparable. There are advantages and disadvantages of such analyses (the advantages generally outweigh the disadvantages) as described in the following.

12.9.3 Repeated Measures: Advantages

1. Between-subject variations are excluded from the experimental and stochastic errors.

2. Only the within-subject variation is included in the mean square error (*MSE*) term.
3. Each subject becomes its own control.
4. Economizes on the number of subjects in an experiment.
5. Minimizes both type I (false positive) and type II (false negative) error rates, thereby increasing power of the test statistic to be employed while decreasing inconsistent significant effects.

12.9.4 Repeated Measures: Disadvantages

1. Order of the treatment may cause interference which can be avoided by appropriate randomization.
2. There is the possibility of carry-over effects. This is more crucial in Latin square and other crossover designs. Knowledge of pharmacokinetics and metabolism of a compound under study generally helps in avoiding this problem.
3. Exact permutation and distribution-free techniques are not as widely developed as in the cases of one-way methods.
4. Power and sample size computations are a little more difficult to compute than one-way designs.
5. Generally requires computers for performing the analyses using specialized software (not a major issue in most societies nowadays).
6. A little more difficult to interpret the results than their one-way counterparts.

12.9.5 Common Practices in Large Animal Toxicology

Older (and some newer) literature in large animal toxicology is full of two-sample, one-way parametric, and distribution-free techniques. Some of the newer works use repeated measures and even multivariate techniques. The following is a brief exposé of various methods used in the field:

1. One-way analysis of variance/covariance/regression and preplanned and post hoc group comparisons
2. Two-sample Student's *t*-test, Wilcoxon–Mann–Whitney rank test, and so on
3. Graphic display of response over time (as two- or three-dimensional plots)
4. Univariate repeated measures analysis of variance/covariance techniques
5. Multivariate analysis of variance/covariance (MANOVA/MANCOVA) techniques

Methods 1 and 2 in the preceding text should not be preferred in global analyses. Graphic displays have tremendous values as exploratory data analysis (EDA) techniques

with the type of data one encounters in these studies. For formal analyses, one could weigh univariate repeated and other factorial designs against their true multivariate counterparts.

12.9.6 Univariate (Repeated Measures) Techniques: Advantages

1. Easier to compute.
2. Less susceptible to violation of normality.
3. Exact and distribution-free tests are easier to compute.
4. Require smaller sample sizes; there is more power.
5. Very few test statistics to deal with: classical ANOVA F , Greenhouse–Geisser and Huynh–Feldt adjusted df , and ANOVA F .
6. Biologically meaningful and easier to resolve contrasts and multiple comparison tests.
7. Missing values are easily handled.

12.9.7 Univariate (Repeated Measures) Techniques: Disadvantages

1. Susceptible to heteroscedasticity (heterogeneity of variances)
2. Less fancy compared to multivariate techniques

12.9.8 Multivariate Techniques: Advantages

1. Less susceptible to heteroscedasticity
2. Handles multiple dependent variables
3. Real fancy compared to univariate ANOVA/ANCOVA techniques

12.9.9 Multivariate Techniques: Disadvantages

1. More susceptible to violation of normality
2. Less power than univariate ANOVA, particularly with small sample sizes
3. Contrasts and multiple comparisons that are difficult to construct
4. Missing values that are more difficult to handle
5. Computationally more difficult (a mute point nowadays with personal computers)
6. Too many test statistics, sometimes giving contradictory answers, to deal with

12.9.10 Some Other Design Factors to Be Considered in Analysis

Most of the toxicological studies are designed to evaluate efficacy and safety in both sexes. With small sample sizes, one can increase the power efficiency of the particular test statistic by including sex as a factor in a full factorial

analysis (not combining the two sexes) where appropriate. The factorial analysis will reveal whether there is any need to separate the two sexes. The other design fact that should be weighed carefully is the presence of any concomitant variables or covariates. For example, most large animal studies will involve collection of data both prior to the beginning of the experiments as well as after. Thus pretreatment values and other characteristics that control variables (e.g., body weights) may be important covariates in the analysis of the data. There are both advantages and disadvantages in including covariates in the analysis.

12.9.11 Covariates: Advantages

1. Increases precision of an analysis (indirect or statistical control of variability)
2. Correction of bias
3. Elimination of extraneous variation in the data

12.9.12 Covariates: Disadvantages

1. Unequal intra- and intergroup covariate slopes—may actually introduce bias as a consequence.
2. Nonlinearity of covariate slopes—may have the same effect as in the preceding text.
3. In some cases the covariates may be affected by treatment.

An example is shown in Table 12.14. A two-factor analysis of variance for the covariate, as shown in Table 12.15, clearly indicates that the two sexes started with approximately the same means ($p=0.5598$). Moreover, there were no differences between the group means in either sex as indicated by the large tail probabilities for treatment ($p=0.8823$) and sex \times treatment interaction ($p=0.6532$). These facts justify using sex as a factor in the analysis, as was done here.

There are various other ways of examining the variate in question in this case. Let us first examine simple one-way ANOVA of the variate by sex as in Table 12.16. In neither of the two cases was there any indication of significant treatment differences at any reasonable level. Because the two sexes did not show any pretreatment differences based on the two-factor analysis of the covariate, let us combine the two sexes and analyze the data by one-way ANOVA as in Table 12.17. In this case, because of the increased sample sizes for combining the two sexes, there was an indication of some treatment differences ($p=0.0454$). Unfortunately, this analysis assumes that because there were no pretreatment differences between the two sexes, that pattern will hold during the posttreatment period. That often may not be the case because of biological reasons.

The earlier analysis establishes that there was no significant sex difference, as indicated by the tail probabilities for sex

TABLE 12.14 Example 1

| Sex | Control | | Treatment 1 | | Treatment 2 | |
|----------|-----------|---------|-------------|---------|-------------|---------|
| | Covariate | Variate | Covariate | Variate | Covariate | Variate |
| Male | 40 | 95 | 30 | 85 | 50 | 90 |
| | 35 | 80 | 40 | 100 | 40 | 85 |
| | 40 | 95 | 45 | 85 | 40 | 90 |
| | 50 | 105 | 40 | 90 | 30 | 80 |
| | 45 | 100 | 40 | 90 | 40 | 85 |
| Raw mean | 42.0 | 95.0 | 39.0 | 90.0 | 40.0 | 86.0 |
| SD | 5.7 | 9.4 | 5.8 | 6.1 | 7.1 | 4.2 |
| Female | 50 | 100 | 50 | 100 | 45 | 95 |
| | 30 | 95 | 30 | 90 | 30 | 85 |
| | 35 | 95 | 40 | 95 | 25 | 75 |
| | 45 | 110 | 45 | 90 | 50 | 105 |
| | 30 | 88 | 40 | 95 | 35 | 85 |
| Raw mean | 38.0 | 97.6 | 41.0 | 94.0 | 37.0 | 89.0 |
| SD | 9.1 | 8.1 | 7.4 | 4.2 | 10.4 | 11.4 |

TABLE 12.15 Two-Factor Analysis of Variance for the Covariate

| Source | Sum of Squares | df | Mean Squares | <i>F</i> | Tail Probability |
|-----------------|----------------|----|--------------|----------|------------------|
| Mean | 46 807.50000 | 1 | 46 807.50000 | 785.58 | 0.0000 |
| Sex | 20.83333 | 1 | 20.83333 | 0.35 | 0.5598 |
| Treatment | 15.00000 | 2 | 7.50000 | 0.13 | 0.8823 |
| Sex × treatment | 51.66667 | 2 | 25.83333 | 0.43 | 0.6532 |
| Error | 1 430.00000 | 24 | 59.58333 | | |

TABLE 12.16 One-Way Analysis of Variance of the Variable of Sex

| Source | Sum of Squares | df | Mean Squares | <i>F</i> | Tail Probability |
|----------------|----------------|----|---------------|----------|------------------|
| <i>Males</i> | | | | | |
| Mean | 122 401.66667 | 1 | 122 401.66667 | 2576.88 | 0.0000 |
| Treatment | 203.33333 | 2 | 101.66667 | 2.14 | 0.1603 |
| Error | 570.00000 | 12 | 47.50000 | | |
| <i>Females</i> | | | | | |
| Mean | 131 227.26667 | 1 | 131 227.26667 | 1841.36 | 0.0000 |
| Treatment | 186.53333 | 2 | 93.26667 | 1.31 | 0.3061 |
| Error | 855.20000 | 12 | 71.26667 | | |

TABLE 12.17 One-Way Analysis of Variance for Combined Sexes

| Source | Sum of Squares | df | Mean Squares | <i>F</i> | Tail Probability |
|-----------|----------------|----|---------------|----------|------------------|
| Mean | 253 552.13333 | 1 | 253 552.13333 | 4549.999 | 0.0000 |
| Treatment | 387.26667 | 2 | 193.63333 | 3.47 | 0.454 |
| Error | 1504.60000 | 27 | 55.72593 | | |

($p=0.2667$) and sex × treatment interaction ($p=0.9784$) (Table 12.18). There was also some indication that there may have been some treatment effect across the treatment groups in both sexes ($p=0.0559$). Examination of the variate means indicated that both sexes seemed to have lower means than their respective controls. The picture was

clouded by the fact that there was such a slightly lower tendency, though not very consistent, in the covariate means as well. Under this circumstance, it is more appropriate to take both the covariate and the variate into any optimal analysis. Table 12.19 shows an analysis of covariance for the factorial model.

TABLE 12.18 Two-Factor Analysis of Variance with Sex as a Factor

| Source | Sum of Squares | df | Mean Squares | F | Tail Probability |
|-----------------|----------------|----|---------------|---------|------------------|
| Mean | 253 552.13333 | 1 | 253 552.13333 | 4269.75 | 0.0000 |
| Sex | 76.80000 | 1 | 76.80000 | 1.29 | 0.2667 |
| Treatment | 387.26667 | 2 | 193.63333 | 3.26 | 0.0559 |
| Sex × treatment | 2.60000 | 2 | 1.30000 | 0.02 | 0.9784 |
| Error | 125.20000 | 24 | 59.38333 | | |

TABLE 12.19 Analysis of Covariance of the Factorial Model

| Source | Sum of Squares | df | Mean Squares | F | Tail Probability |
|-----------------|----------------|----|--------------|-------|------------------|
| Mean | 147.42310 | 1 | 147.42310 | 5.65 | 0.0262 |
| Sex | 292.81064 | 2 | 146.40532 | 5.61 | 0.0104 |
| Treatment | 14.41235 | 2 | 7.20617 | 0.28 | 0.7613 |
| Sex × treatment | 824.75245 | 1 | 824.75245 | 31.59 | 0.0000 |
| Error | 600.44755 | 23 | 26.10642 | | |

| | Control | Treatment 1 | Treatment 2 |
|---|----------|-------------|-------------|
| Adjusted cell means and standard errors (males) | | | |
| Mean | 93.10140 | 90.37972 | 85.62028 |
| Standard error | 2.30985 | 2.28601 | 2.28601 |
| Adjusted cell means and standard errors (females) | | | |
| Mean | 98.73916 | 92.86084 | 90.89860 |
| Standard error | 2.29398 | 2.29398 | 2.30985 |

As the ANCOVA table indicates, there was a definite significant treatment effect ($p=0.0104$), but this effect was not sex specific because there was no significant sex × treatment interaction ($p=0.7613$). Furthermore, there was a significant difference between the two sexes in terms of magnitude but not in the direction of the effect. These findings are apparent in the covariate-adjusted means in all groups in both sexes. The magnitude of the treatment effect became amplified by introducing the covariate in the model. As can be seen from the two ANOVA and ANCOVA tables in the preceding text, despite the fact that the ANCOVA error term lost one degree of freedom ($df=23$) as opposed to the ANOVA error term ($df=24$), the former gains some edge over the latter because of increased precision. Precision in this context is defined as the ratio between the *MSEs* of ANOVA and ANCOVA. For this example,

$$\begin{aligned}
 \text{Precision} &= \frac{1/\text{MSE}_{\text{ANCOVA}}}{1/\text{MSE}_{\text{ANOVA}}} \\
 &= \frac{\text{MSE}_{\text{ANOVA}}}{\text{MSE}_{\text{ANCOVA}}} \\
 &= \frac{59.38333}{26.10642} \sim 2.3
 \end{aligned}$$

In other words, we have gained about 2.3-fold precision by ANCOVA over ANOVA in resolving treatment effect.

With the advent of powerful personal computers and the availability of sophisticated “do-it-all” statistical packages, there is a trend among nonstatisticians (even some statisticians) to accept the results from these packages without contemplating twice. Many of these packages have flexible features that allow one to perform different types of analyses with the same data set, inappropriately or appropriately sometimes. What popular statistical packages give is not necessarily correct statistics or they may not be correct under specific designs. Some programs, for example, BMDP’s 2V (1992), have “intelligence” built into them whereby they can identify the design based on the data matrix. By following the data matrix setup specified in the manual correctly, one can simply press the button and get the appropriate analysis needed. On the other hand, incorrect specification of the data matrix will produce incorrect results (although some programs, such as 2V, will often give an error message or prompt to make sure one wants what one is asking for; some such as SAS’s PROC GLM may not and give results that are not even remotely related to the design). In other words, one must know some statistics and must be well versed in the features of the particular package before using them. The one-time famous mathematician, statistician, composer, pianist, singer, producer, and recording artist Tom Lehrer (1959), in one of his famous monologues, said, “Life is a sewer; what one gets out of it depends on what one puts into it.” Statistical packages are exactly like that.

12.9.13 Missing Values

All investigators know that missing values are a nuisance. They also create statistical nightmares. Classical statistical techniques were not geared toward having missing values in experiments. Unfortunately, in real-life situations, it just happens. Animals may die or are censored for various reasons. There are various techniques of calculating missing values for specific designs (Miller, 1981) just like there are for extreme values or outliers (SAS, 2015). In neither case there is a unique way of handling them that is completely agreed upon by statisticians. One should remember that every time a missing value is computed and used in statistical analyses, one loses a degree of freedom. In large animal toxicology, with small sampling sizes, one must be very careful about dealing with missing values. In a repeated measures analysis, if one observation is missing from an animal during one interval, classical techniques automatically will exclude observations from that animal for all remaining intervals. Newer techniques based on regression or imputation have been developed in recent years and have been implemented in popular packages such as BMDP (5V) or SAS (PROC MIXED). Within a single package, there may be various techniques based on assumptions on covariance structures (unstructured, compound symmetry, etc.) and statistical algorithms (maximum likelihood, restricted minimum likelihood, etc.). The results sometimes could be very different under the same assumptions and algorithms. As a result, given the same compound symmetry assumption and using the same restricted maximum likelihood (REML) algorithm, two well-known programs give different quantitative results. These methods are still experimental in nature and should not be taken for granted. Actually, the BMDP manual clearly warns users about the nature of this method. Consequently, the best way to avoid confusion is to try to make sure that missing values do not occur in key parameters in *large animal* studies (Thakur, 2000).

12.10 SUMMARY

While there are advantages and disadvantages associated with all three of these nonrodent species, the dog is probably the nonrodent species most frequently used in safety assessment studies. This is because dogs are relatively docile and even tempered, they are generally easier to obtain and relatively less expensive than monkeys, they carry less serious diseases than the ferret and the monkey, and they have a more extensive historical database in safety studies. It should be noted, however, that if the technical and health problems associated with the ferret can be overcome, its small size in terms of compound requirements, cost, and housing may make it an ideal nonrodent species for future use in safety assessment studies.

REFERENCES

- Andrews PLR, Bower AJ, Illman O. (1979) Some aspects of the physiology and anatomy of the cardiovascular system of the ferret, *Mustela putorius furo*. *Lab. Anim.* 13:215–220.
- Animal and Plant Health Inspection Service (APHIS) (1989) 54 FR 36112, August 31, 1989, as amended at 78 FR 57227, September 18, 2013, Animal Welfare Act. United States Department of Agriculture (USDA), Washington, DC.
- Barnett J, Hensworth P. (1986) The impact of handling and environmental factors in the stress response and its consequences in swine. *Lab. Anim. Sci.* 36:366–369.
- Black LE, DeGeorge JJ, Cavagnaro JA, Jordan A, Ahn CH. (1993) Regulatory considerations for evaluating the pharmacology and toxicology of antisense drugs. *Antisense Res. Dev.* 3:399–404.
- Black LE, Farrelly JG, Cavagnaro JA, Ahn CH, DeGeorge JJ, Taylor AS, DeFelice AF, Jordan A. (1994) Regulatory considerations for oligonucleotide drugs: updated recommendations for pharmacology and toxicology studies. *Antisense Res. Dev.* 4:299–301.
- Blumel J, Korte S, Schenck E, Weinbauer GF. (2015) *The Nonhuman Primate in Nonclinical Drug Development and Safety Assessment*. Academic Press, San Diego, CA.
- Brechbuler T, Kaeslin M, Wyler F. (1984) Reference values of various blood constituents in young minipigs. *J. Clin. Chem. Clin. Biochem.* 22:301–304.
- Burks M, Ymbleson M, Hicklin K, Hutcheson D, Middleton C. (1977) Age and sex related changes of hematologic parameters in sinclair (S-1) miniature swine. *Growth.* 41:51–62.
- Campbell SA, Hughes HC, Griffin HE, Landis MS, Mallon FM. (1988) Some effects of limited exercise on purpose-bred beagles. *Am. J. Vet. Res.* 49:1298–1301.
- Clausing P, Beitz H, Gericke S, Solecki S. (1986) On the usefulness of minipigs in toxicology testing of pesticides. *Arch. Toxicol.* 9(Suppl):225–271.
- Dean JH, Cornacoff JB, Labrie T, Barbolt TA. (1990) Assessment of immune responses in rodents and non-human primates—implications in preclinical evaluation of proteins. In: *Preclinical Evaluation of Peptides and Recombinant Proteins* (Sundwall A, Ed.). Malmo Skogs Grafiska AB, Stockholm, pp. 23–34.
- DeSesso JM, Williams AL. (2008) Contrasting the gastrointestinal tracts of mammals: factors that influence absorption. *Annu. Rep. Med. Chem.* 43:353–371.
- Detweiler DK. (1981) The use of electrocardiography in toxicological studies with beagle dogs. In: *Cardiac Toxicity* (Balazs T, Ed.). CRC Press, Boca Raton, FL, pp. 33–82.
- Detweiler DK, Patterson DF, Buchanan JN, Knight DH. (1979) The cardiovascular system. In: *Canine Medicine* (Catcott EJ, Ed.), Vol. 2. American Veterinary Publications, Santa Barbara, CA.
- Diehl KH, Hull R, Morton D, Pfister R, Rabmampianina Y, Smith D, Vidal J-M, Vortembosch C. (2001) A good practice guide to the administration of substances and removal of blood, including routes and volumes. *J. Appl. Toxicol.* 21:21–23.
- DiGiacomo RF, Shah KV. (1972) Virtual absence of infection with *Herpesvirus simiae* in colony-reared rhesus monkeys (*Macaca*

- mulatta*) with a literature review on antibody prevalence in natural and laboratory populations. *Lab. Anim. Sci.* 27:61–67.
- Earl FL, Miller E, Van Loon EJ. (1973) Teratogenic research in beagle dogs and miniature swine. In: *The Laboratory Animal in Drug Testing: 5th Symposium, Internal Committee on Laboratory Animals, September 19–21, 1971 in Hanover, Germany* (Spiegel A, Ed.). Fisher, Stuttgart, Germany, pp. 233–247.
- Ellegaard L, Jorgensen KD, Klastbrück S, Hansen AK, Svendsen O. (1995) Haematologic and clinical chemical values in three- and six-month old Göttinger minipigs. *Scand. J. Lab. Anim. Sci.* 22:239–248.
- Evans RH. (1982) *Nonhuman Primates*. Ralston Purina Co., St. Louis, MO.
- Fox JG. (1988) *Biology and Diseases of the Ferret*. Lea and Febiger, Philadelphia.
- Fox JG, Marini RP. (2014) *Biology and Diseases of the Ferret*, 3rd Ed. Wiley Blackwell, Philadelphia, PA.
- Freudenthal R, Leber P, Emmerling D, Kerchner G, Campbell D. (1976) Characterization of the hepatic microsomal mixed-function oxidase system in miniature pigs. *Drug Metab. Dispos.* 4:25–27.
- Gad SC. (2000) Pigs and ferrets as models in toxicology and biological safety assessment. *Int. J. Toxicol.* 19:149–168.
- Gad SC. (2015) *Animal Models in Toxicology*, 3rd Ed. CRC Press, Boca Raton, FL.
- Gad SC, Gad SD, Gad SE. (2003) The dog functional observational (FOB) for use in pharmaceutical safety evaluation studies. *Int. J. Toxicol.* 22:1–8.
- Galbraith WM, Hobson WC, Giclas PC, Schechter PJ, Agrawal S. (1994) Complement activation and hemodynamic changes following intravenous administration of phosphorothioate oligonucleotides in the monkey. *Antisense Res. Dev.* 4:201–206.
- Gulamhusein AP, Harrison-Sage C, Beck F, Al-Alousi A. (1980) Salicylate-induced teratogenesis in the ferret. *Life Sci.* 27:1799–1805.
- Haggerty GC, Miller G, Port C, Gad S. (1989) *Four-week oral toxicity study of diclofenac in the ferret*. Presented at the National Meeting of the American College of Toxicology, October 30–November 1, 1989, Williamsburg, VA.
- Hanhijarvi H, Nevalainen T, Mannisto P. (1985) A six-month dermal irritation test with anthralins in the Göttinger miniature swine. *Arch. Toxicol.* 8(Suppl):463–468.
- Hart JE. (1986) The ferret as a replacement for the dog in toxicity studies. *Anim. Technol.* 37:201–206.
- Heiser HJ. (1970) *Atlas of Comparative Hematology*. Academic Press, New York.
- Herman E, Ferrans V, Young R. (1988) Examination of minoxidil-induced acute cardiotoxicity in miniature swine. *Toxicology*. 48:41–51.
- Herman E, Ferrans V, Young R, Balazs T. (1989) A comparative study in minoxidil-induced myocardial lesions in beagle dogs and miniature swine. *Toxicol. Pathol.* 17:189–191.
- Hoar RM. (1984) Use of ferrets in toxicity testing. *J. Am. Coll. Toxicol.* 3:325–329.
- Hobson W. (2000) Safety assessment studies in nonhuman primates. *Int. J. Toxicol.* 19:141–147.
- Hunsaker H, Norden S, Allen K. (1984) An inexpensive caging method for miniature swine suitable for trace-element studies. *Lab. Anim. Sci.* 2:386–387.
- Hutcheson D, Tumbleson M, Middleton C. (1979) Serum electrolyte concentrations in Sinclair (S-1) miniature swine from 1 through 36 months of age. *Growth*. 43:62–70.
- Jackson RK, Kieffer VA, Sauber JJ, King GL. (1988) A tether-restraint system for blood collection from ferrets. *Lab. Anim. Sci.* 38:625–628.
- Khan M. (1984) Minipig: advantages and disadvantages as a model in toxicity testing. *J. Am. Coll. Toxicol.* 3:337–342.
- Koch W, Windt H, Walles M, Borlak J, Clausen P. (2001) Inhalation studies with the Göttinger minipig. *Inhal. Toxicol.* 13:249–259.
- Konrad M. (1989) Immunogenicity of proteins administered to humans for therapeutic purposes. *Trends Biotechnol.* 7(7):175–179.
- Lee K. (1986) Swine as animal models in cardiovascular research. *Swine Biomed. Res.* 3:1491–1496.
- Lehrer T. (1959) *An Evening Wasted with Tom Lehrer*. Lehrer Records TL202. Decca, London.
- Line SW, Morgan KN, Markowitz H, Strong S. (1989) Influence of cage size on heart rate and behavior in rhesus monkeys. *Am. J. Vet. Res.* 50:1523–1526.
- Mannisto P, Havas A, Haasio K, Hahnijarvi H, Mustakallio K. (1984) Skin irritation by dithranol (anthralin) and its 10-aceyl analogues in three animal models. *Contact Dermatitis*. 10:140–145.
- Mazue G, Richez P. (1982) Problems in utilizing monkeys in toxicology. In: *Animals in Toxicological Research* (Bartosek I, Ed.). Raven Press, New York, pp. 147–163.
- McAnulty PA, Dayan AD, Ganderup N-C, Hastings KL. (2010) *The Minipig in Biomedical Research*. CRC Press, Boca Raton, FL.
- McLain D, McCartney M, Giovanetto S, Martis L, Greener Y, Youkilis E. (1987) Assessment of the subchronic intravenous toxicity and disposition of (¹⁴C) acrolein in the rat and the acute and subchronic toxicity in ferrets. *Toxicologist*. 7:208.
- Miller RG Jr. (1981) *Simultaneous Statistical Inference*, 2nd Ed. Springer-Verlag, New York.
- Moody KD, Bowman TA, Lang CM. (1985) Laboratory management of the ferret for biomedical research. *Lab. Anim. Sci.* 35:272–279.
- Morton F. (2010) In: *Ferrets: A Complete Pet Owner's Manual* (Vriends MM, consult. Ed.). Barron's Educational Series, Hauppauge, NY.
- Mueller D, Nelles J, Deparde E, Arni P. (1980) The activity of S-9 liver fractions from seven species in salmonella/mammalian-microsome mutagenicity tests. *Mutat. Res.* 70:279–300.
- National Research Council (NRC). (2003) *Nutrient Requirements of Nonhuman Primates*. Ad Hoc Committee on Nonhuman Primate Nutrition, National Academies Press, Washington, DC.
- National Research Council (NRC). (2006) *Nutrient Requirements of Dogs and Cats*. Subcommittee on Dog and Cat Nutrition, National Academies Press, Washington, DC.

- National Research Council (NRC). (2011) *Guide for the Care and Use of Laboratory Animals*, 8th Ed. National Academies Press, Washington, DC.
- Norris WP, Poole CM, Frye RJ, Kretz ND. (1968) *A Study of Thermoregulatory Capabilities of Normal, Aged and Irradiated Beagles*. ANL-7535. Argonne National Laboratory, Argonne, IL, pp. 166–169.
- Oldigs B. (1986) Effects of internal factors upon hematological and clinical chemical parameters in the Göttinger miniature pig. *Swine Biomed. Res.* 2:809–813.
- Olson H, Betton G, Robinson D, Karluss T, Monro A, Kolaja G, Lilly P, Sanders J, Sipes G, Bracken W, Dorato M, van Deun K, Smith P, Berger B, Heller A. (2000) Concordance of the toxicity of pharmaceuticals in humans and in animals. *Regul. Toxicol. Pharmacol.* 32:56–67.
- Panepinto L. (1986) Laboratory methodology and management of swine in biomedical research. *Swine Biomed. Res.* 1:97–109.
- Panepinto L, Phillips R. (1986) The Yucatan miniature pig: characterization and utilization in biomedical research. *Lab. Anim. Sci.* 36:344–347.
- Parsons A, Wells R. (1986) Serum biochemistry of healthy Yucatan miniature pigs. *Lab. Anim. Sci.* 36:428–430.
- Peggins J, Shipley L, Weiner M. (1984) Characterization of age-related changes in hepatic drug metabolism in miniature swine. *Drug Metab. Dispos.* 12:379–381.
- Perkin CJ, Stejskal R. (1994) Intravenous infusion in dogs and primates. *J. Am. Coll. Toxicol.* 13:40–47.
- Phillips R, Tumbleson M. (1986) Models. *Swine Biomed. Res.* 1:437–440.
- Radin M, Weiser M, Frettmann M. (1986) Hematologic and serum biochemical values for Yucatan miniature swine. *Lab. Anim. Sci.* 36:425–427.
- Rendic S, Ruf S, Weber P, Kajfez F. (1984) Cimetidine and ranitidine: their interaction with human and pig liver microsomes and with purified cytochrome P-450. *Eur. J. Drug Metab. Pharmacokinet.* 9:195–200.
- Rettie A, Bogucki B, Lim I, Meier P. (1990) Stereoselective sulfadiazine oxidation of a series of alkyl P-tolyl sulfides by microsomal and purified flavin-containing monooxygenases. *Mol. Pharmacol.* 37:643–651.
- Sambuco P. (1985) Miniature swine as an animal model in photodermatology: factors influencing sunburn cell formation. *Photodermatology.* 2:144–150.
- Schalm DW. (2010) Materials and methods for the study of the blood including brief comments on factors to be considered in interpretation. In: *Schalm's Veterinary Hematology* (Weiss DJ and Wardrup KJ, Eds.), 6th Ed. Lea and Febiger, Philadelphia, PA, p. 24.
- Schneider N, Bradley S, Andersen M. (1977) Toxicology of cyclomethylenetrinitramine: distribution and metabolisms in the rat and the miniature swine. *Toxicol. Appl. Pharmacol.* 39:531–541.
- Smith G, Watkins J, Thompson T, Rozman K, Klassen C. (1984) Oxidative and conjugative metabolism of xenobiotics by livers of cattle, sheep, swine and rats. *J. Anim. Sci.* 58:386–395.
- SAS. (2015) Statistical Analysis System (SAS). SAS Institute, Cary, NC.
- Swindle M, Smith A, Hepburn B. (1988) Swine as models in experimental surgery. *J. Invest. Surg.* 1:65–79.
- Thakur AK. (2000) Statistical issues in large animal toxicology. *Int. J. Toxicol.* 19:133–140.
- Thornton PC, Wright PA, Sacra PJ, Goodier TEW. (1979) The ferret, *Mustela putorius furo*, as a new species in toxicology. *Lab. Anim.* 13:119–124.
- Tynes R, Hodgson E. (1984) The measurement of FAD-containing oxygenase activity in microsomes containing cytochrome P-450. *Xenobiotica.* 14:515–520.
- Van Ryzin R, Trapold J. (1980) The toxicology profile of the anti-inflammatory drug proquazone in animals. *Drug Chem. Toxicol.* 3:361–379.
- Van Vleet J, Herman E, Ferrans V. (1984) Cardiac morphologic alterations in acute minoxidil cardiotoxicity in miniature swine. *Exp. Mol. Pathol.* 41:10–25.
- Walker C, Bentley P, Oesch P. (1978) Phylogenetic distribution of epoxide hydratase in different vertebrate species, strains and tissues measured using three substrates. *Biochim. Biophys. Acta.* 539:427–434.
- Walker MD, Templin MV, Bernal JC, Jacobson SB, Gad SC, Kemper CJ. (2015) Primates. In: *Animal Models in Toxicology* (Gad SC, Ed.), 3rd Ed. CRC Press, Boca Raton, FL, pp. 671–762.
- Watkins J, Klaassen C. (1986) Xenobiotic biotransformation in livestock: comparison to other species commonly used in toxicity testing. *J. Anim. Sci.* 63:933–942.
- Wilson MS, Donnoghue PND. (1982) A mobile rack of cages for ferrets (*Mustela putorius furo*). *Lab. Anim.* 16:278–288.

13

DEVELOPMENTAL AND REPRODUCTIVE TOXICITY TESTING

13.1 INTRODUCTION

The objective of evaluating the potential developmental and reproductive toxicity of drug candidates in laboratory animals is to predict which agents would adversely affect the ability to achieve and maintain pregnancy and for normal development of offspring in humans and to allow evaluation of the potential risks to patients at levels of clinical use. The need for specific developmental testing of drugs evolved over a period of 40 years, with four historic events heightening concern:

- 1941—Human malformations linked to rubella virus.
- 1950s—Methylmercury recognized as developmental toxicant (from Japanese environmental exposure).
- 1960s—Thalidomide (a sedative and antinausea drug) found to cause human malformations.
- 1970s—Alcohol related to developmental effects—fetal alcohol syndrome (FAS). Lead came to be a similar concern.

This testing involves an extensive battery of studies based on experience with animal models (see Table 13.1), historically on guidelines promulgated by the Food and Drug Administration (FDA) of the United States in 1966 (see Food and Drug Administration, 1966, 1984, 2000; D'Aguanno, 1973), and subsequently modified by ICH. These guidelines established three basic types of studies, segments I, II, and III, that are based on dosing during sequential phases of the reproductive cycle. These guidelines represented a dramatic increase in the extent and

sophistication of testing expected of new drug candidates. The impetus for this intensified interest was the tragic epidemic of phocomelia and other congenital malformations caused in the early 1960s by the exposure of pregnant women to the sedative thalidomide (for an excellent discussion of the history of the thalidomide tragedy, see Schardein, 2000). Table 13.2 presents the most recent guidelines.

The types of developmental and reproductive toxicity studies performed prior to 1993 and the methods used have been extensively documented (see Palmer, 1981; Christian, 1983; Persaud, 1985; Schardein, 1988; Christian and Hoberman, 1989; Khera, 1991; DeSesso and Willhite, 2009; Makori et al., 2009; Cooper and Goldman, 2010; Kapp, 2010; Klinefelter and Veeramachaneni, 2014; York et al., 2014). Since June 20, 1979, the FDA has required that these studies be conducted according to Good Laboratory Practice (GLP) regulations (see FDA, 1978, 2007). The conduct of these studies had been complicated by the need to satisfy worldwide regulatory guidelines that varied from country to country. As a result, studies were conducted for regulatory purposes that, from a scientific viewpoint, were redundant, superfluous, and/or unnecessarily complex. This situation was changed in 1993 when the International Conference on Harmonisation of Technical Requirements for Registration of Pharmaceuticals for Human Use (ICH) standardized worldwide requirements in the guideline “Detection of Toxicity to Reproduction for Medicinal Products” (S5 A and B).

This chapter briefly describes the current standard study designs and then focuses on current issues in developmental and reproductive toxicity testing (also see Hood, 2012).

TABLE 13.1 Teratological Classification in Animal Species of Groups of Substances Universally Recognized as Human Teratogens: Summary of Classifications by Animal Species

| Species | Substance Group | | | | | | | | | | | | | | Results | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
|---------|---------------------|----------------------------|------------------------------|------------|------------|------------|-------------------|------------|------------|------------|-------------------------|-------------------|----------------------------|-------------------|---------|------|---------------|-------|----|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|
| | Androgenic Hormones | Anticancer Antimetabolites | Anticancer Alkylating Agents | | | | Anticancer Diones | | | | Coumarin Anticoagulants | Antithyroid Drugs | Aminoglycoside Antibiotics | Vitamin A Analogs | | PCBs | Tetracyclines | +/-/- | %+ | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| | | | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | | | | | | | | | | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer | Anticancer |

Eleven groups of substances known to be teratogenic in humans are listed above, along with their teratogenic classifications in 12 species of animal (Schardein, 2000). + = teratogenic; ± = variably teratogenic; - = not teratogenic. No entry = no classification available. The number of positive, variable, and negative results in each species for all groups of substances is summarized, along with the combined percentage of positive results for all groups of substances in each species. The maximum and minimum concordance of results with respect to human classifications are 100 and 0% in the ferret and the cat, respectively, although only 2 results are listed for both of these species. While positive predictability is 75% for the hamster, it is just 40% for the rabbit, which also exhibits a false-negative rate of 40%. The mean positive predictability rate in those species tested for at least 9 of the 21 agents (viz., the first six categories, i.e., mouse, rat, rabbit, hamster, primate, dog) was under 55%, and the number of equivocal results remained high across these six species at just under 25%. Taking all 70 results into account leads to an overall positive rate of only 56%, only slightly better than by chance.

TABLE 13.2 Current Regulatory Guidelines—ICH and FDA

| Medical Agents | |
|--------------------|--|
| ICH S5A(R2) (2000) | Detection of Toxicity to Reproduction for Medicinal Products (ICH 55) |
| ICH S5B(R2) (2000) | Detection of Toxicity to Reproduction for Medicinal Products. (ICH 55) |
| FDA | International Conference on Harmonisation: Guideline on Detection of Toxicity to Reproduction for Medicinal Products. Federal Register, September 22, 1994, Vol. 59, No. 183 |
| FDA | International Conference on Harmonisation: Guideline on Detection of Toxicity to Reproduction for Medicinal Products; Addendum on Toxicity to Male Fertility. Federal Register, April 5, 1996, Vol. 61, No. 67 |

13.2 ICH STUDY DESIGNS

The ICH S5 A and B guideline allows for various combinations of studies. The studies conducted must include evaluation of the following components:

1. Male and female fertility and early embryonic development to implantation
2. Embryo–fetal development
3. Pre- and postnatal development including maternal function

These components would normally be evaluated in a rodent species (preferably rats) and, in addition, embryo–fetal development would be evaluated in a second species, typically the rabbit. The “most probable option” in the ICH guideline is the case where three rodent studies would be conducted that separately addressed each of the components listed in the preceding text. These study designs are described in the succeeding text. Table 13.3 presents a comparison of

TABLE 13.3 Comparison of ICH Stages and Study Types with Similar Observations Made

| ICH Stage | FDA Guidelines | Great Britain and EEC Guidelines | Japanese Guidelines | EPA OPPTS, OECD, and FDA Redbook Guidelines |
|--|--|--|---------------------------|--|
| A—Premating to conception: reproductive functions in adult animals, including development and maturation of gametes, mating behavior, and fertilization | Segment I | Segment I | Segment I | Multigeneration One generation |
| B—Conception to implantation: reproductive functions in the adult female and preimplantation and implantation stages of the conceptus | Segment I | Segment I | Segment I | Multigeneration One generation Developmental Toxicity |
| C—Implantation to closure of the hard palate: adult female reproductive functions and development of the embryo through major organ formation | Segment I Segment II | Segment I Segment II | Segment II | Multigeneration One generation Developmental Toxicity Developmental Neurotoxicity |
| D—Closure of the hard palate to the end of pregnancy: adult female reproductive function, fetal development, and growth and organ development and growth | Segment I Segment II Segment III | Segment I Segment II | Segment II | Multigeneration One generation Developmental Toxicity Developmental Neurotoxicity |
| E—Birth to weaning: adult female reproduction function, adaptation of the neonate to extrauterine life, including preweaning development and growth (postnatal age optimally based on postcoital age) | Segment I Segment III Pediatric | Segment I Segment II Segment III | Segment II Segment III | Multigeneration One generation Developmental Toxicity Developmental Neurotoxicity |
| F—Weaning to sexual maturity: (pediatric evaluation when treated) postweaning development and growth, adaptation to independent life, and attainment of full sexual development | Pediatric | Segment I | Segment II Segment III | Multigeneration Developmental Neurotoxicity Developmental Immunotoxicity |

Bolded information indicates treatment interval.

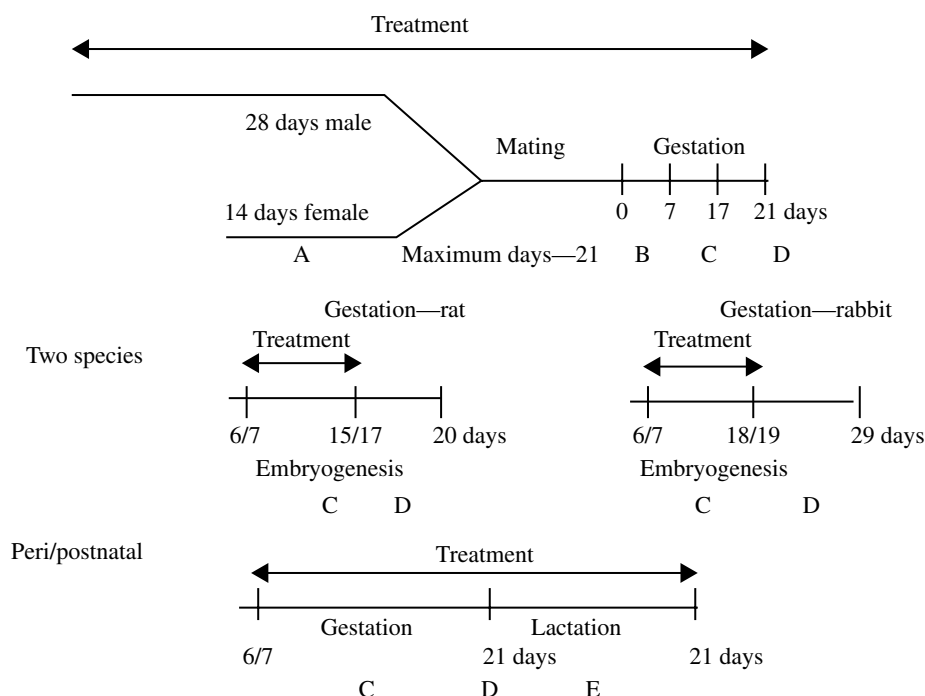


FIGURE 13.1 ICH line charts of nonclinical reproductive and developmental toxicology studies. A, B, C, D, and E here refer to ICH stages as explained in Table 13.3.

ICH, FDA, European, and Japanese guidelines. The day of insemination or detection of evidence of mating is considered day 0 of gestation, and the day of birth is considered postpartum and postnatal day 0. Figure 13.1 presents line charts for the ICH stage study designs.

13.2.1 Male and Female Fertility and Early Embryonic Development to Implantation

The purpose of this component is to assess the effects that result from treatment during maturation of gametes, during cohabitation, and, in females, during gestation up through the time of embryo implantation (typically the last dose on day 6 of gestation) (DeSesso and Willhite, 2009; Gupta, 2011; York et al., 2014). Assuming that the findings from a toxicity study of at least 1 month in duration do not contraindicate, the treatment period begins in males 4 weeks before male/female cohabitation and, in females, 2 weeks prior to cohabitation. A group size of 16–24 litters would generally be considered acceptable.

Minimal in-life observations include:

1. Clinical signs and mortality daily
2. Body weight twice weekly
3. Food consumption weekly

4. Vaginal cytology daily during cohabitation
5. Valuable target effects seen in previous toxicity studies

Females are sacrificed after the middle of the gestation period. Males are sacrificed at any time after the end of the cohabitation period, but it is generally advisable to retain the males until after the outcome of the first mating is known, to ensure that a repeat cohabitation with untreated females will not be needed to determine if an observed effect on mating performance is a male effect. Males are treated until termination. Terminal examination of adults includes:

1. Necropsy
2. Preservation of organs with gross changes and sufficient control organs for comparison
3. Preservation of testes, epididymides, ovaries, and uteri
4. Sperm count and sperm viability
5. Count of corpora lutea and implantation sites
6. Count of live and dead conceptuses

Among the study designs conducted before the ICH guidelines, the segment I fertility study conducted according to Japanese guidelines is most similar to this ICH study design. The major differences are the shortening of the treatment period of males prior to cohabitation from the duration of

spermatogenesis (60–80 days) to 4 weeks and the addition of sperm evaluation. The justifications given for shortening the treatment period of males are the following:

1. Careful organ weight and histopathological evaluation of testes in general toxicity studies will detect most testicular toxins.
2. Fertility is an insensitive measure of testicular effects.
3. Compounds known to affect spermatogenesis generally exert their effects during the first 4 weeks of treatment.

Sperm counts can be performed with sperm from either the testis or the epididymis. Sperm motility is commonly being treated as a measure of sperm viability. The addition of sperm evaluation greatly increases the sensitivity of the study to detect effects on sperm maturation, and the current study design will likely detect more male effects than previous designs even though the treatment period has been shortened.

13.2.2 Embryo–Fetal Development

The purpose of this component is to detect anatomical effects on the developing conceptus by treating during the period of organogenesis from implantation to closure of the secondary palate. The study design is very similar to the historical segment II developmental toxicity study. A group size of 16–24 litters would generally be considered acceptable. The following is demonstrated in Table 13.4.

Minimal in-life observations include:

1. Clinical signs and mortality daily
2. Body weight twice weekly
3. Food consumption weekly
4. Valuable target effects seen in previous toxicity studies

Females are sacrificed at the end of the gestation period, about one day prior to parturition (day 20 or 21 for rats, day 28 or 29 for rabbits, and day 17 or 18 for mice). Terminal examinations include:

1. Necropsy
2. Preservation of organs with gross changes and sufficient control organs for comparison
3. Count of corpora lutea and live and dead implantations

TABLE 13.4 Recommendations for Acceptable Group Size (Litters)

| | Rat | Rabbit | Mouse |
|-------------------------------------|-------|--------|-------|
| Treatment period (gestational days) | 6–17 | 6–18 | 6–15 |
| Group size (mated or inseminated) | 16–24 | 16–24 | 16–24 |

4. Fetal body weight
5. External, visceral, and skeletal examination of fetuses
6. Gross evaluation of placenta

A minimum of 50% of fetuses are to be examined for visceral alterations and a minimum of 50% for skeletal abnormalities. When a fresh tissue microdissection technique is being used for the visceral examination of rabbit fetuses, all fetuses should be examined for both visceral and skeletal abnormalities.

Interpretation of results requires understanding and utilizing the following definitions:

Malformation Structural change that is likely to be permanent and detrimental to the survival or well-being of the fetus, in the species/strain of animal being tested

Alteration Change that is, in isolation, unlikely to be detrimental to the survival or well-being of the fetus, in the species/strain of animal being tested

Variant Observation occurring frequently in a particular strain of animal

13.2.3 Adverse Effects

The following definitions should be referred to when considering whether an observed effect of treatment is adverse or not:

1. Treatment-related trend in incidence of specific or related malformations
2. Treatment-related increase in alterations, the cumulative effect of which is considered to be detrimental to the well-being of the fetus
3. Treatment-related increase in alterations, which are related in nature or derivation to treatment-related malformations evident on the study
4. Treatment-related marked change in the incidence of a group of alterations, in which although their form is normal for a previous or future stage of development, that is, their occurrence suggests precocious or delayed development, their presence in a marked degree suggests some permanent change in the rate of development of the fetus and could be detrimental to its future development
5. Marked treatment-related increase in the occurrence of a specific alteration, in which the form is not predictive of the normal chronological order of development (e.g., bent scapula)

13.2.4 Pre- and Postnatal Development

The purpose of this component is to detect effects of treatment from implantation through lactation on the pregnant and lactating female and on the development of the conceptus and

offspring through sexual maturity. The study design is similar to the previous segment III study design except that dosing begins on day 6 of gestation instead of day 15. A group size of 16–24 litters would generally be considered acceptable (with 25 mated females being recommended).

Minimal in-life observations for parental (F_0 generation) females include:

1. Clinical signs and mortality daily
2. Body weight twice weekly
3. Food consumption weekly
4. Valuable target effects seen in previous toxicity studies
5. Length of gestation
6. Parturition

Parental females are sacrificed after weaning of the F_1 generation. The age of sacrifice of the F_1 generation animals is not specified in the ICH guideline and varies among laboratories. Typically, they are sacrificed intermittently with some laboratories reducing litter size on postnatal day 0, 3, or 4, on postnatal day 21 or at weaning, at male/female cohabitation to produce an F_2 generation, and at the terminal sacrifice, after production of the F_2 generation. Terminal examinations for maternal animals and offspring include:

1. Necropsy of all parental and F_1 adults
2. Preservation of organs with gross changes and sufficient control organs for comparison
3. Count of implantations

Additional observations of the F_1 generation include:

1. Abnormalities
2. Live and dead offspring at birth
3. Body weight at birth
4. Pre- and postnatal survival, growth, maturation, and fertility
5. Physical development including vaginal opening and preputial separation
6. Sensory function, reflexes, motor activity, learning, and memory

13.2.5 Single-Study and Two-Study Designs for Rodents

Except for the embryo–fetal development component in rabbits, the components described previously can be combined into fewer, larger studies instead of conducting each component separately. Acceptable alternatives include the “single-study design” and “two-study design.” The choice may be made based on when study results are

needed (how soon are females to be incorporated in clinical studies) and compound availability.

In the “single-study design,” all of the aforementioned components are combined into one study. The dosing period, extending from before mating to lactation, is a combination of that for the fertility study together with that for the pre- and postnatal development study. Subgroups of animals are terminated at the end of gestation for fetal examination.

There are a variety of possible “two-study designs.” One is to conduct the single study described in the preceding text except that, instead of having subgroups for fetal examination, a separate embryo–fetal development study in rodents is conducted. Another two-study design consists of combining the embryo–fetal development study with the pre- and postnatal development study such that the two studies to be conducted would be (i) the fertility study and (ii) the pre- and postnatal development study with subgroups terminated at the end of gestation for fetal examination. A third possible two-study design is to combine the fertility study with the embryo–fetal development study. In the first study, treatment would extend through the end of organogenesis, and then, at termination at the end of gestation, there would be a complete fetal examination. The second study would be the pre- and postnatal development study.

For all the options described earlier, effects on male and female fertility can be evaluated separately by conducting separate studies in which only one sex is treated. The treatment periods are the same, but the treated animals are cohabited with untreated animals of the opposite sex. In the male fertility study, the untreated females are terminated after the middle of gestation, and terminal observations include embryo survival and possibly external examination of fetuses (if terminated at the end of gestation) (Tanimura, 1990). The advantage of conducting separate male and female studies is that, if there are effects, it is clear which sex was affected by treatment. Often when effects are seen in a combined male and female study, additional work is required to resolve which sex was affected. Either a second cohabitation of the treated males with untreated females is added, or studies with only one sex treated must then be conducted.

With the possible exception of combining the female fertility component with the embryo–fetal development component, the combined-study approach is used often. The female fertility and embryo–fetal development components are needed to support clinical trials in women of childbearing potential in most countries and thus will be conducted early in the development of a drug. However, since the pre- and postnatal development component is not routinely required for clinical studies of women of childbearing potential and represents a large commitment of resources, it will not generally be conducted until late in the drug development process.

13.2.6 Preliminary Studies

According to the ICH guideline, “some minimal toxicity is expected to be induced in the high dose dams” in the reproductive and developmental toxicity studies. In some cases, particularly for the fertility and early embryonic development study, available information from general toxicity studies in the selected rodent species may be sufficient to allow the selection of dosage levels for a reproductive toxicity study with the goal of achieving minimal toxicity in high-dose dams. However, pregnant females sometimes respond differently to toxins than nonpregnant females, the duration of dosing for reproductive toxicity studies is different than for general toxicity studies, and toxicity may not have been achieved in the subacute toxicity studies. Thus, it is often necessary to conduct a range-finding study in pregnant rodents prior to the embryo–fetal development study. A range-finding study in rabbits is almost always required since only rarely are results available from other toxicity studies.

The range-finding study in pregnant animals (rodents or rabbits) is similar to the embryo–fetal development study discussed previously except that there may be more dosage groups, group size is smaller (6–10 inseminated or mated females per group), and there is generally no need to examine fetuses for visceral or skeletal abnormalities. Evaluating litters from range-finding studies for resorption, fetal weight, and external abnormalities is valuable for providing an early indication of marked developmental toxicity. This is particularly important if conceptus survival at a particular dosage level would be inadequate to evaluate effects on development in the subsequent embryo–fetal development study. Once it has been determined during a range-finding study that a particular dosage level causes toxicity exceeding the minimal toxicity desired for the embryo–fetal development study, it is best to terminate that dosage group since continued treatment and evaluation unnecessarily expose animals to toxicity, any subsequent data collected are not useful for risk assessment (since it is known that excessive maternal toxicity itself causes developmental toxicity), and investment of resources is therefore unwarranted.

13.2.7 Toxicokinetics

The ICH guidelines do not require that toxicokinetic studies be conducted except that “at the time of study evaluation further information on kinetics in pregnant or lactating animals may be required according to the results obtained.” In addition, the guidelines state that “it is preferable to have some information on kinetics before initiating reproduction studies.” In practice however, at least some degree of toxicokinetic sampling and analysis is performed to verify exposure and allow risk assessment.

The major toxicokinetic issue for reproductive toxicity studies is whether systemic exposure in the selected species

and route is adequate relative to the systemic exposure with the clinical regimen. Often, this information is available for the selected rodent species from studies conducted independently from the reproductive toxicity studies. For rabbits, though, there is rarely toxicokinetic information available from other studies. Accordingly, it is advisable to conduct at least a crude evaluation of systemic exposure in the rabbit. It is best if these data are available prior to the embryo–fetal development study so that, if the rabbit is found to have inadequate systemic exposure, an alternative species may be selected before the investment of resources in a large rabbit study. The collection of blood samples for toxicokinetic evaluations may be incorporated into the range-finding study in pregnant rabbits. However, rabbits stressed by multiple bleedings should not be retained for evaluation of developmental toxicity, and satellite groups of toxicokinetic animals for bleeding only may be needed.

It would be ideal to have data from five to eight time points following the first and last doses to examine accumulation and other changes in kinetic parameters during pregnancy and, since physiology changes rapidly during gestation, to have data periodically during gestation as well. However, from a practical point of view, the question being asked (what is the approximate systemic exposure?) does not justify a comprehensive kinetic evaluation. When circumstances dictate that a toxicokinetic evaluation be performed, determining maternal plasma levels at a few postdosing intervals during a single 24 h period of gestation, preferably during the period when serious adverse effects are most likely to be induced (days 9 through 12 of gestation), will generally provide adequate information.

Only in special circumstances will the determination of embryo levels of drug add meaningfully to the assessment of human risk from a drug. In such studies, even if it is found that the embryo is not exposed, the lack of exposure of the embryo would not necessarily indicate an invalid study or increased human risk since there may also be no exposure in human embryos. When embryo level studies are conducted, the selection of day(s) of gestation to harvest embryos is severely restricted by the sensitivity of the assay. Often, the earliest day that allows the collection of sufficient tissue for assay is gestational day 10 or 11.

13.2.8 Timing of Studies

The definition of which studies need to be performed in advance of clinical trials has not been harmonized yet by the ICH process and is currently determined by the regulatory agencies of individual countries and institutional review boards (IRBs) (or ethics committees (ECs)). Embryo–fetal development studies in two species are almost universal prerequisites prior to clinical studies in women of child-bearing potential. Some regulatory agencies also request that a fertility study in female rodents be conducted before

clinical trials in women of childbearing potential. A fertility study in male rodents is required before clinical trials in men in Japan. Some pharmaceutical companies have internal guidelines that specify compliance with all the guidelines listed earlier, regardless of the location of the clinical trials.

The most conspicuous exception to the policy described in the preceding text is the position of the US FDA (1993). The FDA withdrew the restriction on the participation of women of childbearing potential in early clinical trials, citing "(i) exclusion of women from early trials is not medically necessary because the risk of fetal exposure can be minimized by patient behavior and laboratory testing, and (ii) initial determinations about whether that risk is adequately addressed are properly left to patients, physicians, local IRBs and sponsors with appropriate review and guidance by FDA, as are all other aspects of the safety of proposed investigations." The policy of excluding women has been replaced by one that specifies that "the patients included in clinical trials should, in general, reflect the population that will receive the drug when it is marketed." In fact, inclusion of women at the earliest possible stages is frequently mandated. This led to the current situation where requirements as to when developmental and reproductive testing must be performed are very different between ICH regions, with US FDA requirements generally not mandating any such studies until after phase I unless a potential risk/concern is indicated by drug mechanism or finding in other studies (most commonly the initial repeat-dose toxicity studies).

To comply with FDA policy, at least for the conduct of clinical trials in the United States, pharmaceutical companies have a few choices. They can conduct the standard battery of reproductive studies prior to enrolling women of childbearing potential in early clinical trials. The possible negative impact would be a delay in the initiation of clinical trials. Alternatively, pharmaceutical companies can enroll women of childbearing potential in early clinical trials without having conducted any reproductive toxicity studies and accept the additional risk resulting from exposure to untested drugs during inadvertent or undetected pregnancy. In either case, the incidence of pregnancy during clinical trials can be decreased by pregnancy testing and/or assurances of contraception.

13.3 METHODOLOGICAL ISSUES

13.3.1 Control of Bias

An important element to consider when designing developmental and reproductive toxicity studies is the control of bias. For example, animals should be assigned to groups randomly and preferably blocked by body weight. This can be accomplished by first ranking the animals in order of

body weight and then, starting with the lightest or heaviest, assigning by rank to groups based on a list of sets of random permutations of numbers (e.g., 1, 2, 3, and 4 if there are four groups, where 1 represents the control group, 2 represents the low-dose group, etc.). Housing of treatment groups should also be unbiased. This can be done by "Latin square" design where each block of four cages (if there are four groups) includes an animal from each group. It is often an acceptable compromise to have animals from different groups in alternating vertical columns with all the animals in a column from the same group. This provides equal vertical balancing for all groups. Historically, it has proven unwise to have groups segregated on separate racks.

The order of sacrifice on the day of cesarean sectioning should be balanced by group (again using random permutations) since fetuses continue to grow during the day and an unbalanced time of sacrifice would bias fetal weights, particularly for rodents. Alternatively, all animals can be killed at about the same time in the morning and the fetuses stored for examination later the same day. Fetal examinations should be conducted blind, that is, without knowledge of treatment group.

13.3.2 Diet

It is known that rodents require a diet relatively rich in protein and fats for successful reproduction (Zeman, 1967; Chow and Rider, 1973; Turner, 1973; Mulay et al., 1982). Consequently, rodents are fed high-protein, high-fat diets *ad lib* for reproductive toxicity studies and also generally as a maintenance diet for all toxicity studies. Female rats fed in this manner begin to show decreases in fertility, litter size, and the incidence of normal estrus cycling at the age of 6 months (Matt et al., 1986, 1987). The disadvantage of this feeding practice is that the animals more quickly acquire age-related diseases and sexual dysfunction and die sooner than if they are fed a restricted amount of calories (for review, see Weindrich and Walford, 1988). In relatively short-term studies (such as standard ICH studies), this rapid aging does not present a problem. However, for male breeding colonies or multigeneration studies with multiple litters per generation, it could be advantageous to restrict caloric intake, at least when the animals are not being bred. Restriction of food intake to achieve a 30% decrease in body weight gain compared to *ad lib*-fed controls has no adverse effect on male rat reproduction (Chapin et al., 1991), although it does affect reproduction in mice (Gulati et al., 1991) and female rats (Chapin et al., 1991).

Dietary restriction is even more important for rabbits. Rabbits fed *ad lib* fare very poorly. Some laboratories restrict New Zealand white rabbits to 150–180 g day⁻¹ of a high-fiber (at least 13.5%) diet. However, even this regimen results in some rabbits going off feed late in gestation. It has been observed that by restricting New Zealand white rabbits to only 125 g of food per day, nearly all control animals

retain appetite throughout gestation and fewer of these animals abort (Clark et al., 1991). More uniform food consumption late in gestation is associated with greater uniformity in maternal body weight change and fetal weight. This decreased variability makes these measures more sensitive indicators of maternal and developmental toxicity. Thus, 125 g is the preferred daily ration for New Zealand white rabbits.

13.3.3 Clinical Pathology

Regulatory guidelines require that there be maternal toxicity at the highest dosage level in embryo–fetal developmental toxicity studies. It is also important to avoid excessive toxicity in these studies since it is known that marked maternal toxicity can cause secondary developmental toxicity (see discussion in Section 13.5.3). This secondary developmental toxicity is irrelevant to the assessment of the developmental hazard of the test agent and thus simply confounds the interpretation of the data.

The traditional indicators of maternal toxicity in range-finding studies in pregnant animals (mortality, body weight, food consumption, and clinical signs) do not always provide a sensitive measure of toxicity. This insensitivity is a particular problem for rabbit studies since typically no other toxicity studies are conducted in rabbits and body weight change in rabbits is very variable (typically –100 to +400 g during gestation), making it a particularly insensitive indicator of toxicity.

Thus, it is desirable to improve the assessment of toxicity in range-finding studies in pregnant animals. Complete histopathologic examination is not practical. However, it is often feasible to perform hematologic and serum biochemical analyses that can significantly increase the changes of detecting significant toxicity and provide important information for selecting an appropriate highest dosage level for the embryo–fetal developmental toxicity study.

Based on more than 20 years of experience, body weight effects most often provided the basis for the selection of dosage levels in the segment II study. However, there have been cases where clinical pathology was or would have been useful to justify dosage selection. For example, the nonsteroidal anti-inflammatory drug diflunisal caused a decrease in erythrocyte count from 6.0 (million mm^{-3}) to 2.9 at a dosage level ($40 \text{ mg kg}^{-1} \text{ day}^{-1}$) that caused only a 1% decrease in body weight in pregnant rabbits. The severe hemolytic anemia caused by this excessively high dosage level in turn caused secondary axial skeletal malformations in the fetuses (Clark et al., 1984). Also, the angiotensin-converting enzyme (ACE) inhibitor enalapril caused an increase in serum urea nitrogen from 16 to 46 mg dL^{-1} (highest value = 117) at a dosage level ($10 \text{ mg kg}^{-1} \text{ day}^{-1}$) that had no apparent effect on body weight but caused a significant ($p < 0.05$) increase in resorptions (Minsker et al., 1990). Serum urea nitrogen

concentration was used to select dosage levels for a subsequent ACE inhibitor, lisinopril. Likewise, the routine use of clinical pathology in range-finding studies has previously been proposed (Wise et al., 1988). The animals can be bled on the day after the last dose or sooner to detect transient effects or to allow an evaluation of the data prior to cesarean section.

13.3.4 Gravid Uterine Weights

Effects of treatment on maternal body weight gain are commonly evaluated as indicators of maternal toxicity. However, maternal body weight gain is influenced by fetal parameters such as live fetuses per litter and fetal body weight. Thus, effects indicative of developmental toxicity could contribute to decreased maternal body weight gain and confound the interpretation of maternal toxicity. In addition, other maternal but pregnancy-related parameters, such as volume of intrauterine fluid, could be affected by treatment and contribute to effects on overall body weight gain.

In an attempt to correct this complication, some laboratories weigh the gravid uterus at cesarean section and then subtract the weight of the gravid uterus from the body weight gain to obtain an adjusted weight gain that is more purely maternal. This adjustment is imprecise, but not inappropriate for rats for which gravid uterine weight is correlated with and generally substantially less than maternal body weight change during gestation (e.g., see Figure 13.2 for which the correlation coefficient r was 0.63 and $p < 0.001$). However, the subtraction of gravid uterine weight from maternal weight gain is an overadjustment for rabbits. The maternal body weight gain of rabbits during gestation is generally less than the weight of the gravid uterus (see Figure 13.3). Moreover, gravid uterine weight is correlated with maternal body weight change in some but not all studies. For example, in the 53 untreated rabbits from the study shown in Figure 13.3, $r = 0.54$ and $p < 0.001$. However, in a study of 32 rabbits treated with a methylcellulose vehicle, $r = 0.21$ and $p = 0.25$. Thus, subtracting the gravid uterine weight from the maternal weight gain is not always appropriate. A preferred method for adjusting maternal body weight gain for possible developmental effects is to test and, if appropriate, use gravid uterine weight as a covariate (J. Antonello, personal communication, 1990). This method can be used for both rats and rabbits and for body weight change intervals in addition to those ending at study termination.

Alternatively, to avoid weighing the uterus (or if the analysis is being performed retrospectively and uterine weights are unavailable) or if a more purely fetal adjustment is desired, one can use the sum of the live fetal weights within the litter (total live fetal weight) as the covariate instead of gravid uterine weight. As expected, total live fetal weight is very highly correlated with gravid uterine weight in control

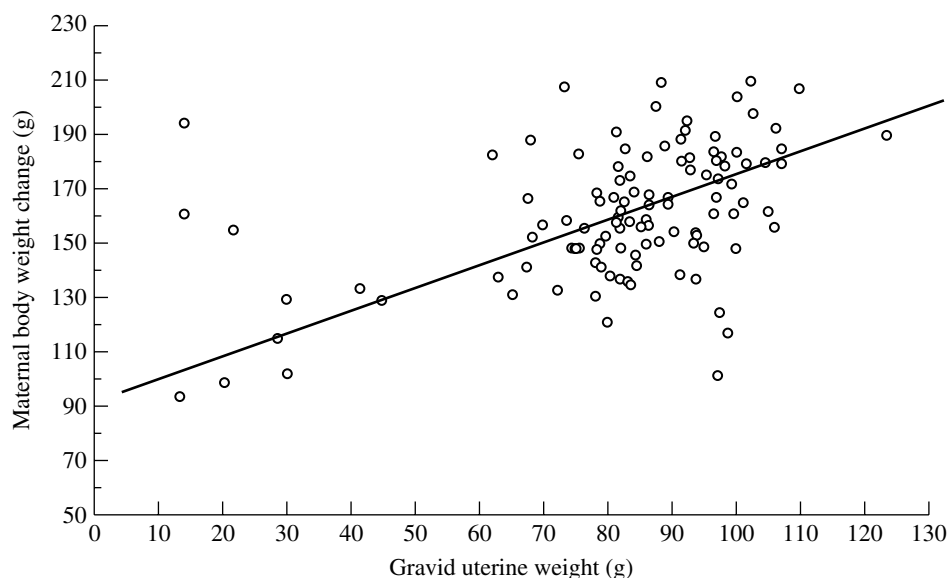


FIGURE 13.2 Relationship of maternal body weight and gravid uterine weight in rats. Relationship between gravid uterine weight and maternal body weight change in control rats between days 0 and 20 of gestation. One hundred and twenty pregnant Sprague–Dawley (CrI:CD(SD)BR) rats were dosed orally with 0.5% aqueous methylcellulose on days 6–17 of gestation and cesarean sectioned on day 20 of gestation. The gravid uterus from each animal was removed and weighed.

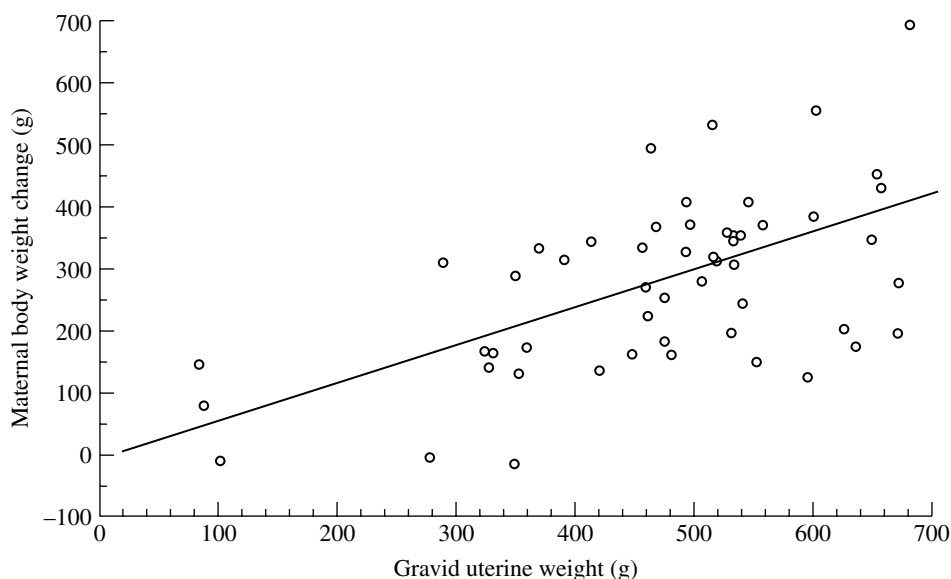


FIGURE 13.3 Body weight changes versus gravid uterine weight changes in rabbits. Relationship between gravid uterine weight and maternal body weight change in untreated rabbits between days 0 and 28 of gestation. Fifty-three pregnant New Zealand white rabbits that had not been treated with control article or test agent were cesarean sectioned on day 28 of gestation. The gravid uterus from each animal was removed and weighed.

animals ($r=0.99$ in control rats and 0.95 in control rabbits; J. Antonello, personal communication, 1990). Thus, in general, using either gravid uterine weight or total live fetal weight as the covariate will yield similar results. However, if treatment were to have an effect on gravid uterine weight that was not reflected in total live fetal weight (e.g., if the

volume of amniotic, extracoelomic, or intrauterine fluid was affected), then total live fetal weight may not be highly correlated with gravid uterine weight and, hence, not interchangeable as a covariate. In that case, only weighing the gravid uterus would allow the detection of these effects not revealed by total live fetal weight.

13.3.5 Implant Counts and Determination of Pregnancy

Two observations suggest that the remnants of embryos that die soon after implantation are not apparent at gross examination of the uterus near term: First, embryos that were observed to be resorbing at laparotomy early in gestation left no readily visible trace near term (Staples, 1971). Second, occult implantation sites can be revealed near term by staining the uterus with ammonium sulfide or sodium hydroxide (Salewski, 1964; Yamada et al., 1988). It is not known if the uterine staining techniques reveal all implantation sites. It is clear, though, that when uterine staining techniques are not used, very early resorptions may not be included in what is termed the "resorption rate" but instead may contribute to the apparent "preimplantation loss" or, if no implantation sites were detected, the rate of "nonpregnant" animals.

In normal circumstances, probably very few implantation sites are not detected without staining. However, cases have occurred in which probably treatment effects were detected only as a result of uterine staining. For example, in one rabbit study with drug treatment starting on day 6 of the gestation, a drug-treated group had four litters that had implantation sites that were seen only after staining with ammonium sulfide, indicating very early drug-induced resorption. For critical studies in rabbits designed to determine early effects on resorption and abortion rates, it would be advantageous to measure plasma levels of progesterone on day 6 of gestation since low levels indicate nonpregnant animals (Adams et al., 1989, 1990).

13.3.6 Fetal Examinations

Many fetal anomalies, such as cleft palate, exencephaly, ectrodactyly, and missing vertebra, are discrete and distinct and therefore easy to recognize objectively. Some anatomical structures, though, occur along a continuous gradation of size and shape and are only considered anomalous if the deviation from the typical exceeds a somewhat arbitrarily selected threshold. These anomalies are observed in all examination types and include micrognathia, reduced gallbladder, enlarged heart, distended ureter, wavy rib, and incomplete ossification at many sites. In many cases, it cannot be said with certainty whether a specific degree of variation from normal would have resulted in an adverse consequence to the animal and should therefore be considered abnormal. In the absence of certainty about outcome, the best approach is to uniformly apply a single criterion within a study (and preferably among studies) so that all treatment groups are examined consistently. The subjectivity (and hence fetus-to-fetus variability) of the examination can be minimized by having the criteria be as clear and objective as possible. For example, when examining for incompletely ossified thoracic centra or supraoccipitals, it can be required that the ossification pattern be absent (unossified), unilateral,

or bipartite (which are objective observations) before recording as an observation. Subjective criteria such as being dumbbell or butterfly shaped would not be applied.

13.3.6.1 Examination of External Genitalia One aspect of external anatomy that is largely overlooked in the examination of offspring exposed *in utero* to test agents is the external genitalia, even though major malformations can occur in those structures. For example, hypospadias is a malformation in the male in which the urethra opens on the underside of the penis or in the perineum. Hypospadias can occur in the male rat following *in utero* exposure to antiandrogens (e.g., Neumann et al., 1970), testosterone synthesis inhibitors (e.g., Bloch et al., 1971), or finasteride, a 5 α -reductase inhibitor (Clark et al., 1990b). However, it is impractical to detect hypospadias in fetuses or young pups. Although the genital tubercle of the normal male rat fetus is grossly distinguishable from that of the normal female as early as day 21 of gestation (the female has a groove on the ventral side), the difference is very subtle, and partial feminization of the male genital tubercle would be very difficult to ascertain. Routine histological examination is obviously too labor intensive to be considered. Hypospadias can readily be determined, though, by expressing and examining the penis of the adult. Thus, it is recommended that adult F₁ males be examined for hypospadias. If the timing of the separation of the balanopreputial membrane is being included in the pre- and postnatal development study as a developmental sign (see Korenbrot et al., 1977), the examination of the penis for hypospadias can be conducted at the same time.

The critical period for the induction of hypospadias by finasteride in rats is days 16–17 of gestation (Clark et al., 1990a). It is unlikely that other agents would have a much earlier critical period since testosterone synthesis, which is required for the development of the penile urethra, begins in the rat on day 15 of gestation (Habert and Picon, 1984). Thus, if treatment in the embryo–fetal development study terminates on day 15 of gestation (as is done in some laboratories), it is doubtful that hypospadias could be induced. However, hypospadias could be induced in the pre- and postnatal development study. Since the formation of the penile urethra in the rat is not completed until day 21 of gestation (Anderson and Clark, 1990), it could be argued that "major organogenesis" continues until that time.

One parameter that is readily and commonly measured as an indicator of effects on differentiation of the external genitalia in rodent fetuses is the sexually dimorphic distance between the anus and the genital tubercle (anogenital distance). However, it should not be assumed that anogenital distance is synonymous with hypospadias, since effects on anogenital distance are not necessarily predictive of hypospadias. Finasteride caused both hypospadias and decreased anogenital distance in male offspring but with very different dose–response relationships and only a slight tendency for animals

with hypospadias to have a shorter anogenital distance (Clark et al., 1990b). Also, the effects on anogenital distance were largely reversible, whereas hypospadias was not. Another agent, triamcinolone acetonide, caused dramatic (reversible) decreases in anogenital distance in male rat fetuses on day 20 of gestation but did not affect the development of the genital tubercle and did not cause hypospadias (Wise et al., 1990b). Thus, decreased anogenital distance per se does not necessarily indicate a serious congenital anomaly.

When evaluating effects of treatment on fetal anogenital distance, it is obviously important to correct for effects on fetal weight. One approach is to calculate "relative" anogenital distance, the ratio between anogenital distance and another linear measure, for example, biparietal diameter (head width). The cube root of fetal weight simulates a linear measure (Wise et al., 1990b) and can also be used to normalize anogenital distance. Another approach is to compare the anogenital distance in a weight-reduced treatment group to that in a weight-matched control group at a younger age.

13.3.6.2 Visceral Fetal Examinations The examination of the abdominal and thoracic viscera of fetuses is performed either fresh without fixation ("Staples" technique") or after Bouin's fixation by making freehand razor blade sections ("Wilson's technique"; Wilson, 1965). Both techniques have advantages. The fresh examination technique, which may require less training to become thoroughly proficient, provides a more easily interpreted view of heart anomalies. The examination must be performed on the day the dam is terminated, though, so having a large number of litters to examine in one day requires that a large team of workers be committed to the task.

With both techniques, the heads of one-half of the fetuses can be fixed in Bouin's fixative for subsequent freehand sectioning and examination. A common artifact induced by fixation in rabbit fetal heads is retinal folding.

Whether or not the kidneys are sliced transversely to examine the renal pelvis varies among laboratories. Hydronephrosis, delayed papillary development, and distended renal pelvis are most readily detected in this manner. However, it is not necessary to slice the kidneys to detect the urinary retention that can lead to distended renal pelvis and hydronephrosis. This point was demonstrated in a study in which 200 000 IU kg⁻¹ day⁻¹ of vitamin A administered orally on days 8–10 of gestation induced hydronephrosis and/or distended renal pelvis in 29 fetuses (R. Clark, personal communication, 2004). In all of these 29 fetuses (and two others), distended ureter also occurred. Thus, distended ureter may be a more sensitive indicator of urinary retention than distended renal pelvis.

13.3.6.3 Skeletal Fetal Examination There is variability in the development of the fetal skeleton, including the number of vertebrae and ribs, patterns of sternebral ossifica-

tion, alignment of ribs with sternbrae, and alignment of ilia with lumbar and sacral vertebrae. There is also extensive plasticity in the development of the skeleton beyond the fetal stage. For example, it is known that markedly wavy ribs in fetuses can resolve so that the ribs in the adult are normal (Saegusa et al., 1980; Nishimura et al., 1982) and supernumerary ribs can be resorbed (Wickramaratne, 1988). This variability and plasticity complicates the classification of anomalies as true malformations as opposed to variations of normal. There is no unanimity on terminology, but, in general, a variation tends to be an alteration that occurs at relatively high spontaneous incidence (>1%), is often reversible, and has little or no adverse consequence for the animal.

When tabulating and interpreting fetal skeleton data, a distinction is made between alterations in the pattern of development and simple delays in development that are considered to be less serious. A delay in skeletal development is usually apparent as a delay in ossification, as evidenced by an increased incidence of specific, incompletely ossified sites or decreases in counts of ossified bones in specific regions (e.g., sacrocaudal vertebrae). These delays are normally associated with decreases in fetal weight and commonly occur at dosage levels of the test agent that also cause decreased maternal body weight gain.

When determining the criteria for recording skeletal alterations, particularly sites of incomplete ossification, it is legitimate to consider the resulting incidences. For example, including an unossified fifth sternbra in the criteria for recording incomplete sternebral ossification may increase the control incidence to a very high proportion (over 95%) of fetuses affected, which would then reduce the sensitivity for detecting treatment effects. The additional effort expended in recording the extra observations due to fifth sternbra would be wasted. In addition, recording high incidences of incomplete ossification at many sites is not worth the effort involved. The ossification at various sites is highly correlated, so recording at multiple sites is redundant. In some cases, the incidences can be reduced to reasonable levels (1–20% of control fetuses) and the criteria simultaneously be made more objective by requiring that the bone be entirely unossified before recording.

13.3.7 Developmental Signs

The postnatal evaluation of F₁ pups includes the observation of developmental signs in two or more pups per sex per litter. In general, the acquisition of these developmental landmarks, including anatomical changes (e.g., ear pinna detachment, incisor eruption, hair growth, and eye opening) and reflexes (negative geotaxis, surface righting, and free-fall righting), is highly correlated with body weight, but as indicators of developmental toxicity, they are not as sensitive as body weight (Lochry et al., 1984; Lochry, 1987) and thus have minimal value. Possible exceptions to this generality are the

ontogeny of the auditory startle reflex and the markers of sexual maturation (vaginal patency, testes descent, and balanopreputial separation in males).

The examinations for developmental signs should be performed daily starting before and continuing until the criterion is achieved. The separation of the balanopreputial membrane of the penis (occurring at postnatal week 6–7; Korenbrot et al., 1977) is becoming the preferred landmark of sexual maturation in males. The timing of the testes descent is more variable and very dependent on the achievement criteria used. Another advantage of determining the time of the balanopreputial separation is that anomalies of the penis may be observed at the same time (as noted in the preceding text).

13.3.8 Behavioral Tests

The trend within reproductive toxicology is to move from simple determinations of developmental landmarks and reflexes to more sophisticated and sensitive behavioral tests. This process was accelerated by the Environmental Protection Agency (EPA) of the United States, which issued guidelines requiring a “developmental neurotoxicity” study of compounds that meet any of several broad criteria (EPA, 1991). The behavioral tests to be performed in this study are extensive and rigidly defined. As laboratories become equipped and trained to meet these guidelines, they are adding such tests to their evaluations of pharmaceuticals. The suggestions for routine testing made in the succeeding text are considered reasonable for pre- and postnatal development studies intended as routine screens. It is suggested that testing be conducted on one or two adults per sex per litter, keeping the range of actual ages as tight as possible.

Measurement of motor activity is commonly performed in the dark in cages or plastic boxes (open field) or residential mazes in which movement is quantitated by infrared detectors or by recording the interruption of light beams as the test subject moves through a horizontal grid of light beams. Possible parameters to evaluate include horizontal activity (light beams interrupted), number of movements, and time spent in the middle of the cage. The test period is selected to be long enough (normally 30–50 min) to allow the activity of the animals to decrease to an approximately constant level (asymptote). Testing of young pups (e.g., 13 days of age) is not recommended as their activity level is fairly constant during the test period and young unweaned pups should not be separated from their mothers for extended periods of time.

Another test paradigm for detecting treatment effects on brain functioning in F_1 offspring measures auditory startle habituation. In this test, the animal is placed in a chamber with a floor that detects movement. The animal is exposed to a sequence of 50–60 auditory stimuli, each at 110–120 dB for 20–50 s and separated by 5–20 s.

The gradual diminution of the animal’s movement response is indicative of normal habituation.

There is not a consensus about the procedures to use to test for effects on learning and memory. The two most commonly used techniques are the water-filled maze, which is preferred for measuring learning, and passive avoidance, which is preferred for measuring memory (see Buelke-Sam et al., 1985). Retention is tested in a repeat test conducted approximately 1 week later.

13.3.9 Detecting Effects on Male Reproduction

Male fertility studies with typical group sizes (15–30 males per group) are very insensitive for detecting effects on male fertility. If the control fertility rate is 80%, even a group size of 30 will only detect (at the 5% significance level) a 38% decrease in fertility 80% of the time and a 50% decrease 95% of the time (J. Antonello, personal communication, 1990). To detect slight effects on male fertility would require enormous group sizes. Mating each male with more than one female provides a more precise estimate of the reproductive capacity of each male, but does not greatly increase statistical power. If multiple matings are to be done, it is recommended that the cohabitations with multiple females be sequential rather than concurrent.

Not only is it difficult to detect effects on male fertility because of group-size considerations, effects on male fertility mediated by decreased sperm production are also difficult to detect because of the normally huge excess of sperm included in a rat ejaculate. Sperm production can be decreased by up to 90% without effect on fertility (either pregnancy rate or litter size) in the rat. This is not the case for men, so the sperm excess in the rat represents a serious flaw in the rat model (see Working, 1988). To address this deficiency and improve the sensitivity of the model, it is advisable to determine the effects of the test agent on testes weights, testicular spermatid counts, and histopathology of the testes (preferably plastic sections) in the male fertility study and/or the 14-week toxicity study. In some cases, these parameters may be more predictive of possible effects on male fertility in humans than the fertility rate in rats.

To improve sensitivity of the process, the FDA recently promulgated guidelines (FDA, 2015a) for assessing testicular toxicity in studies involving male test animals. Table 13.5 presents a summary of findings in such studies to be considered further.

13.4 DEVELOPMENTAL STUDIES IN PRIMATES

With the increasing development and use of highly humanized biological molecules as therapeutics in humans, it has become necessary to develop and validate an increasing range of study types which employ primates to evaluate risks to humans.

TABLE 13.5 General Nonclinical Findings to Consider in Male Fertility Risk Assessment**Nonclinical Findings That May Increase the Level of Concern for Infertility in Men**

Finding is dose dependent
 Similar findings in multiple species
 Finding persists or increases in severity with increasing duration of exposure
 Finding persists after drug withdrawal, especially if withdrawal period is an entire spermatogenic cycle
 Finding occurs in bilateral tissues
 Finding is rare in healthy untreated animals
 Maximum dose without adverse effect occurs at exposures that are clinically relevant
 Reproductive organ weight change (increased or decreased weight) correlates with adverse histology
 Decreased male fertility and impaired mating behavior
 Sperm quality adversely affected (count, motility, or morphology)
 Adverse effects on reproductive tissues and function at multiple stages of life (repeat-dose study in adults, adult fertility assessment, effects in adulthood after exposure during pre/postnatal period, toxicity to the reproductive tissues during development)
 Antiandrogenic signs—reduced body weight, decreased weight and maturation of male sexual organs, clinical signs suggestive of reduced aggressiveness (e.g., lethargic or reduced mating behavior)
 Androgenic signs—masculinization of females (decreased fertility, female sexual organ pathology, or estrus cyclicity), decreased testes size, and impaired spermatogenesis

Confounding Issues

Use of immature animals
 Pharmaceuticals that cause weight loss—in some cases, findings observed only in animals with weight loss may be difficult to ascribe to the drug exposure because weight loss alone may adversely affect male fertility independent of drug exposure. Also, weight loss may be secondary to overt toxicity and may not be clinically relevant
 Pharmaceuticals that impair mating behavior or neuromuscular function
 Inappropriate animal model—pharmaceutical is not active in the species or has different metabolite profile, tissue distribution, or extent of elimination

One of these needs is to be able to evaluate developmental toxicity in a model responsive to primate-specific pharmacodynamics.

Figure 13.4 presents a line chart for the study design for an embryo–fetal developmental toxicity study (segment II in old parlance) to be conducted in nonhuman primates (NHPs); the studies present a number of difficulties and need for compromise, starting with the facts that the animals are limited in supply and have small litters (1 or 2 is common).

13.5 DATA INTERPRETATION**13.5.1 Use of Statistical Analyses**

Statistical analysis is a very useful tool for evaluating the effects of treatment on many developmental and reproductive toxicity parameters. For some parameters, such as maternal body weight changes, fetal weight, and horizontal activity in an open field, the comparison to the concurrent control is the primary consideration, and, assuming adequate group size, the investigator relies heavily on the results of appropriate statistical analyses to interpret differences from control.

For other parameters, though, statistical analysis is just one of several considerations that include historical control data and other relevant information about the test agent and related test agents. For example, statistical analysis of a low incidence of an uncommon fetal malformation will usually not be significant ($p > 0.05$) even if treatment related, due to the low power for detecting such effects with typical group sizes. In such cases, examination of the historical control data becomes paramount. If two fetuses with a particular malformation occur in separate litters only in a high-dose group, the finding is of much more concern if it is a very rare malformation than if recent historical control groups have had a few fetuses with that malformation.

Other known effects of the test agent or related agents also sometimes contribute to data interpretation. For example, a low incidence of a malformation may be considered treatment related if it is at the low end of a typical dose–response curve or if it is in a high-dose group and that malformation is an expected effect of the test agent. In general, though, a single occurrence of a fetal malformation in a treatment group (with none in control) is not cause for alarm, since this occurs in almost every study (together with occurrences of some malformations only in the control group).

Statistical methods exist to appropriately analyze most developmental and reproductive toxicity parameters. Exceptions to this are the “ r/m ” litter parameters in which, for each litter, there is a number affected divided by the number in the litter. These parameters include preimplantation loss (r =corpora lutea–implants, m =corpora lutea), resorption rate (r =resorptions, m =implants), and the family of alteration rates (r =affected fetuses, m =fetuses). There are two factors complicating the statistical analysis of these data that have heretofore been inadequately handled (Clark et al., 1989). One is that almost all of these parameters have a strong dependence on m . For example, both preimplantation loss (Figure 13.5) and resorption rate (Figure 13.6) are normally higher at both the low and high extremes of M . In contrast, supernumerary rib tends to occur at higher incidences in average-size litters (Figure 13.7). The second factor that complicates the statistical analysis of r/m data is that affected implants tend

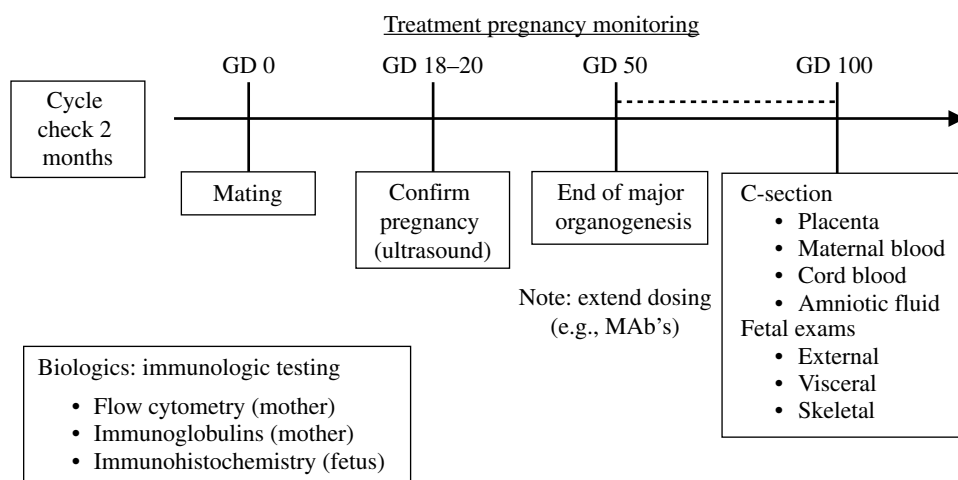


FIGURE 13.4 Embryo–fetal development (EFD) study in nonhuman primates.

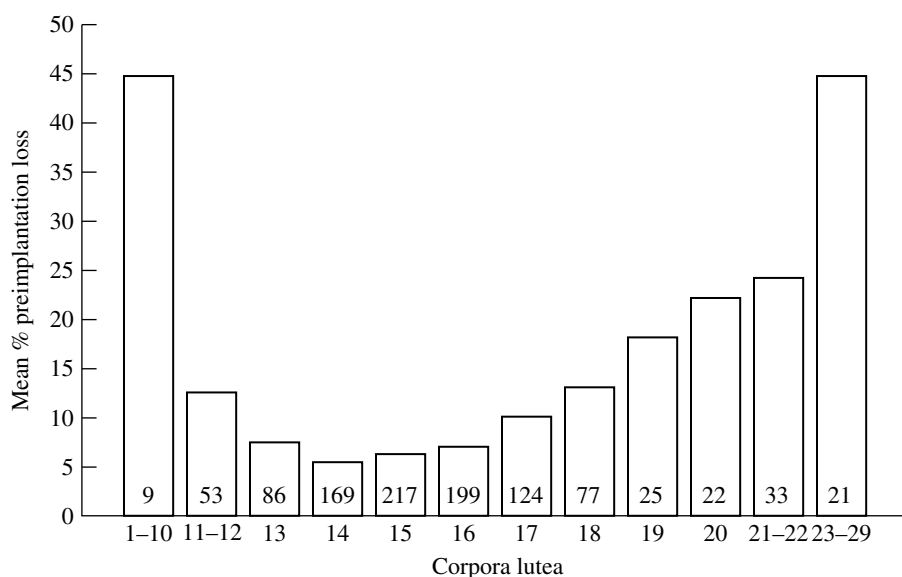


FIGURE 13.5 Percentage preimplantation losses versus corpora lutea in rats. Effect of litter size on mean percentage preimplantation loss in 1035 control rat litters. Between 1970 and 1988, 1035 control rats were cesarean sectioned on day 20 of gestation, and the numbers of resorptions and implants were counted. Numbers within the bars indicate the number of litters.

to occur in clusters within litters (“litter effects”); that is, the intralitter correlation is greater than the interlitter correlation. For example, the total number of litters affected with anasarca, missing vertebra, and supernumerary rib is much less than would be expected by change based on the number of affected fetuses (Table 13.6 and Figure 13.8).

These problems have been resolved for analysis of resorption rate (and preimplantation loss) in Sprague–Dawley rats using a three-step process (Soper and Clark, 1990). First, based on an analysis of data from 1379 control rat litters examined since 1978, a likelihood score was derived for each (r , m) couplet based on the incidence of that couplet

given that value of m . These scores were approximately equal to r . Second, an analysis of 136 litters from groups with slight effects on resorption rate revealed that, at low-effect doses of embryocidal test agents, the increases in resorptions tended to occur as an increased number of resorptions within affected litters rather than as an increased proportion of affected litters. To maximize the difference in scores between control and affected litters, the scores for control-like litters ($r=1, 2$, or 3) were downgraded from r (1, 2, and 3) to 0.4, 1, and 2.4, respectively. Third, to arrive at the final score for each litter, the modified r score for each litter was divided by the expected control value for that value

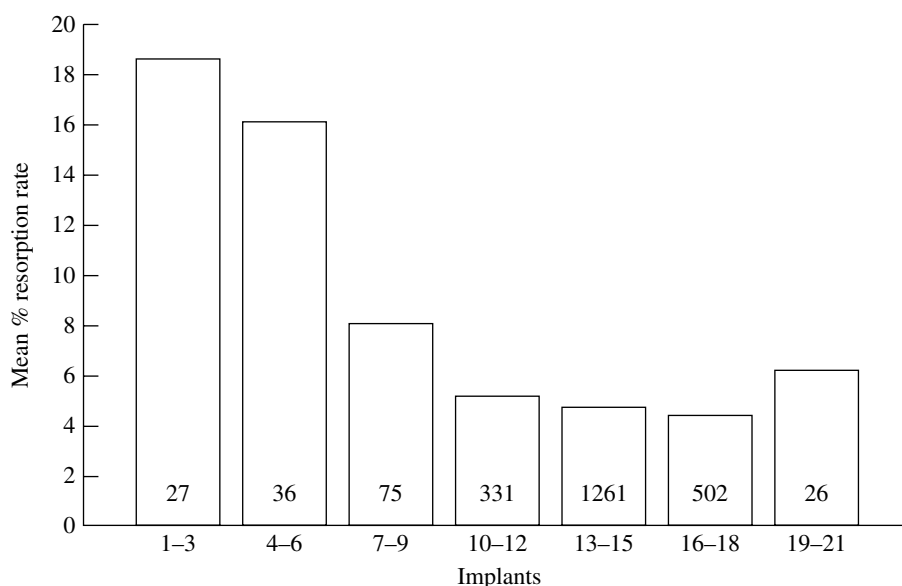


FIGURE 13.6 Resorption rate versus number of implants in rats. Effect of litter size on mean percentage resorption rate in 2258 control rat litters. Between 1970 and 1988, 2258 control rats were cesarean sectioned on day 20 of gestation and the numbers of resorptions and implants were counted. Numbers within the bars indicate the number of litters.

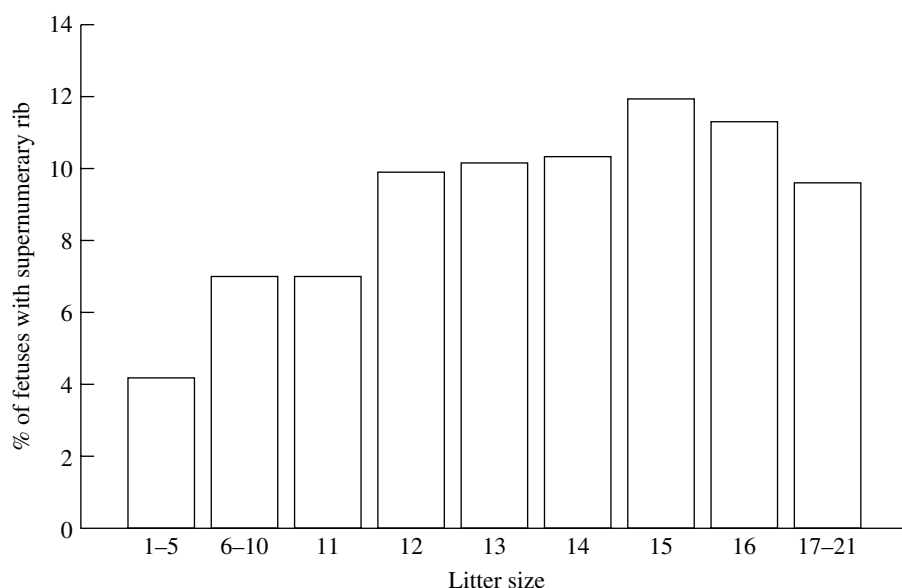


FIGURE 13.7 Effect of litter size (live fetuses per litter) on incidence of supernumerary rib in 1379 control rat litters. Between 1978 and 1988, fetal skeletons from 1379 L of control rats were stained with alizarin red and examined for supernumerary rib.

of m . This last step makes the litter score immune to spontaneous or treatment-related effects on m . The final “robust” scores have more power for detecting effects than various other measures (raw r/m , affected litters/litters, r , $\Sigma r/\Sigma n$, and the likelihood score) and have a lower false-positive rate with fluctuations in m .

Covariance analysis (Snedecor and Cochran, 1989) can be used to reduce variability in a parameter and thereby increase sensitivity. For example, much of the variability in

fetal weight data is due to variable litter size and, for rats, litters being sacrificed at different times during the workday. The variability due to these sources can be reduced by using litter size and time of sacrifice as potential covariates. Similarly, litter size and length of gestation can be used as covariates for neonatal pup weights, and body weight at the beginning of treatment can be used as a covariate for maternal body weight changes during the treatment period of an embryo–fetal development study.

TABLE 13.6 Examples of Litter Effects in Control Litters

| Effect | Affected Fetuses | Affected Litters | Litters Examined |
|-------------------|------------------|------------------|------------------|
| Anasarca | 22 | 12 | 2203 |
| Missing vertebra | 53 | 13 | 1951 |
| Supernumerary rib | 1817 | 621 | 1379 |

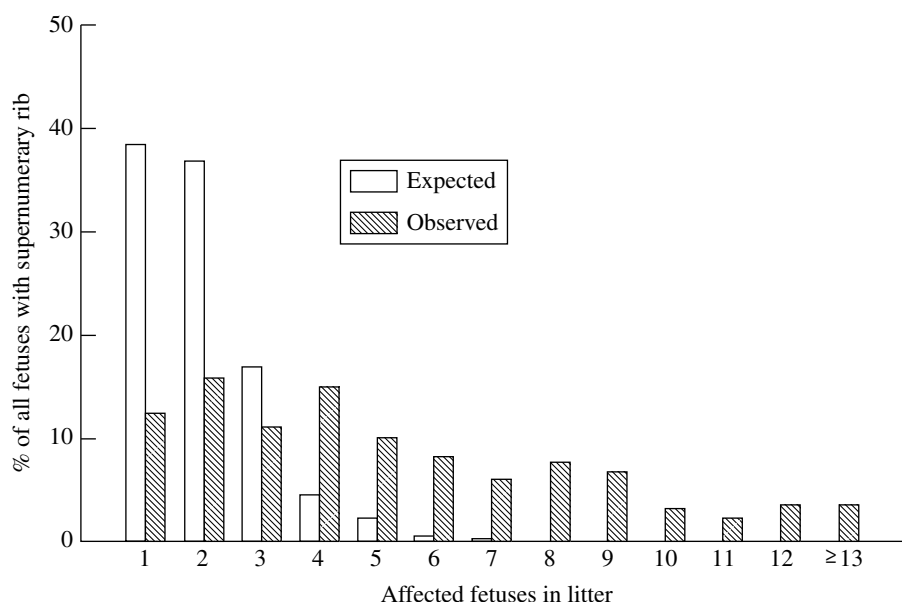


FIGURE 13.8 Litter effect with supernumerary rib in 1379 control litters. Between 1978 and 1988, fetal skeletons from 1379 litters of control rats were stained with alizarin red and examined for supernumerary rib in addition to other anomalies. The calculation of the expected number of fetuses with supernumerary rib in each litter was based on the assumption that each fetus had an equally likely chance of having supernumerary rib independent of the incidence among littermates (K. Soper, personal communication, 1990). *Source:* Soper and Clark (1990).

13.5.2 Potential Hazard Categories of Developmental Toxins

It is generally agreed that an agent that causes developmental toxicity in laboratory animals at dosage levels that cause no maternal toxicity (i.e., “selective” developmental toxins) is potentially more hazardous to humans than agents that cause developmental toxicity only at maternotoxic dosages (“nonselective” developmental toxins; e.g., see Johnson, 1981, 1984; Schwetz, 1981; Fabro et al., 1982; Johnson and Christian, 1984). This position is based on the supposition that pregnant women will avoid being exposed to toxic dosages of pharmaceuticals (which is usually but not always true). Developmental toxins can also be categorized as acting directly on the embryo or indirectly via an effect on the mother. All selective developmental toxins are presumably direct acting. Nonselective developmental toxins can either act directly or indirectly.

Direct-acting developmental toxins may be potentially more hazardous to humans than indirectly acting ones even if the direct developmental toxicity occurred only at maternotoxic dosages in the species of laboratory animals tested. When the developmental toxicity of an agent is secondary to

maternal toxicity in all species tested, the dose–response curves for developmental and maternal toxicity in various species may be invariably linked, and developmental toxicity would never occur at nonmaternotoxic dosages. However, when an agent acts directly on the embryo to cause developmental toxicity, the dose–response curves may not be linked and, although they may be superimposed in the species of laboratory animals tested, they may not be superimposed in other species including humans. Thus, a direct-acting developmental toxin that is nonselective in one species may be selective in another species.

The ranking of potential developmental hazard in terms of selective, direct/nonselective, and indirect is more meaningful than the use of the terminology of specific/non-specific and malformation/variation. However, when it cannot be determined if observed developmental toxicity is a direct or indirect effect, the alternative terminology becomes useful. A nonspecific effect (including in some cases decreased fetal weight, supernumerary rib, cleft palate in mice, and abortion in rabbits) is one that occurs commonly in response to high toxic dosages of a test agent. What makes a nonspecific effect generally less important than a specific

effect is that nonspecific effects commonly occur only at maternally toxic dosages (“coeffective”) and may be secondary to maternal toxicity—that is, indirect. However, when an apparently nonspecific adverse developmental effect is selective (direct), that is, it occurs at nonmaternotoxic dosages, it may nevertheless be indicative of a potential developmental hazard.

In general, an agent that induces a malformation (i.e., a teratogen) is considered to be more of a potential hazard than one that induces only a minor variation. Also, there has traditionally been more of a stigma associated with an agent that induces malformations than one that causes resorptions, even though embryo death is obviously a seriously adverse outcome. The point that makes the distinction among malformations, variations, or resorptions less important is that an agent that perturbs development to cause one effect in one species may cause a different effect in another species. Thus, any developmental toxic effect at nonmaternotoxic dosages should be considered carefully.

13.5.3 Associations between Developmental and Maternal Toxicity

The developmental toxicity of many pharmaceuticals occurs only at maternally toxic dosages (Khera, 1984, 1985; Schardein, 1987, 2000). Also, there are several compounds for which there is evidence that their developmental toxicity is secondary to their maternal toxicity. The decreased uterine blood flow associated with hydroxyurea treatment of pregnant rabbits may account for the embryotoxicity observed (Millicovsky et al., 1981). The teratogenicity of diphenylhydantoin in mice may be secondary to decreased maternal heart rate (Watkinson and Millicovsky, 1983) as supported by the amelioration of the teratogenicity by hyperoxia (Millicovsky and Johnston, 1981) and the dependence on maternal genotype in genetic crosses between sensitive and resistant strains (Johnston et al., 1979; Hansen and Hodes, 1983). The hemolytic anemia caused in pregnant rabbits by diflunisal was severe enough to explain the concomitant axial skeletal malformations (Clark et al., 1984). Acetazolamide-induced fetal malformations in mice are apparently related to maternal hypercapnia (Weaver and Scott, 1984a, b) and hypokalemia (Ellison and Maren, 1972). The increased resorption rate induced in rabbits by the antibiotic norfloxacin depends on exposure of the maternal gastrointestinal tract (Clark et al., 1991).

In addition, various treatments that simulate effects that can result from pharmaceutical treatment have been shown to cause developmental toxicity. Food deprivation can cause embryo–fetal toxicity and teratogenicity in mice (Szabo and Brent, 1975; Hemm et al., 1977) and rats (Ellington, 1980) and fetal death, decreased fetal weight, and abortions in rabbits (Matsuzawa et al., 1981; Clark et al., 1991). Treatments that result in maternal hypoxia,

such as hypobaric exposure (Degenhardt and Kladetzky, 1955) and blood loss (Grote, 1969), have been shown to be teratogenic. Also, the results from testing with numerous agents suggest that supernumerary rib in mice is caused by maternal stress (Kavlock et al., 1985; Beyer and Chernoff, 1986).

Thus, in any case where developmental toxicity occurs at dosage levels with only moderate to severe maternal toxicity, the possibility of the developmental toxicity being secondary to the maternal toxicity can be considered. That is not to say, however, that it can be concluded that the developmental toxicity is secondary any time there is coincident maternal toxicity. On the contrary, it is usually very difficult to establish a causal relationship. Superficially similar types of maternal toxicity do not always cause the same pattern of developmental toxicity (Chernoff et al., 1990). This may be because the developmental toxicity is secondary to maternal toxicity, but, since typical developmental toxicity studies include only a very cursory evaluation of maternal toxicity, the developmental toxicity may be secondary to an aspect of maternal toxicity that is not even being measured.

To demonstrate that a developmental effect is secondary to a particular parameter of maternal toxicity, it is necessary but not sufficient to show that all mothers with developmental toxicity also had maternal toxicity and that the severity of the developmental effect was correlated with the maternal effect. An example of such a correlation is shown in Figure 13.9, in which a drug-induced effect on maternal body weight change in rabbits is correlated ($r=0.45$; $p<0.05$) with a drug-induced decrease in fetal weight. Other examples where this approach has been used to evaluate the relationship between maternal and developmental toxicity include (i) the negative correlation between resorption rate and maternal body weight change in norfloxacin-treated rabbits (Clark et al., 1991), supporting the contention that the developmental toxicity was secondary, and (ii) the lack of correlation between embryotoxicity and maternal body weight change in pregnant mice treated with caffeine and L-phenylisopropyladenosine (Clark et al., 1987), suggesting no causal relationship between developmental and maternal toxicity may be required.

13.5.4 Assessment of Human Risk

Most test agents can be demonstrated to be developmentally toxic if tested under extreme conditions. This fact has been popularized as Karnofsky’s law: “Any drug administered at the proper dosage, at the proper stage of development, to embryos of the proper species ... will be effective in causing disturbances in embryonic development” (Karnofsky, 1965, p. 185). In practice, about 37% of 3301 chemicals tested have been found to be teratogenic according to one tabulation (Schardein, 2000, p. viii; see also Shepard and Lemire, 2010). Contributing to this high rate is the practice of testing

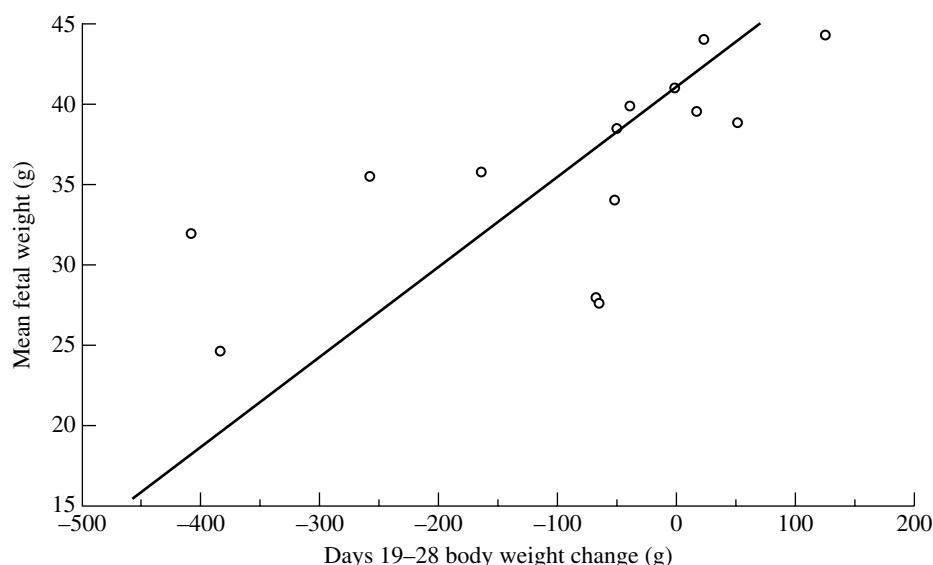


FIGURE 13.9 Maternal weight change versus fetal weight changes. Correlation between drug-induced effects on maternal body weight change and fetal weight in rabbits. The data were collected from the high-dosage group of a developmental toxicity study of a prospective drug candidate. The rabbits were dosed orally with the test agent from days 6 to 18 of gestation. On day 28 of gestation, the rabbits were cesarean sectioned and the live fetuses weighed.

maternotoxic doses (to satisfy regulatory guidelines) that in some cases result in developmental toxicity secondary to maternal toxicity. Despite the high rate of positives in animal tests, very few xenobiotics are known to cause developmental toxicity in humans as commonly used. Thus, simply the induction of developmental toxicity by a test agent in animals does not necessarily indicate that that test agent will be a developmental hazard to human conceptuses under normal exposure conditions.

When a prospective drug under development for use in women of childbearing potential is determined to cause developmental toxicity in laboratory animals, the first question to be considered is whether that agent would cause developmental toxicity in humans at the anticipated therapeutic dosage level. This assessment and the related decision of whether to continue development of the drug candidate are currently based on the following:

1. The ratio between the estimated systemic exposure at the lowest effect level (or highest no-observed-effect level (NOEL)) and the estimated systemic exposure at the anticipated therapeutic dosage level (the “safety factor”)
2. Whether the effect is selective, direct, and/or specific
3. The potential benefit to the patient population (compared to other available therapies)

The most common finding is that minor, nonselective, non-specific developmental toxicity (e.g., decreased fetal weight) is observed at dosages at least 10-fold above the anticipated therapeutic dosage level. In this situation, development of

the agent would normally proceed even if the “safety factor” were only 3–5. This is the case since (i) many pharmaceuticals cause maternal toxicity in laboratory animals at low multiple (e.g., 10) of the clinical exposure, (ii) nonspecific developmental toxicity commonly accompanies maternal toxicity, and (iii) pharmaceuticals fitting this pattern do not usually cause developmental effects as used clinically (which often includes the practice of not prescribing for women known to be pregnant).

In contrast, a drug candidate that selectively causes major malformations at a dosage threefold higher than the clinical dosage would likely not be developed to treat a nonlife-threatening disease. However, it might be developed if the disease to be treated was particularly debilitating, no other effective therapy was available, and it was felt that the exposure of pregnant women could be largely avoided.

Once a new pharmaceutical is approved by the FDA, it is placed in one of five pregnancy categories (A, B, C, D, or X) based on the results of animal developmental toxicity studies and, when available (usually not), information from human usage experience (see Table 13.7 and Frankos, 1985). Note that the categorization does not depend on the safety factor for a developmental effect or whether the effect is major, selective, direct, or specific (although these factors may be considered when determining if a drug is to be approved). Most often, there are positive findings in animals and no experience in pregnant women, and the drug is placed in pregnancy category C, indicating that it is to be used in pregnancy only if the potential benefit justifies the risk to the conceptus. Thus, it is left to the prescribing physician to regulate the exposure of pregnant women to the

TABLE 13.7 Pregnancy Categories^a

| Outcome of Human Studies | Outcome of Animal Studies | | |
|--------------------------|---------------------------|--------|----------------|
| | + | – | Not Available |
| + | X or D | X or D | X or D |
| – | B | A | A or B |
| Not available | C ₁ | B | C ₂ |

^aA, B, C₂: use during pregnancy only if clearly needed.

C₁: use during pregnancy only if the potential benefit justifies the potential risk to the fetus.

D: if used during pregnancy, the patient should be apprised of the potential hazard to the fetus.

X: contraindicated in women who are or may become pregnant.

TABLE 13.8 Pregnancy Categories

A: adequately tested in humans, no risk (0.7% of approved drugs)

B/C/D: increasing levels of concern (92.3% of approved drugs—66% in C)

X: contraindicated for use in pregnancy (7.0% of approved drugs)

Use-in-pregnancy ratings, FDA (1979).

A: animal studies and well-controlled studies in pregnant women failed to demonstrate a risk to the fetus.

B: animal studies have failed to demonstrate risk to fetus; no adequate and well-controlled studies in pregnant women.

C: animal studies showed adverse effect on fetus; no well-controlled human studies.

D: positive evidence of human fetal risk based upon human data, but potential drug benefit outweighs risk.

X: studies show fetal abnormalities in animals and humans; drug is contraindicated in pregnant women.

drug. If animal studies were negative and there is no information on effects on pregnant women, the agent is placed in pregnancy category B, indicating that the agent is to be used in pregnancy only if clearly needed. If developmental toxicity has been established in women (or, in some cases, if only strongly suspected), the agent is placed in category D or X. With category D, women may be prescribed the drug if the benefit outweighs the risk and the patient is informed of the potential hazard to the conceptus. Category X drugs are contraindicated in women who are or may become pregnant. Table 13.8 summarizes these categories and the proportions of drugs in each in 2000 (according to the PDR). It should be noted that the FDA proposed a change to these in May of 2008.

It should be noted that the FDA proposed a change to these in May of 2008. This new FDA Pregnancy Labeling Guidance was proposed for labeling of drugs for reproductive and pregnancy risks. The proposed rule would remove the letter categories from the pregnancy section. The new format would have three sections:

1. *Fetal risk summary* This section would begin with a one-sentence risk conclusion that characterizes the likelihood that the drug increases the risk of four types

of developmental abnormalities: structural anomalies, fetal and infant mortality, impaired physiologic function, and alterations to growth. An example of a risk conclusion based on human data is that “human data do not indicate that Drug X increases the overall risk of structural anomalies.” Many of the risk conclusions in the proposed rule are standardized statements that must be used. This would be followed by a summary of significant effects.

2. *Clinical considerations* This component would address three main topics important when counseling with, and prescribing for, women who are pregnant, lactating, or of childbearing age.
3. *Data* This section would have a more detailed discussion of available data. Human data would appear before animal.

The pregnancy section would also include information about whether there is a pregnancy registry for the drug. Pregnancy exposure registries collect and maintain data on the effects of approved drugs that are prescribed to and used by pregnant women.

The lactation section of prescription drug labeling would provide information on use of the drug while breastfeeding and would use the same format as the pregnancy section.

Using the highest NOEL for determining a safety factor has the following flaws:

1. The determination of a true no-effect level (should one actually exist, which is debatable in some cases) is impossible given the statistical power associated with the group sizes typically used; thus, the reported NOEL is very dependent on the selected group size.
2. The NOEL depends greatly on the selection of dosage levels; unless the selected dosage is just below the threshold for detectable effects, the reported NOEL is an underestimate; thus, tightly spaced dosage levels favor the determination of a higher NOEL.

Accordingly, the FDA has developed a sequential method of evaluating and dealing with reproductive and developmental analysis. This is called wedge analysis and is demonstrated in Figures 13.10, 13.11, and 13.12.

13.6 JUVENILE AND PEDIATRIC TOXICOLOGY

Safety of drug products to developing children has been a concern for more than a decade. Historically, healthcare providers rarely had explicit guidance on either dose levels to administer for drugs or of unusual risks that might be expected in these patients when drugs were given. The usual approach for adjusting dose was to scale approximately based on body weight.

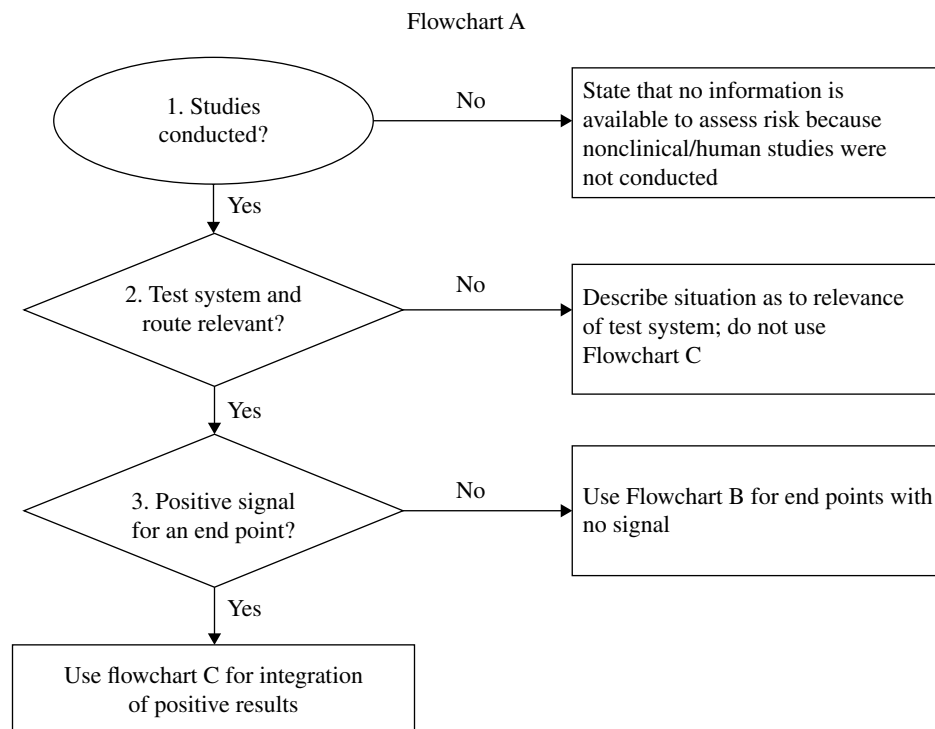


FIGURE 13.10 Integrative evaluation of potential reproductive risk. Flowchart A. Overall decision tree for evaluation on reproduction/developmental toxicity risk. *Source:* FDA (2011). © Food and Drug Administration.

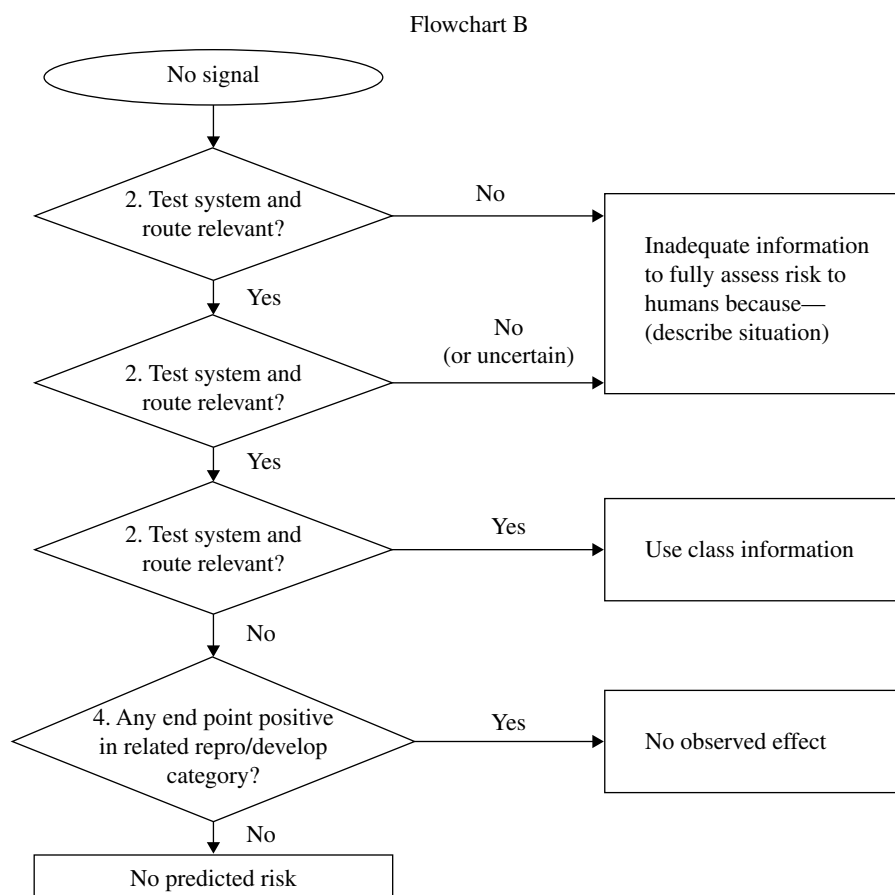


FIGURE 13.11 Flowchart for reproductive risk assessment of a drug. Flowchart B. Decision tree for end points with no signal. *Source:* FDA (2011). © Food and Drug Administration.

Flowchart C

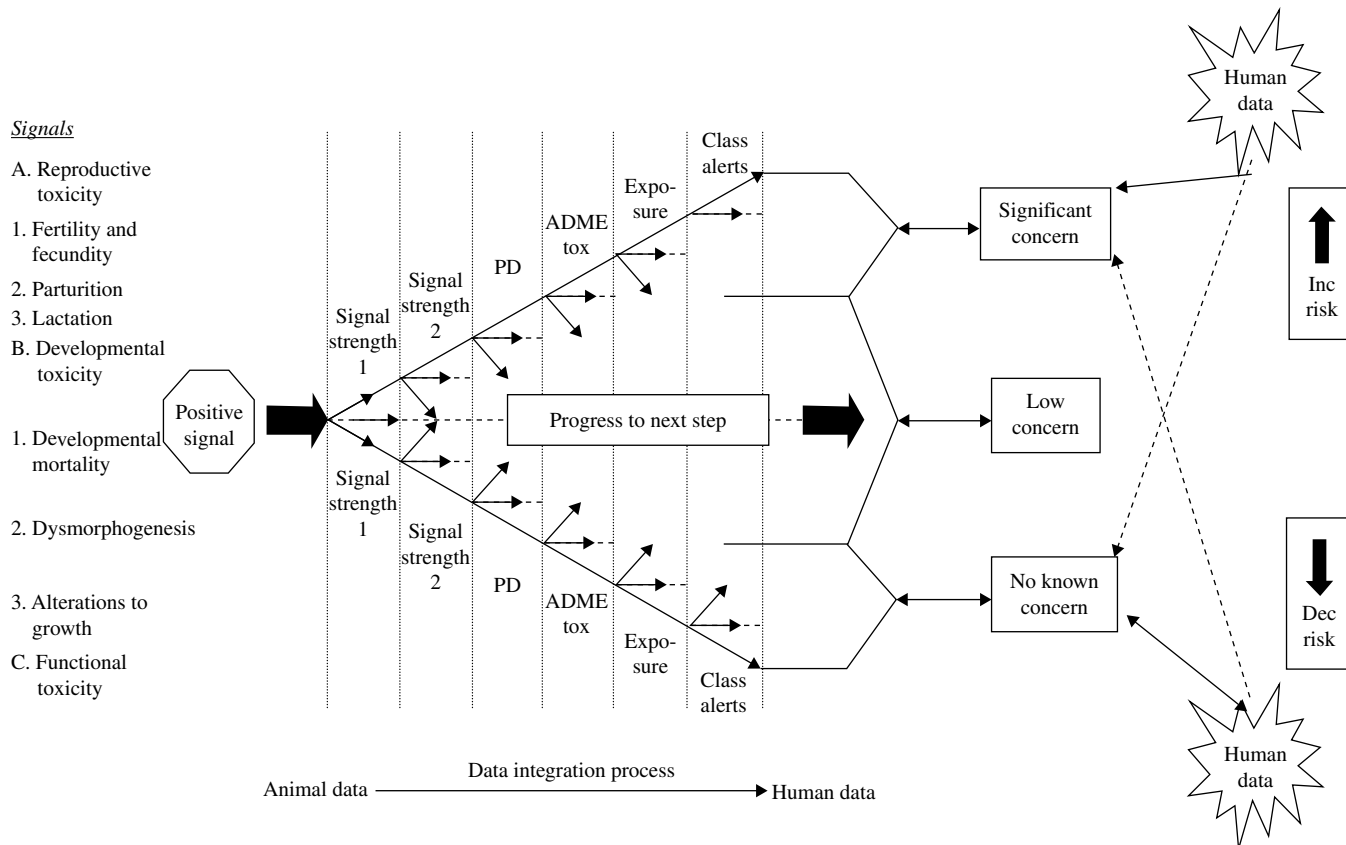


FIGURE 13.12 Scheme for integrative assessment of reproductive/developmental risk of a drug. Flowchart C. Integration of positive reproduction/ancillary study results. *Source:* FDA (2011). © Food and Drug Administration.

TABLE 13.9 Cross-Species Postnatal Maturation

| Organ System | Maturation |
|---------------------|-------------------------------|
| Neurologic | Adolescent/adult |
| Reproductive | Adolescent/puberty |
| Pulmonary (alveoli) | Infant (1–2 years) |
| Renal function | Infant (1 year) |
| Skeletal | Adolescent/adult |
| Immune (IgG, IgA) | Infant/child (5 and 12 years) |

Unfortunately, the differences between an adult and a developing child (usually considered as ranging from neonatal to 16 years of age) may be much more complex (Hoberman and Lewis, 2011):

- Developmental differences and sensitivity differences may result in altered PF/PD in juveniles and require consideration of:
 - Development (structure/function), which is continuous
 - Role of developmental versus chronologic age-matched animal/kids
 - Cross-species postnatal maturation (Table 13.9)

In 2005, 2006, and 2014 both the EMEA and the FDA provided guidance as to study requirements (CHMP, 2005; FDA, 2006; FDA, 2014).

While helpful, these guidances also introduced a number of points of uncertainty (Which species? What ages in test species equate to human developmental situation at different human ages?) and to the need to develop proficiency in areas of technique not previously commonly available (such as how does one go about dosing preweaning animals in toxicology studies—see Zoetis and Walls, 2003).

Tables 13.10, 13.11, 13.12, 13.13, 13.14, 13.15, 13.16, 13.17, and 13.18, originally provided in the FDA 2006 guidance, provide organ system-specific guidances for equivalent developmental ages between the principal test species and humans. Figure 13.13 provides a chronological comparison of timing of key developmental stages for five primary species.

13.7 *IN VITRO* TESTS FOR DEVELOPMENTAL TOXICITY

Many *in vitro* systems have been proposed as tests for developmental toxicity (for review, see Brown and Freeman, 1984; Lin, 1987; *In Vitro* Teratology Task Force, 1987;

TABLE 13.10 Nervous System

| Developmental Event | Postnatal Developmental Period | | | |
|--|--------------------------------|-----------------|-------------|----------------------|
| | Human (Years) | Primate (Weeks) | Dog (Weeks) | Rat (Days) |
| Glutamate receptors ^a (maximal binding) | 1–2 cortex | | | 28 |
| | Decline to adult 2–16 | | | Decline to adult >28 |
| Monoamine system ^b | 2–4 | | | 21–30 |
| | Maximum receptor density | | | Adult levels |
| Ocular dominance ^c | 0–3 | | | 21–35 |
| Cerebellum persistent external germinal layer ^c | 0.6–2 | | | 0–21 |
| Rapid phase of myelination ends ^d | 2 | | | 25–30 |
| Cognitive development Delayed response learning ^e | 1–2 | 9–36 | 12–16 | 10–35 |

^aIkonomidou et al. (1999).^bRice and Barone (2000).^cSidhu et al. (1997) and Kimmel and Buelke-Sam (1994).^dRadde (1985).^eWood et al. (2003).**TABLE 13.11 Reproductive System**

| Developmental Event | Postnatal Developmental Period | | | | |
|----------------------|--------------------------------|-----------------------|------------|--------------|------------|
| | Human (Years) | Rhesus Monkey (Years) | Dog (Days) | Mouse (Days) | Rat (Days) |
| Puberty ^a | 11–12 | 2.5–3 | 180–240 | 35–45 | 40–60 |

^aDeSesso and Harris (1995), Marty et al. (2003), Beckman and Feuston (2003), and Lewis et al. (2002).**TABLE 13.12 Skeletal System**

| Developmental Event | Postnatal Developmental Period | | | | | |
|---|--------------------------------|----------------|-------------|----------------|-------------|---------------|
| | Human (Years) | Monkey (Years) | Dog (Years) | Rabbit (Weeks) | Rat (Weeks) | Mouse (Weeks) |
| Fusion of Secondary Ossification Centers ^a | | | | | | |
| Femur distal epiphysis | 1419 | 3–6 | 0.7–0.9 | 32 | 15–162 | 12–13 |

^aZoetis (2003).**TABLE 13.13 Pulmonary System^a**

| Developmental Event | Postnatal Developmental Period (Days) | | |
|------------------------------------|---------------------------------------|-----|-------|
| | Human | Rat | Mouse |
| Alveoli Formation ^{b,c,d} | | | |
| Onset | Prenatal | 1–4 | 1–2 |
| Completion | 730 | 28 | 28 |

^aThe stages of lung development (glandular, canalicular, saccular, alveolar) at birth vary with the species. Human lungs have few alveoli and are considered in the alveolar stage at birth. Rodent lungs are less developed and considered in the saccular stage without alveoli at birth (Zoetis and Hurtt, 2003).^bBurri (1997).^cMerkus et al. (1996).^dTschanz and Burri (1997).**TABLE 13.14 Immune System**

| Developmental Event | Postnatal Developmental Period (Days) | |
|--|---------------------------------------|-------------------|
| | Human | Mice |
| B-cell development ^a | Prenatal | Prenatal |
| T-cell development ^a | Prenatal | Prenatal |
| NK cell development ^a | Prenatal | 21 |
| T-dependent antibody response ^a | 0 | 14 |
| | | 41–56 adult level |
| T-independent antibody response ^a | 45–90 | 0 |
| | | 14–21 adult level |
| Adult level IgG ^a | 1825 | 42–56 |

^aHolladay and Smialowicz (2000).

TABLE 13.15 Renal—Functional

| Developmental Event | Postnatal Developmental Period (Days) | |
|--|---------------------------------------|-------|
| | Human | Rat |
| Glomerulo-/nephrogenesis ^{a,b} | Prenatal | 8–14 |
| Adult GFR and tubular secretion ^{a,b} | 45–180 | 15–21 |

^aSnodgrass (1992).^bTravis (1991).**TABLE 13.16 Renal—Anatomical**

| Developmental Event | Postnatal Developmental Period (Weeks) | | | | | |
|--|--|-----|--------|-----|----------|-----|
| | Human | Dog | Rabbit | Rat | Mouse | Pig |
| Completion of nephrogenesis ^a | Prenatal week 35 | 2 | 2–3 | 4–6 | Prenatal | 3 |

^aZoetis (2003).**TABLE 13.17 Metabolism**

| Developmental Modulation of Phase I/II Metabolism | | | |
|---|-------------------------------|--|----------------------------|
| Enzyme | Maturation of Enzyme Activity | | |
| | Human (Years) | Rat (Days) | Rabbit (Days) |
| CYP2D6 ^{a,b} | 0–3 | NA | NA |
| CYP2E1 ^{b,c,d} | 0–1 | 4–17 ↓ postweaning Male > female | 14–35 2× adult At 35 |
| CYP1A2 ^{a,e,f,g} | 0.5 1 (>adult) | 7–100 low levels | 21–60 |
| CYP2C8 ^{a,b} | <1 | NA | NA |
| CYP2C9 ^{a,b} | <0.5 0.5 (>adult) | NA | NA |
| CYP3A4 ^b | 0–2 | NA | NA |
| Acetylation ^{a,b} | 1 (35% adult) | NA | NA |
| Methylation ^{a,b} | <1 (50% adult) | NA | NA |
| Glucuronidation ^{a,b} | 0 (>adult) | | NA |
| Sulfation ^{a,b} | 12 0 | NA | NA |

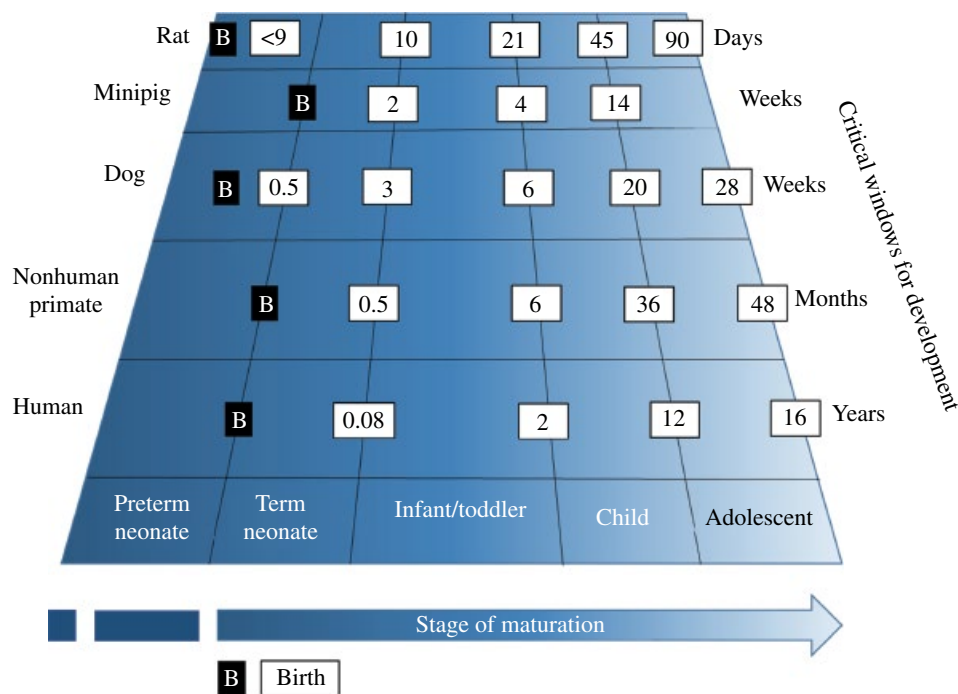
NA, not available.

^aKearns and Reed (1989).^bLeeder and Kerns (1997).^cWaxman et al. (1989).^dPeng et al. (1991).^eDing et al. (1992).^fImaoka et al. (1991).^gPineau et al. (1991).

TABLE 13.18 Cardiac^a

| Cardiac Parameter | Postnatal Developmental Period | | |
|--------------------------------------|---|--|---|
| | Human (Years) | Dog | Rat (Weeks) |
| Electrophysiology (ECG) | 5–7 | NA | 3–8 |
| Cardiac output (CO) and hemodynamics | Birth 138 bpm; adults 8 bpm. <2 years: smaller ventricular vol., stroke index, ejection fraction versus adult Birth BP 62/40; 2 months 85/47; 0.5–8 years. Diastolic 58–62 | Increase in BP and decrease in HR from 1 week to 0.5 years | Early increase in HR and then constant into adulthood High CO and low PVR Neonate–puberty systolic BP doubles and reaches maturity by 10 weeks Primarily diploid in infant and adult |
| Myocytes | Diploid at birth compared to 60% in adults (40% polyploidy) | NA | |
| Coronary vasculature | Diameter of arteries doubled at 1 year max at 30 years. Capillary angiogenesis occurs postnatally and density decreases with age | Capillary angiogenesis occurs postnatally and density decreases with age | Capillary angiogenesis occurs postnatally arterial maturation by 1 month |
| Cardiac innervation | Neuron number increases and reaches adult pattern/density in childhood | Continued development during 2–4 months | Adrenergic pattern matures by 3 weeks and nerve density matures by 5 weeks. Cholinergic matures postnatally |

NA, not available.

^aHew and Keller (2003).**FIGURE 13.13** Comparative stages of development for neonatal and juvenile toxicity studies. *Source:* Beulke-Sam (2003).

Gad, 2000). Various uses have been suggested for these *in vitro* tests including the following:

1. A general prescreen to select likely developmental toxins for subsequent whole-animal studies
2. A prescreen to select among possible backups to a lead drug candidate that had been found to be developmentally toxic
3. To study mechanisms of developmental toxicity
4. To provide supplementary information about developmental toxicity in addition to that provided by whole-animal studies
5. To replace whole animals for evaluating developmental toxicity

Uses one and five in the preceding text are very unlikely to be applicable to the pharmaceutical industry. One problem with *in vitro* systems for these purposes is that the percentages of the agents that are positive are very high, for example, 69% of agents tested in the mouse ovarian tumor (MOT) cell attachment test and 72% of agents tested in the mouse limb bud (MLB) assay. High correlations between *in vivo* and *in vitro* results have been reported based on the limited number of validation work completed. But these correlations have compared an *in vitro* end point to teratogenicity in laboratory animals without regard to maternal toxicity. Thus, the question that these screens seem to be answering is: Can this agent be teratogenic or developmentally toxic in laboratory animals at any dosage level? However, as discussed in the preceding text, it is not important for the purpose of safety assessment if an agent can be developmentally toxic in laboratory animals at high, maternotoxic dosages. The important question for prospective screens to answer is this: Is the agent a selective or direct developmental toxin? For these reasons, a promising drug candidate would not be dropped from development due to a positive result in a current *in vitro* test, and a negative result would not preclude the need for whole-animal studies.

To relate a positive finding in an *in vitro* test to the *in vivo* situation, one must either compare the concentration that caused the positive developmental effect *in vitro* to the exposure level of the embryo *in vivo* or compare the *in vitro* concentration for a developmental effect to the maternal toxicity that would be associated with exposure at that concentration *in vivo*. To do the necessary pharmacokinetic studies *in vivo* would defeat the purpose of using an *in vitro* test. It would be very desirable and may be possible, though, to have an end point in an *in vitro* test that would correlate with maternal toxicity.

Currently, only the *Hydra* system incorporates a measurement of "toxicity" to the adult to provide a comparison of the sensitivity of the "embryo" with that of the adult (Johnson et al., 1988). However, the *Hydra* screen

has not been fully validated as being predictive of results in mammals and has fallen from favor. Thus, a major goal of research directed toward developing an *in vitro* teratogen screen should be to find a simple yet appropriate measure of toxicity unrelated to development. This would allow the comparison of the dose for a 50% effect (ED_{50}) of "developmental toxicity" as measured *in vitro* to an ED_{50} for "adult" toxicity *in vitro*. The validation of such a dual *in vitro* system would involve comparing the developmental selectivity *in vitro* to that *in vivo* for a large number of compounds. In a preliminary effort in this regard, effects on cell division in the rat limb bud micromass assay were considered as a possible correlate of maternal toxicity (Wise et al., 1990a).

Another possible use of *in vitro* developmental toxicity tests would be to select the least developmentally toxic backup from among a group of structurally related compounds with similar pharmacological activity (use 2 from the preceding text), for example, when a lead compound causes malformations *in vivo* and is also positive in a screen that is related to the type of malformation induced. However, even for this limited role for a developmental toxicity screen, it would probably also be desirable to have a measure of the comparative maternal toxicity of the various agents and/or information on the pharmacokinetics and distribution of the agents *in vivo*.

In vitro developmental toxicity systems have clearly been useful for studies of mechanisms of developmental effects (e.g., Daston et al., 1989)—use 3 from the preceding text. It is unclear, though, whether *in vitro* developmental toxicity tests will provide useful information about developmental toxicity that is not derived from whole-animal studies (use 4). As is true for a possible use as a prescreen, the interpretation of a positive finding in an *in vitro* test will depend on knowing the exposure level *in vivo*. When this is known, the *in vitro* information could be helpful. The results of *in vivo* studies, though, would still likely be considered definitive for that species.

13.8 APPRAISAL OF CURRENT APPROACHES FOR DETERMINING DEVELOPMENTAL AND REPRODUCTIVE HAZARDS

The current system for testing new pharmaceuticals for developmental and reproductive toxicity has been largely intact since 1966. In that time, no thalidomide-like disasters have occurred. It cannot be proven, but there is a good chance that these two statements are linked, that is, that the testing system has prevented potent, selective, human teratogens from being marketed. Indeed, the development of many compounds has been terminated because of positive findings in standard developmental toxicity studies. We do not know for certain if any of these agents would have been developmental hazards in humans, but it seems very likely. Due to the limited information

on developmental toxicity of chemical agents in humans and the obvious inability to conduct controlled human studies, the correlation between animal studies and human findings is uncertain, and it is difficult to extrapolate precisely from animals to humans (see Frankos, 1985). However, the worst hazards—the few dozen selective developmental toxins that are known to be teratogens in humans—are generally also selective teratogens in animals. Thus, although the current battery of animal studies is not perfect, it appears to have been adequate and effective in performing the important task of preventing the widespread exposure of pregnant women to dangerous developmental toxins. In the few cases where new pharmaceuticals have been shown to cause malformations in humans, animal studies had been positive and provided an early warning to the potential problem. Concern with the risk of potential male-mediated developmental toxicity not being adequately assessed has led to the issuance of a new guidance (FDA, 2015b) and to suggesting steps such as monitoring drug in male test animal ejaculate, with concentration on *in vitro* genotoxicity as means of addressing the issue.

REFERENCES

- Adams MJ, Hicks PG, York MJ. (1989) Detection of pregnancy in the rabbit by assay of plasma progesterone. *Teratology*. 40:275.
- Adams MJ, Hicks PG, York MJ. (1990) Detection of pregnancy in the rabbit by assay of progesterone: an update. *Teratology*. 42:25A–26A.
- Anderson C, Clark RL. (1990) External genitalia of the rat: normal development and the histogenesis of 5 α -reductase inhibitor-induced abnormalities. *Teratology*. 42:483–496.
- Beckman DA, Feuston M. (2003) Landmarks in the development of the female reproductive system. *Birth Defects Res., Part B*. 68:137–143.
- Beulke-Sam J. (2003) *Comparative schedules of development in rats and humans: implications for developmental neurotoxicity testing*. Presented at the 42nd Annual Meeting of the Society of Toxicology, March 9–13, 2003, Salt Lake City, UT.
- Beyer PE, Chernoff N. (1986) The induction of supernumerary ribs in rodents: the role of maternal stress. *Teratog. Carcinog. Mutagen*. 6:149–429.
- Bloch E, Lew M, Klein M. (1971) Studies on the inhibition of fetal androgen formation. Inhibition of testosterone synthesis in rat and rabbit fetal testes with observations on reproductive tract development. *Endocrinology*. 89:16–31.
- Brown NA, Freeman SJ. (1984) Alternative tests for teratogenicity. *ATLA*. 12:7–23.
- Buelke-Sam J, Kimmel CA, Adams J. (1985) Design considerations in screening for behavioral teratogens: results of the collaborative behavioral teratology study. *Neurobehav. Toxicol. Teratol*. 7:537–789.
- Burri P. (1997) Structural aspects of prenatal and postnatal development and growth of the lung. In: *Lung Growth and Development* (McNoald JA, Ed.). Marcel Dekker, Inc., New York, pp. 1–35.
- Chapin RE, Gulati DR, Barnes LH. (1991) The effects of dietary restriction on reproductive endpoints in Sprague-Dawley rats. *Toxicologist*. 11:112.
- Chernoff N, Setzer RW, Miller DB, Rosen MB, Rogers JM. (1990) Effects of chemically induced maternal toxicity on prenatal development in the rat. *Teratology*. 42:651–658.
- CHMP. (2005) *Guideline on the Need for Non-Clinical Testing in Juvenile Animals on Human Pharmaceuticals for Paediatric Indication*. European Medicines Agency (EMA), London, October 2005.
- Chow BF, Rider AA. (1973) Implications of the effects of maternal diets in the various species. *J. Anim. Sci.* 36:167–173.
- Christian MS. (1983) Assessment of reproductive toxicity: state of the art. In: *Assessment of Reproductive and Teratogenic Hazards* (Christian MS, Glabraith M, Voytek P and Mehlman MA, Eds.). Princeton Scientific Publishers, Princeton, NJ, pp. 65–76.
- Christian MS, Hoberman AM. (1989) Current in vivo reproductive toxicity and developmental toxicity (teratology) test methods. In: *A Guide to General Toxicology* (Marquis JA and Hayes AW, Eds.), 2nd Ed. S. Karger, Basel, pp. 91–100.
- Clark RL, Robertson RT, Minsker DH, Cohen SM, Tocco DJ, Allen HL, James ML, Bokelman DL. (1984) Association between adverse maternal and embryo-fetal effects in norfloxacin-treated and food-deprived rabbits. *Fundam. Appl. Toxicol.* 7:272–286.
- Clark RL, Eschbach K, Cusick WA, Heyse JF. (1987) Interactions between caffeine and adenosine agonists in producing embryo resorptions and malformations in mice. *Toxicol. Appl. Pharmacol.* 91:371–385.
- Clark RL, Antonello JM, Soper KA, Bradstreet TC, Heyse JF, Ciminera JL. (1989) Statistical analysis of developmental toxicity data. *Teratology*. 39:445–446.
- Clark RL, Anderson CA, Prahalada S, Leonard YM, Stevens JL, Hoberman AM. (1990a) 5 α -reductase inhibitor-induced congenital abnormalities in male rat external genitalia. *Teratology*. 41:544.
- Clark RL, Antonello JM, Grossman JT, Wise LD, Anderson C, Bagdon WJ, Prahalada S, MacDonald JS, Robertson RT. (1990b) External genitalia abnormalities in male rats exposed *in utero* to finasteride, a 5 α -reductase inhibitor. *Teratology*. 42:91–100.
- Clark RL, Antonello JM, Wenger JD, Deyerle-Brooks K, Duchai DM. (1991) Selection of food allotment for New Zealand white rabbits in developmental toxicity studies. *Fundam. Appl. Toxicol.* 17:584–592.
- Cooper RL, Goldman JM. (2010) Female reproductive toxicity. In: *Reproductive Toxicology* (Kapp RW and Tyl RW, Eds.), 3rd Ed. Informa, New York, pp. 270–290.
- D'Aguzzo W. (1973) *Guidelines of Reproduction Studies for Safety Evaluation of Drugs for Human Use*. In: FDA Introduction to Total Drug Quality. DHEW Publication no. (FDA) 74-3006, Department of Health, Education and Welfare, Washington, DC.
- Daston GP, Yonker JE, Powers JF, Heitmeyer SA. (1989) Difference in teratogenic potency of ethylenethiourea in rats and mice: relative contribution of embryonic and maternal factors. *Teratology*. 40:555–556.
- Degenhardt K, Kladetzky J. (1955) Spinal deformation and chordal attachment. *Z. Mensch. Vererb. Konstitutionsl.* 33:151–192.

- DeSesso JM, Harris SB. (1995) Principles underlying developmental toxicity. In: *Toxicology Risk Assessment* (Fan A and Cgabg LW, Eds.). Marcel Dekker, New York.
- DeSesso JM, Willhite CC. (2009) Developmental toxicology. In: *General and Applied Toxicology* (Ballantyne B, Marrs TC and Syverson T, Eds.), 3rd Ed. John Wiley & Sons, Ltd, Chichester, pp. 1909–1952.
- Ding X, Peng HM, Coon MF. (1992) Cytochromes P450 NMa, Nmb (2G1) and LM4 (1A2) are differentially expressed during development in rabbit olfactory mucosa and liver. *Mol. Pharmacol.* 42(N6):1027–1032.
- Ellington S. (1980) *In vivo* and *in vitro* studies on the effects of maternal fasting during embryonic organogenesis in the rat. *J. Reprod. Fertil.* 60:383–388.
- Ellison AC, Maren TH. (1972) The effect of potassium metabolism on acetazolamide-induced teratogenesis. *Johns Hopkins Med. J.* 130:105–115.
- Environmental Protection Agency. (1991) *Pesticide Assessment Guidelines—Subdivision F, Hazard Evaluation: Human and Domestic Animals*. Addendum 10—Neurotoxicity Series 81, 82, and 83. Publication No. PB-91-154617. National Technical Information Service, Washington, DC.
- Fabro S, Shull G, Brown NA. (1982) The relative teratogenic index and teratogenic potency: proposed components of developmental toxicity risk assessment. *Teratog. Carcinog. Mutagen.* 2:61–76.
- FDA. (1966) *Guidelines for Reproduction Studies for Safety Evaluation of Drugs for Human Use*. Drug Review Branch, Division of Toxicological Evaluation, Bureau of Science, Food and Drug Administration, Washington, DC.
- FDA (1978) *Good laboratory practice regulations*. Federal Register, December 22, 1978, Title 21, Part II, Part 58, Washington, DC.
- FDA. (1979) Pregnancy category classification of fetal risks. *FDA Drug Bull.* 9(4):23–24.
- FDA. (1984) *Final Report of Task Force on Reproductive and Developmental Toxicity*. Review of Current Recommendations and Requirements for Reproductive and Developmental Toxicity Studies, Department of Health and Human Services, Washington, DC, 24 pp.
- FDA. (1993) Guideline for the study and evaluation of gender differences in the clinical evaluation of drug. *Fed. Regist.* 58(139):39405–39416.
- FDA. (2000) *Red Book 2000, Toxicological Principles for Safety assessment of Direct Food Additives and Color Additives and Color Used in Food*. National Technical Information Service, Washington, DC. (Also cited as (Red Book). Guidelines for a Three-Generation Reproduction Toxicity Study with Optional Teratology Phase. Guidelines for Reproduction Testing with a Teratology Phase.)
- FDA. (2006) *Guidance for Industry: Nonclinical Safety Evaluation of Pediatric Drug Products*. CDER, Rockville, MD, February, 2006.
- FDA. (2007) *Good Laboratory Practice for Nonclinical Laboratory Studies*. 21 CFR: Volume 1 Part 58. 1978 updated since, most recently. CDER, Rockville, MD, July 2007.
- FDA. (2011) *Guidance for Industry Reproductive and Developmental Toxicities: Integrating Study Results to Assess Concerns*. Center for Drug Evaluation and Research (CDER), Silver Spring, MD.
- FDA. (2014) *Guidance to Industry: Non-Clinical Safety Evaluation of Pediatric Drugs*. February 2006, CDER, Rockville, MD.
- FDA. (2015a) *Testicular Toxicity: Evaluation during Drug Development*. CDER, Rockville, MD, July 2015.
- FDA. (2015b) *Assessment of Male-Mediated Developmental Risk for Pharmaceuticals*. CDER, Rockville, MD, June 2015.
- Frankos V. (1985) FDA perspectives on the use of teratology data for human risk assessment. *Fundam. Appl. Toxicol.* 5:615–625.
- Gad SC. (2000) *In Vitro Toxicology*, 2nd Ed. Taylor & Francis, Philadelphia, PA.
- Grote W. (1969) Trunk skeletal malformations following blood loss in gravid rabbits. *Z. Anat. Entwicklungsgesch.* 128:66–74.
- Gulati K, Hope E, Chapin RE. (1991) The effects of dietary restriction on reproductive endpoints in Swiss mice. *Toxicologist.* 11:112.
- Gupta RC. (Ed.) (2011) *Reproductive and Developmental Toxicology*. Elsevier, San Diego, CA.
- Habert R, Picon R. (1984) Testosterone, dihydrotestosterone and estradiol-17 β levels in maternal and fetal plasma and in fetal testes in the rat. *J. Steroid Biochem.* 21:183–198.
- Hansen DK, Hodes ME. (1983) Comparative teratogenicity of phenytoin among several inbred strains of mice. *Teratology.* 28:175–179.
- Hemm R, Arslanoglou L, Pollock J. (1977) Cleft palate following prenatal food restriction in mice: association with elevated maternal corticosteroids. *Teratology.* 15:243–248.
- Hew KW, Keller KA. (2003) Postnatal anatomical and functional development of the heart: a species comparison. *Birth Defects Res., Part B.* 68:309–320.
- Hoberman AM, Lewis EM. (2011) *Pediatric Non-Clinical Drug Testing*. John Wiley & Sons, Inc., Hoboken, NJ.
- Holladay SD, Smialowicz R. (2000) Development of the murine and human immune system: different effects of immunotoxicants depend on time of exposure. *Environ. Health Perspect.* 108:463–473.
- Hood RD. (2012) *Developmental and Reproductive Toxicology: A Practical Approach*, 3rd Ed. Informa, Philadelphia, PA.
- ICH S5A(R2). (2000) *Detection of Toxicity to Reproduction for Medicinal Products*. The International Council on Harmonisation, Geneva, Switzerland.
- ICH S5B(R2). (2000) *Detection of Toxicity to Reproduction for Medicinal Products: Addendum on Toxicity to Male Fertility*. The International Council on Harmonisation, Geneva, Switzerland.
- Ikonomidou C, Bosch F, Miksa M, Bittigau P, Vockler J, Dikranian K, et al. (1999) Blockade of NMDA receptors and apoptotic neurodegeneration in the developing brain. *Science.* 283:70–74.
- Imaoka S, Fujita S, Funae Y. (1991) Age dependent expression of cytochrome P450s in Rat Liver. *Biochem. Biophys. Acta.* 1097(N3):187–192.
- In Vitro Teratology Task Force. (1987) Report. *Environ. Health Perspect.* F72:201–249.
- Johnson EM. (1981) Screening for teratogenic hazards: nature of the problems. *Annu. Rev. Pharmacol. Toxicol.* 21:417–429.
- Johnson EM. (1984) A prioritization and biological decision tree for developmental toxicity safety evaluations. *J. Am. Coll. Toxicol.* 3:141–147.

- Johnson EM, Christian MS. (1984) When is a teratology study not an evaluation of teratogenicity? *J. Am. Coll. Toxicol.* 3:431–434.
- Johnson EM, Newman LM, Gabel BEG, Boerner TF, Dansky LA. (1988) An analysis of the *Hydra* assay's applicability and reliability as a developmental toxicity prescreen. *J. Am. Coll. Toxicol.* 7:111–126.
- Johnston MC, Sulik KK, Dudley KH. (1979) Genetic and metabolic studies of the differential sensitivity of A/J and C57BL/6J mice to phenytoin ('Dilantin')-induced cleft lip. *Teratology.* 19:33A.
- Kapp RW. (2010) Male reproductive toxicity. In: *Reproductive Toxicology* (Kapp RW and Tyl RW, Eds.), 3rd Ed. Informa, New York, pp. 249–269.
- Karnofsky D. (1965) Mechanisms of action of certain growth-inhibiting drugs. In: *Teratology: Principles and Techniques* (Wilson J and Warkany J, Eds.). University of Chicago Press, Chicago, pp. 185–213.
- Kavlock RJ, Chernoff N, Rogers EH. (1985) The effect of acute maternal toxicity on fetal development in the mouse. *Teratog. Carcinog. Mutagen.* 5:3–13.
- Kearns LK, Reed MD. (1989) Clinical pharmacokinetics in infants and children. A reappraisal. *Clin. Pharmacokin.* 17(supp 1):29–67.
- Khera KS. (1984) Maternal toxicity—a possible factor in fetal malformations in mice. *Teratology.* 29:411–416.
- Khera KS. (1985) Maternal toxicity: a possible etiological factor in embryo-fetal deaths and fetal malformations of rodent-rabbit species. *Teratology.* 31:129–153.
- Khera KS. (1991) Chemically induced alterations in maternal homeostasis and histology of conceptus: their etiologic significance in rat fetal anomalies. *Teratology.* 44:259–297.
- Kimmel CA, Buelke-Sam J. (1994) *Developmental Toxicology*. Raven Press, New York.
- Klinefelter GR, Veeramachaneni DNR. (2014) Assessment of male reproductive toxicity. In: *Principles and Methods of Toxicity* (Hayes AW and Kruger CL, Eds.), 6th Ed. CRC Press, Boca Raton, FL, pp. 1601–1636.
- Korenbrot CC, Huhtaniemi IT, Weiner RI. (1977) Preputial separation as an external sign of pubertal development in the male rat. *Biol. Reprod.* 17:298–303.
- Leeder JS, Kearns GL. (1997) Pharmacogenetics in pediatrics: implications for practice, new frontiers in pediatric drug therapy. *Pediatr. Clin. North Am.* 44(1):55–77.
- Lewis EM, Barnett JF Jr, Freshwater L, Hoberman AM, Christian MS. (2002) Sexual maturation data for CRL Sprague-Dawley rats: criteria and confounding factors. *Drug Chem. Toxicol.* 25(4):437–458.
- Lin GHY. (1987) Prediction of teratogenic potential and a proposed scheme for teratogenicity screening of industrial research and development materials. *In Vitro Toxicol.* 1:203–217.
- Lochry EA. (1987) Concurrent use of behavioral/functional testing in existing reproductive and developmental toxicity screens: practical consideration. *J. Am. Coll. Toxicol.* 6:433–439.
- Lochry EA, Hoberman AM, Christian MS. (1984) Positive correlation of pup body weight with other commonly used developmental landmarks. *Teratology.* 29:44A.
- Makori N, Oneda S, McElhatton PR, Ratcliffe JM, Sullivan F. (2009) Reproductive toxicology. In: *General and Applied Toxicology* (Ballantyne B, Marrs TC and Syverson T, Eds.), 3rd Ed. John Wiley & Sons, Ltd, Chichester, pp. 1873–1908.
- Marty MS, Chapin R, Parks L, Thorsrud B. (2003) Development and maturation of the male reproductive system. *Birth Defects Res., Part B.* 68:125–136.
- Matsuzawa T, Nakata M, Goto I, Tsushima M. (1981) Dietary deprivation induces fetal loss and abortions in rabbits. *Toxicology.* 22:255–259.
- Matt DW, Lee J, Sarver PL, Judd HL, Lu JKH. (1986) Chronological changes in fertility, fecundity, and steroid hormone secretion during consecutive pregnancies in aging rats. *Biol. Reprod.* 34:478–487.
- Matt DM, Sarver PL, Lu JKH. (1987) Relation of parity and estrous cyclicity to the biology of pregnancy in aging female rats. *Biol. Reprod.* 37:421–430.
- Merkus PJFM, ten Have-Opbroek AAW, Quanjer PH. (1996) Human lung growth: a review. *Pediatr. Pulmonol.* 21:383–397.
- Millicovsky G, Johnston MC. (1981) Maternal hyperoxia greatly reduces the incidence of phenytoin-induced cleft lip and palate in A/J mice. *Science.* 212:671–672.
- Millicovsky G, DeSesso JM, Kleinman LI, Clark KE. (1981) Effects of hydroxyurea on hemodynamics of pregnant rabbits: a maternally mediated mechanism of embryotoxicity. *Am. J. Obstet. Gynecol.* 140:747–752.
- Minsker D, Bagdon W, MacDonald J, Robertson R, Bokelman D. (1990) Maternotoxicity and fetotoxicity of an angiotensin-converting enzyme inhibitor, enalapril, in rabbits. *Fundam. Appl. Toxicol.* 14:461–479.
- Mulay S, Varma DR, Solomon S. (1982) Influence of protein deficiency in rats on hormonal status and cytoplasmic glucocorticoid receptors in maternal and fetal tissues. *J. Endocrinol.* 95:49–58.
- Neumann F, von Berswordt-Wallrabe R, Elger W, Steinbeck H, Hahn JD, Kramer M. (1970) Aspects of androgen-dependent events as studied by antiandrogens. *Recent Prog. Horm. Res.* 26:337–410.
- Nishimura M, Iizuka M, Iwaki S, Kast A. (1982) Repairability of drug-induced "wavy ribs" in rat offspring. *Arzneim.-Forsch./Drug Res.* 32:1518–1522.
- Palmer AK. (1981) Regulatory requirements for reproducing toxicology: theory and practice. In: *Developmental Toxicology* (Kimmel CA and Buelke-Sam J, Eds.). Raven Press, New York, pp. 259–287.
- Pen HM, Porter TD, Ding XX, Coon MJ. (1991) Differences in the developmental expression of rabbit cytochromes P450 2E1 and 2E2. *Mol. Pharmacol.* 40(N1):58–62.
- Pineau T, Daujat M, Pichard L, Girard F, Angevain J, Bonfils C, Maurel P. (1991) Developmental expression of rabbit cytochrome P450 CYP1A1, CYP1A2, CYP3A6 genes, effect of weaning and Rifampicin. *Curr. J. Biochem.* 197(N1):145–153.
- Persaud TVN. (1985) Teratogenicity testing. In: *Basic Concepts in Teratology* (Persaud TVN, Chudley AE and Skalko RG, Eds.). Alan R. Liss, New York, pp. 155–181.
- Radde IC. (1985) Mechanism of drug absorption and their development. In: *Textbook of Pediatric Clinical Pharmacology*

- (Macleod SM and Radde IC, Eds.). PSG Publishing Co., Littleton, MA, pp. 17–43.
- Rice D, Barone S Jr. (2000) Critical periods of vulnerability for the developing nervous system: evidence from humans and animal models. *Environ. Health Perspect.* 108(Suppl. 3):511–533.
- Saegusa T, Kaneko Y, Sato T, Nrama I, Segima Y. (1980) BD40A-induced wavy ribs in rats. *Teratology.* 22:14A.
- Salewski VE. (1964) Färbemethode zum makroskopischen Nachweis von Implantationsstellen am Uterus der Ratte. Naunyn-Schmiedeberg's. *Arch. Exp. Path. U. Pharmacol.* 247:367.
- Schardein J. (1987) Approaches to defining the relationship of maternal and developmental toxicity. *Teratog. Carcinog. Mutagen.* 7:255–271.
- Schardein J. (1988) Teratologic testing: status and issues after two decades of evolution. *Rev. Environ. Contam. Toxicol.* 102:1–78.
- Schardein J. (2000) *Chemically Induced Birth Defects*, 3rd Ed. Marcel Dekker, New York.
- Schwetz BA. (1981) Monitoring problems in teratology. In: *Scientific Considerations in Monitoring and Evaluating Toxicological Research* (Gralla EJ, Ed.). Hemisphere, Washington, DC, pp. 179–192.
- Shepard T, Lemire RJ. (2010) *Catalog of Teratogenic Agents*, 13th Ed. Johns Hopkins University Press, Baltimore, MD.
- Sidhu RS, Del Bigio MR, Tuor UI, Sechia SS. (1997) Low-dose Vigabatrin (gamma-vinyl GABA)-induced damage in the immature rat brain. *Exp. Neurol.* 144:400–405.
- Snedecor GW, Cochran WG. (1989) Covariance analysis. In: *Statistical Methods*, 8th Ed., Chapter 10. Iowa State University Press, Ames, IA.
- Snodgrass WR. (1992) Physiological and biochemical differences between children and adults as determinants of toxic response to environmental pollutants. In: *Similarities and Differences Between Children and Adults, Implications for Risk Assessment* (Guzelian PS, Henry CJ and Olin SS, Eds.). ILSI Press, Washington, DC, pp. 35–42.
- Soper KA, Clark RL. (1990) Exact permutation trend tests for fetal survival data. *Proc. Biopharm. Sect. Am. Stat. Assoc.* 263–268.
- Staples RE. (1971) Blastocyst transplantation in the rabbit. In: *Methods in Mammalian Embryology* (Daniel JC Jr, Ed.). W. H. Freeman, San Francisco, CA, pp. 290–304.
- Szabo K, Brent R. (1975) Reduction of drug-induced cleft palate in mice. *Lancet.* 305:1296–1297.
- Tanimura T. (1990) The Japanese perspectives on the reproductive and developmental toxicity evaluation of pharmaceuticals. *J. Am. Coll. Toxicol.* 9:27–38.
- Travis LB. (1991) Chapter 25: The kidney and urinary tract morphogenic development and anatomy. In: *Rudolph's Pediatrics*, 19th Ed. Appleton & Lange, New York, pp. 1223–1236.
- Tschanz SA, Burri PH. (1997) Postnatal lung development and its impairment by glucocorticoids. *Pediatr. Pulmonol.* 16(Suppl):247–249.
- Turner MR. (1973) Perinatal mortality, growth, and survival to weaning in offspring of rats reared on diets moderately deficient in protein. *Br. J. Nutr.* 29:139–147.
- Watkinson WP, Millicovsky G. (1983) Effect of phenytoin on maternal heart rate in A/J mice: possible role in teratogenesis. *Teratology.* 28:1–8.
- Waxman DJ, Morrissey JJ, Le Balnc GA. (1989) Female predominant rat hepatic P450 forms (IIE1) and 3(IIA1) are under hormonal regulatory controls distinct from those of sex specific P450 forms. *Endocrinology.* 270(N2):458–471.
- Weaver TE, Scott WJ. (1984a) Acetazolamide teratogenesis: association of maternal respiratory acidosis and ectrodactyly in C57Bl/6J mice. *Teratology.* 30:187–193.
- Weaver TE, Scott WJ. (1984b) Acetazolamide teratogenesis: interaction of maternal metabolic and respiratory acidosis in the induction of ectrodactyly in C57Bl/6J mice. *Teratology.* 30:195–202.
- Weindruch R, Walford RL. (1988) *The Retardation of Aging and Disease by Dietary Restriction*. Charles C. Thomas, Springfield, IL.
- Wickramaratne GA de S. (1988) The post-natal fate of supernumerary ribs in rat teratogenicity studies. *J. Appl. Toxicol.* 8:91–94.
- Wilson J. (1965) Methods for administering agents and detecting malformations in experimental animals. In: *Teratology, Principles and Techniques* (Wilson JG and Warkany J, Eds.). The University of Chicago Press, Chicago, pp. 262–277.
- Wise LD, Clark RL, Minsker DH, Robertson RT. (1988) Use of hematology and serum biochemistry data in developmental toxicity studies. *Teratology.* 37:502–503.
- Wise LD, Clark RL, Rundell JO, Robertson RT. (1990a) Examination of a rodent limb bud micromass assay as a pre-screen for developmental toxicity. *Teratology.* 41:341–351.
- Wise LD, Vetter CM, Anderson C, Antonello JM, Clark RL. (1990b) Reversible effects on external genitalia in rats exposed *in vitro* to triamcinolone acetonide. *Teratology.* 41:600.
- Wood SL, Beyer BK, Cappon GD. (2003) Species comparison of postnatal CNS development: functional measures. *Birth Defects Res.* 68(5):391–407.
- Working PK. (1988) Male reproductive toxicology: comparison of the human to animal models. *Environ. Health Perspect.* 77:37–44.
- Yamada T, Ohsawa K, Ohno H. (1988) The usefulness of alkaline solutions for clearing the uterus and staining implantation sites in rats. *Exp. Anim. (Tokyo).* 37:325–332.
- York RG, Parker RM, Haber LT. (2014) Test methods for assessing female reproductive and developmental toxicology. In: *Principles and Methods of Toxicology* (Hayes AW and Kruger CL, Eds.), 6th Ed. CRC Press, Boca Raton, FL, pp. 1637–1722.
- Zeman FJ. (1967) Effect on the young rat of maternal protein restriction. *J. Nutr.* 93:167–173.
- Zoetis T. (2003) Species comparison of anatomical and functional renal development. *Birth Defects Res., Part B.* 68:111–120.
- Zoetis T, Hurtt ME. (2003) Species comparison of lung development. *Birth Defects Res. Part B.* 68:121–124.
- Zoetis T, Walls I. (2003) *Direct Dosing of Pre-Weaning Mammals in Toxicity Testing and Research*. ILSI Press, Washington, DC.

CARCINOGENICITY STUDIES

14.1 INTRODUCTION

Cancer is a group of diseases characterized by the uncontrolled growth and spread of abnormal cells. If the spread is not controlled, it can result in death by a variety of mechanisms. Cancer is caused by both external factors (physical effects, infectious organisms, chemicals, radiation) and internal factors (inherited mutations, hormones, suppressed immune function, and mutations that can occur due to metabolism). These causes, as will be discussed later, can be genetic or not. These causal factors may act together or in sequence to initiate or promote the development of cancer. Usually, but not always, decades pass between external causes and the appearance of detectable cancer.

While the causes of all cancers (also called malignant neoplasia or malignant oncogenicity) are not known, more than half a million deaths a year are attributed directly to the presence of cancer in the United States. It is estimated that there will be 1 660 290 new cancer cases in the United States in 2013 (American Cancer Society, 2013) and a sizable portion of these are believed to be due to a wide range of environmental factors. In 1775, Percivall Pott first described the association between exposure to a xenobiotic (chimney soot) and an increase in the incidence of cancers (scrotal and nasal). Since Pott's early observations and particularly since World War II, the understanding of the associations and mechanisms underlying chemically and physically induced cancer has grown due to epidemiological observations and from the results of experimental studies performed in mammalian and *in vitro* studies (Crum Brown and Fraser, 1869; Diamandopoulos, 1996; Creech, 2000).

We do know that specific factors (intrinsic and external) associated with the development of cancer are correlated with specific forms of neoplasia (not with generalized multiple organ effects). For external factor-associated cancers, we can identify causes and their relative risks using well-established experimental models. At the same time, as will be explored later in this chapter, there are predictive tests (short-term *in vitro* and computational or QSAR) which are quite effective in identifying potential risks of carcinogenicity.

The multistep, multifactorial process by which normal cells are thought to be transformed into neoplastic cells involves the loss of genomic stability and the sequential acquisition of genetic alterations (Loeb and Loeb, 2000; Eyfjord and Bodvardsdottir, 2005). Driver mutations in cells are those which are causally involved in the neoplastic process and are selected for during the process of tumorigenesis. Passenger mutations are neutral and nonlethal; therefore they are retained by chance during cell population expansion. Proto-oncogenes and tumor suppressor genes have been identified as the main mutational targets (Bishop, 1991), and carcinogen-specific mutational patterns have been observed in these genes in tumors of laboratory animals (Balmain and Brown, 1988) as well as in human cancers (Semenza and Weasel, 1997). This association suggests a multistep mechanistic link between carcinogen exposure, genetic alterations, and cancer (Hussain and Harris, 1999). The highly increased tumor incidence in subjects with defects in nucleotide excision repair (NER) supports the key role of DNA alterations in the process of cancer development (Stary and Sarasin, 2002). Moreover, organ- and cell-type-specific differences in DNA repair capacity have been demonstrated to correlate with site of tumor formation under a variety of experimental situations.

In the hands of skilled practitioners, the results of *in vitro* studies, QSAR evaluations, and shorter-term (3–6-months) animal studies can serve to identify those substances that have a low probability of being carcinogenic (Sistare et al., 2011).

14.1.1 History of Xenobiotic Carcinogenesis

A brief outline of the history of xenobiotic-related carcinogenesis, together with details of the causative agents and typical patients, is provided in Table 14.1.

TABLE 14.1 Historical Identification of Chemically Induced Cancer

| Date | Investigator(s) | Causative Agent |
|------|--------------------------------------|---|
| 1775 | Pott | Soot and chimney sweeps |
| 1822 | Ayrton | Arsenic-containing metal |
| 1875 | Thiersch | Sunlight |
| 1876 | Manourriez | Coal tar |
| 1879 | Harting and Hesse | Lung cancer and uranium |
| 1892 | Butlin | Soot and chimney sweeps |
| 1895 | Rehn | Manufacture of aniline dyes |
| 1902 | Frieben | X-rays |
| 1915 | Davis | Pipe smokers and bétel nut chewers |
| 1915 | Yamagiwa, Ichikawa, and Tsutsui | Induction of skin cancer in rabbits and mice by coal tar |
| 1920 | Leitch and Seguína | Radium radiation |
| 1928 | Delore and Bergamo | Benzene |
| 1930 | Kennaway and Hieger | Tumor induction by dibenz[<i>a,h</i>]anthracene |
| 1932 | Alwens, Bischoff, and Hexorvalent | Chromium compounds |
| 1933 | Cook, Hewett, and Hieger | Isolation of the carcinogen benzo[<i>a</i>]pyrene from coal tar |
| 1934 | Wood and Gloyne | Arsenicals, beryllium, and asbestos |
| 1934 | Neitzel | Mineral oil mists and radiation |
| 1936 | Yoshida and Kinoshita | Induction of liver cancer in rats by <i>o</i> -aminoazotoluene |
| 1936 | Kawahata | Coal tar fumes |
| 1938 | Hueper, Wiley, and Wolfe | Induction of urinary cancer in dogs by 2-naphthylamine |
| 1941 | Berenblum, Rous, MacKenzie, and Kidd | Initiation and promotion stages in skin carcinogenesis with benzo[<i>a</i>]pyrene |
| 1943 | Yale researchers | Lung cancer: mustard gas |
| 1956 | Doll and Hill | Lung cancer and other causes of death in relation to smoking |
| 1962 | Weiss | Lung cancer: bis(chloromethyl) ether (BCME) |
| 1982 | Dunkelberg | Ethylene oxide |
| 1991 | Fingerhut | Dioxin and liver cancer |
| 2005 | Schiffman | Human papillomavirus |

14.2 MECHANISMS AND CLASSES OF CARCINOGENS

A carcinogen is any substance or agent that significantly increases the incidence of cancer (Herzig and Christofori, 2002; Pierce and Damjanov, 2006; Gad et al., 2015). Carcinogens can be chemical or physical, the latter including UV light and ionizing radiation. The general feeling among cancer researchers is that carcinogenesis is a multistage genetic and epigenetic process, requiring on the average a minimum of two up to approximately six to eight mutations. These changes produce a shift in the homeostatic state of a cell or cells from that which is a highly regulated population of cells, which would normally differentiate and at some time point become apoptotic to a new population of cells which exhibit unregulated control, proliferate freely, demonstrate autonomous growth, exhibit reduced levels of apoptosis, sustain angiogenic activity, become invasive, and show metastatic behavior. So, succinctly put, a mutated or initiated cell divides or proliferates to eventually form a neoplasm. A neoplasm or tumor can be benign or malignant.

While chemicals or agents that cause cancer can be effective at a variety of levels, they can be categorized into two broad groups, those chemicals or agents that are genotoxic and those that are epigenetic. Epigenetic carcinogens are not genotoxic, but do involve interactions with genomic DNA. Genetic carcinogens react with DNA.

14.3 GENOTOXIC CARCINOGENS

A genotoxic carcinogen is an agent or chemical that reacts with DNA, causing some sort of damage to the DNA (McKinnell, 2006; Perantoni, 2006; Klaunig and Kamendulis, 2008; Bhandari et al., 2012; Klaunig, 2013; Preston and Hoffmann, 2013). This damage can include a chemical reaction in which chemical reacts with and binds via a covalent bond to the DNA, forming a DNA adduct. In either case, whether the DNA is just damaged or structurally modified by the formation of an addition product (DNA adduct), a mutation can eventually be produced, and the process of carcinogenesis begun.

In addition to genotoxic chemicals being mutagenic, they can also function as complete carcinogens (see following text), do not exhibit any threshold level of activity, and show a dose response-like tumorigenic effect. It is important to note that a not insignificant number of carcinogens do not cause the development of cancer in and of themselves. A variety of genotoxic carcinogens exist and include such chemicals as polyaromatic hydrocarbons, alkylating agents, aromatic amines, and amides. Damage to DNA or

DNA-adduct formation can result in the development of transversions or transitions:

- Transitions are a substitution of a purine with a different purine base or a pyrimidine with another pyrimidine base.
- Transversions are the substitution of a purine base with a pyrimidine base and vice versa.

Such simple substitution changes can lead to a misreading of the genetic code. Any misread copy (daughter strand of originally mutated DNA) can have the mutation fixed in place, when DNA replicates and become impervious to further repair processes of DNA.

Alternatively, misreading can result in what is known as a frameshift mutation. This type of mutation most commonly occurs in a situation where a DNA adduct has formed. What happens here is that the reading frame is shifted, either upstream or downstream, essentially garbling the genetic message. Finally, the DNA strands can actually break as a result of DNA-adduct formation. This is seen in situations where the formation of DNA adducts involves either the phosphodiester backbone or a DNA base. A single-strand break can lead to the formation of double-strand breaks, recombination, or other alterations in DNA structure and function.

As a complete class, genotoxic carcinogens can be subclassified as to whether or not they require metabolic activation to react with DNA. Chemicals that require metabolic activation to react with DNA are termed indirect-acting carcinogens and include, but are not limited to, such chemicals as aflatoxin, mycotoxins, azo dyes, and *N*-nitrosamines (Miller and Miller, 1981b; Conney, 1982; Miller et al., 1983; Pegg, 1984; Swenberg et al., 1984, 1985; Ames and Shigenaga, 1993; Klaunig and Kamendulis, 2004; Perantoni, 2006; Klaunig and Kamendulis, 2008; Groopman and Wogan, 2011; Penning, 2011; Bhandari et al., 2012; Klaunig, 2013). These chemicals require some sort of metabolic activation. Commonly these unactivated chemicals are referred to as procarcinogens, and the intermediate metabolites and final metabolites are referred to as proximate carcinogens and ultimate carcinogens, respectively. It is these proximate and ultimate forms that react with DNA to cause mutation and start the process of carcinogenesis. Depending upon the metabolic processes and pathways involved, there can be more than one proximate and more than one ultimate carcinogen for each procarcinogen, and in many cases the actual chemical species reacting with DNA is not known.

Keep in mind that while metabolic activation pathways are actively working, detoxification pathways are also functioning and both competing against each other for the same procarcinogen substrate. Some of these detoxification pathways may even result in inactivation of the carcinogen. Indirect-acting carcinogens do not generally manifest their carcinogenic effects at the original site of exposure, but

rather manifest their toxicity at the site or sites of their metabolism and transformation into moieties, which can react with DNA and form DNA adducts.

Direct-acting carcinogens do not need any sort of metabolic change in their structure in order to react with DNA. Hence for these types of compounds, the unmetabolized form of the chemical represents the ultimate carcinogen. Direct-acting carcinogens include, but are not limited to, such agents as dimethyl sulfate, methyl methanesulfonate, nitrogen mustards, and bis(chloromethyl) ether (Sontag, 1981; Pollard and Luckert, 1985; Marselos and Vainio, 1991; Kiaris, 2006; Perantoni, 2006; Klaunig and Kamendulis, 2008; Bhandari et al., 2012; Klaunig, 2013). Direct-acting carcinogens cause the formation of tumors at the site of their initial exposure and can also be carcinogenic at multiple sites. These chemical agents are ultimate carcinogens in their native forms and also in their native and unmetabolized forms possess electrophilic centers, which seek out and bind to nucleophilic centers (DNA and macromolecules).

A characteristic held in common by both indirect- and direct-acting carcinogens is that they all have strongly electrophilic centers. As electrophiles, they are highly reactive, stable, long lived, and capable of forming strong covalent bonds with other molecules possessing nucleophilic centers and seeking out the nucleophilic centers of other molecules, especially DNA. Molecules possessing oxygen, sulfur, and nitrogen atoms are ideal targets for an electrophile, because of the unpaired electrons on these particular atoms. This means that good sources for nucleophilic targets are DNA bases and the phosphodiester backbone of DNA. The strength of a given electrophile is important. Strong electrophiles can readily react with both weak and strong nucleophiles. But weak electrophiles are capable of only reacting with strong nucleophiles. Examples of electrophiles include carbonium ions, diazonium ions, epoxides, strained lactones, and various free radicals. The ability of an electrophile to form an adduct with DNA is in part limited by the structure of DNA itself, the size and spatial configuration of the electrophile itself, and the metabolic environment in the immediate environment of the DNA. While adducts of DNA can form at any site on DNA, the most common sites of alkylation are generally considered to be O⁶, N³, and/or N⁷ of guanine bases and N¹ and/or N³ of adenine bases. The most common adduct reaction with the phosphodiester backbone of DNA is ethylation.

There is some debate as to the effective carcinogenic potency of the various types of DNA adducts formed, but this has not been demonstrated unequivocally. The effectiveness of either direct-acting or indirect-acting genotoxic carcinogens depends highly upon the strength of attraction between the electrophile and DNA, stability of the electrophile, membrane permeability transport, as well as successful competition with detoxification pathways and the presence of, and opportunity for reaction with, other cellular nucleophiles.

Undoubtedly DNA adducts are formed each and every day, but the mechanisms for the repair of DNA in the cell keep the frequencies of these mutations at a low level. The interval between the time of the DNA insult and the time of its repair is critical. Lack of timeliness in repair can fix the mutation in place forever. However, just because DNA damage exists does not necessarily mean that the process of carcinogenesis has started. Indeed the pathway to carcinogenesis depends strongly upon such factors as the frequencies of DNA replication and repair within a given tissue or organ. In actuality, the development of cancer following an exposure to a carcinogen is relatively rare.

While the whole process of carcinogenesis is very complex, with the involvement of many different factors, the repair of DNA is a very potent defense against the development of genotoxicity and the development of cancer (Rupp and Howard-Flanders, 1968; Lehmann, 1972; di Caprio and Cox, 1981; Prakash, 1981; Neumann, 1983; Becker and Shank, 1985; Morrison et al., 1989; Perera et al., 1991; Hsieh and Hsieh, 1993; Volkert, 1998; Masutani et al., 1999; Sancar, 2003; Watson et al., 2004; Zhang and Lawrence, 2005; Kiaris, 2006; Lopes et al., 2006; Perantoni, 2006; Callegari and Kelly, 2007; Klaunig and Kamendulis, 2008; Asagoshi and Wilson, 2011; Basu, 2011; Peng et al., 2011; Bhandari et al., 2012; Klaunig, 2013). During the repair of DNA, the region of DNA that is damaged or has been modified through the formation of a DNA adduct, the repair systems detect, recognize, remove, and repair the DNA with a patch, so to speak, using the intact opposite strand as a building template. The DNA patch or segment is then spliced into the DNA molecule, replacing the defective segment.

It cannot be strongly enough emphasized that any repair of DNA needs to occur before mitosis occurs. If repair has not taken place before mitosis occurs, the damaged areas of DNA become fixed, and this in turn can lead to the incorrect pairing of bases, rearrangements, translocations, and other mutagenic sequelae or clastogenic events, which can start the cell down the path of carcinogenesis. While the carcinogen itself is an important parameter in the process of neoplastic transformation, the rate of mitosis for a given cell is another extremely important factor that needs to be considered. A variety of mechanisms are available to a cell to repair DNA, and there is some redundancy of activity between processes. While these processes are highly effective, they are not 100% efficient and flawless in the conduct of their function. Lack of accuracy in the repair process can leave residual DNA damage, which can ultimately lead to transcription and translation errors from the mutated genetic templates. This in turn can lead to the formation of proteins and other molecules with altered structure and function. These altered proteins and enzymes can impact the carcinogenesis process. Mutations of oncogenes, tumor suppressor genes, and other genes controlling the cell cycle or cell communication can lead to the development of a group of cells with a distinct

advantage for survival as compared to other cells. Common forms of DNA damage include, but are not limited to, adduct formation, DNA cross-linking, hydroxylation of bases, single-strand DNA breaks, double-strand DNA breaks, and the loss of bases. DNA repair processes include excision repair systems, postreplication repair, direct reversal of DNA damage, and nonhomologous end joining.

Regions of DNA that contain bases that are chemically modified or possess adducts are generally repaired by excision repair processes. Excision repair systems include mismatch repair, NER, and base excision repair activities:

- In base excision repair, a single base, whose structure has been altered by oxidation, alkylation, hydrolysis, or deamination, is removed and replaced.
- In NER, bulky helix distorting lesions are recognized and removed.
- In mismatch repair, errors of DNA replication and recombination that have resulted in mispaired but undamaged nucleotides are corrected. Spontaneous mutations occur frequently and are known as point mutations. That is a change in a single base in a base pair unit.

However, the challenge here for the repair system is determining which strand is the damaged or nontrue strand of DNA. Depurination is an activity that occurs fairly frequently and produces a site where the purine base is missing. If not repaired, mechanisms for the synthesis and repair of DNA are unable to determine what the appropriate base should be at the replication fork for proper pairing and a mutation can be the end result. Endonucleases cut DNA near apurinic sites and then extend that cut via the action of exonucleases. The resulting gap is then repaired by the actions of DNA polymerases and ligases. The end-joining repair of nonhomologous DNA involves the joining of pieces of free DNA. Pieces of free DNA typically result from the formation of double-stranded breaks in DNA as a result of exposure to ionizing radiation or some antineoplastic drugs.

The problem with this type of repair is that many times these pieces of DNA are translocated from one chromosome to another chromosome. This leads to the potential of placement of, for example, a proto-oncogene next to, for example, a gene controlling some aspect of cell growth. So the cell growth gene is now under the control of a different gene promoter. So even though the DNA is rejoined and not lost, it is in the incorrect place in the DNA for normal function. Double-stranded breaks in DNA are only correctly repaired when the free ends are rejoined exactly as they were before. Also, when this rejoining of DNA takes place, base pairs are not infrequently lost at the juncture point, which can lead to miscoding and misreading of DNA and a subsequent mutational change.

Postreplicational repair (PRR) is repair that takes place after DNA has replicated. A number of polymerases have been implicated in this process. The elucidation of postreplication error processes is currently an active area of research in molecular biology and still very much in flux in terms of the details. However, the activity of PRR enzymes may be controlled by the postreplication checkpoint response that has been observed in eukaryotic cells. Cells are also known to eliminate some types of damage to their DNA by chemically reversing it. These processes do not require a template, because this type of damage can occur in only one of the four bases. Direct reversal does not involve the breakage of the phosphodiester backbone. An example is the formation of pyrimidine dimers, two pyrimidines directly connected by a covalent bond. This dimerization occurs as a result of exposure to UV light. An enzyme, photolyase, which is activated by a very specific wavelength range of UV light (300–500nm), directly reverses this change. In another example of this direct reversal of damage, the methylation of guanine bases can be directly reversed by the action of guanine methyltransferase (MGMT).

14.4 EPIGENETIC CARCINOGENS

Any carcinogen that is not genotoxic can contribute to the process of carcinogenesis in what is termed an epigenetic manner (Trosko et al., 1998; Powell and Berry, 2000; Moggs et al., 2004; Kiaris, 2006; Klaunig and Kamendulis, 2008; Darwanto et al., 2011; Bhandari et al., 2012; Klaunig, 2013). The key difference from genotoxic carcinogens is that agents or metabolites of agents of this category do not *directly* damage, bind to, or interact with DNA. These agents are typically active only above a certain threshold level, and the process by which they work may be species specific. A number of different subcategories exist within this group, and they and their mechanisms of action are outlined in the following text.

The methylation of DNA is a normal ongoing cellular process, and this activity can profoundly affect the level of gene expression and the process of carcinogenesis (Holliday, 1990; Baylin, 1997; Klaunig and Kamendulis, 2004; Klaunig and Kamendulis, 2008; Klaunig, 2013). If there is a copious amount of methylation within a gene, it is usually silenced. Alternatively if there is little methylation within a gene, it is highly active. DNA methylation takes place as a result of the activity of DNA methylases, which obtain their methyl groups for transfer to the cytosine bases in DNA from *S*-adenosylmethionine or choline. Lower than normal levels of DNA methylation have been associated with increases in the rates of mutation. Chemical agents such as phenobarbital and diethanolamine are associated with methyltransferase activity, the alteration of DNA methylation patterns, and resultant changes in chromosomal structure. Deviations

from the norm in methylation patterns have been associated with the activity of tumor suppressor genes, bladder cancer, retinoblastoma, and metastatic tumors.

Reactive oxygen species (ROS) can contribute to the process of carcinogenesis and are produced by the reduction of molecular oxygen (Rao and Reddy, 1991; Clarkson and Thompson, 2000; Ding et al., 2000; Simeonova and Luster, 2000; Abuja and Albertini, 2001; Martindale and Holbrook, 2002; Klaunig and Kamendulis, 2008; Klaunig, 2013). Two chemicals that have been associated with the production of ROS are ethanol and acrylonitrile. Many different ROS exist and include such moieties as hydrogen peroxide, the superoxide anion, and the hydroperoxyl radical. Both endogenous (e.g., macrophages, oxidative phosphorylation, P450 metabolism, peroxisomes) and exogenous (e.g., metals, radiation, redox-cycling compounds) sources can cause the formation of oxygen radicals. These oxygen radicals can be neutralized by antioxidants, which can be endogenous or exogenous and can be enzymatic (e.g., catalase, superoxide dismutase) or nonenzymatic (e.g., vitamin C, vitamin E).

As an example, peroxisome proliferators (e.g., clofibrate) increase the production of peroxisomes by interacting as an agonist with a peroxisome proliferator-activated receptor (PPAR) by increasing the formation of peroxisomes (Rao and Reddy, 1991; Peters et al., 1998; Pugh et al., 2000; Burkhardt et al., 2001; Klaunig et al., 2003; Klaunig and Kamendulis, 2008; Klaunig, 2013). These receptors exist in different forms (alpha, beta, delta, gamma) and stimulate the formation of hydrogen peroxide, which can cause oxidative stress in the cell. ROS of this oxidative stress can cause damage to cellular macromolecules, biomembranes, and DNA. The damage to DNA can include but is not limited to strand breaks, cross-linking, and base modifications. This DNA damage can result in a lack of or faulty induction, replication errors, faulty or lack of signal transduction, and a generalized instability of the genome. Damage to biomembranes can result in the production of highly reactive electrophilic compounds such as epoxides, which can inflict damage to DNA as mentioned earlier. ROS can also influence various signal transduction pathways altering genetic expression in a positive or negative fashion. Such alteration can lead to the development of apoptosis, altered differentiation, or changes in the rates of cellular proliferation.

A number of chemicals have been shown to be capable of causing the formation of neoplasms through receptor-mediated pathways as well as via the disturbance of hormonal homeostasis (Noller et al., 1972; Tavani et al., 1993; Capen et al., 1995; Li et al., 1997; Hood et al., 1999; Klaunig and Kamendulis, 2008; Klaunig, 2013). Some of these chemicals include but are not limited to tamoxifen, diethylstilbestrol (DES), and phenobarbital. Some hormones are also known to cause cell proliferation in their target organ. A loss of hormonal control can lead to the development of persistently elevated levels of a given hormone and rampant cellular

proliferation and resultant neoplasm formation. Epidemiologic data on patients afflicted with breast or ovarian cancer has revealed that individuals with high levels of estrogen as well as those individuals exposed to DES were at an increased risk to develop cancer. Indeed, a classic example involves the compound DES. Exposed females gave birth to daughters, who had significantly higher incidences of cancers of the vagina and cervix. The mechanism here is ostensibly the ability of DES to cause aneuploidy. Synthetic estrogens and anabolic steroids have been implicated in the development of hepatic adenomas in both laboratory animals and human females. These adenomas shrink or disappear with removal to hormone (oral contraceptive) stimulation. Yet, interestingly enough, estrogens are protective in males with prostate cancer. Thyroid-stimulating hormone (TSH) drives proliferative activity in the thyroid gland, and agents (e.g., phenobarbital) that increase the levels of TSH lead to the development in the thyroid of hypertrophy, hyperplasia, and neoplasia.

The repeated exposure of cells to chemicals or agents that produce cell death cause a cycle of persistent cell growth or hyperplasia in an attempt to regenerate lost organ function (Andersen et al., 1998; Klaunig and Kamendulis, 2008; Klaunig, 2013). Such constant stimulation and resultant replicative activity increases the potential and probability for the development of the spontaneous mutation of DNA. As numbers of mutated cells increase in number and proliferate, pre-neoplastic lesions develop, and these can eventually lead to the formation of neoplasms. An important point to note here is that many carcinogens, whether they are genotoxic or not, at sufficiently high enough doses can cause cell death and an associated regenerative response, which may contribute ultimately to the carcinogenicity profile of some chemicals. A classic carcinogen of this subcategory is chloroform.

Some epigenetic agents work through receptors. The classic example of this subcategory is the aryl hydrocarbon receptor (AhR) (Whysner et al., 1996; Nebert et al., 2000; Ueda et al., 2002; Kodama et al., 2004; Moennikes et al., 2004; Klaunig and Kamendulis, 2008; Klaunig, 2013). Chemical agents that have been associated with the development of tumors are polybrominated biphenyls, polychlorinated biphenyls, and 2,3,7,8-tetrachlorodibenzodioxin (TCDD). The mechanism of action here is that chemical agents bind to the AhR and then this ligand–AhR complex moves into the nucleus, where it dimerizes with the Ah receptor nuclear translocator (ARNT). AhR/ARNT in turn now binds to aryl hydrocarbon response elements (ARE) (also known as xenobiotic response elements (XRE) or dioxin response elements (DRE)). There are AhR/ARNT-dependent genes that are associated with various P450 enzymes and other phase I and phase II enzymes that are involved in metabolic activation and detoxification. It is thought that TCDD works chiefly as a cancer promoter, promoting the carcinogenicity initiated by other compounds. Other proposed mechanisms for TCDD-induced

carcinogenesis include, but are not limited to, the generation of ROS, altered signal transduction, or the disruption of endocrine homeostasis. Another receptor proposed to be involved in the process of carcinogenesis is the constitutive androstane receptor (CAR). CAR is a member of the nuclear receptor family, and the most commonly studied chemical agent here is phenobarbital. It is widely known that phenobarbital induces the activity of P450 enzymes, especially CYP2B. While phenobarbital has many other actions that support the development of cancer, it is argued that the mechanism of action through CAR may well be the causation of oxidative stress.

The protein $\alpha_2\mu$ -globulin is synthesized in the liver of the male rat and filtered through the glomerulus and only partly excreted in the urine (Melnick et al., 1996; Klaunig and Kamendulis, 2008; Klaunig, 2013). Reabsorption of the protein occurs in the P2 segment of the proximal tubule, where it is hydrolyzed into constituent amino acids. Some chemicals, such as D-limonene, have the ability to bind to $\alpha_2\mu$ -globulin and prevent its catabolism. Accordingly, $\alpha_2\mu$ -globulin accumulates in renal lysosomes. This accumulation results in the dysfunction of the organelle with the release of destructive enzymes and the development of cell necrosis. As tubular cells are wiped out, regeneration tries to take place in order to regain lost organ function. Ultimately the mechanism of carcinogenesis here might well be that of cytotoxicity.

Cell-to-cell communication, although not typically classified as an epigenetic contributor to the process of carcinogenesis, is still nonetheless an important contributor to the development of cancer (Pitts, 1994; Colombo et al., 1995; White and Paul, 1999; Kelsell et al., 2001; Willecke et al., 2002; Leithe et al., 2006; Klaunig and Kamendulis, 2008; Klaunig, 2013). While there are a variety of ways by which cells can communicate with each other, for purposes of this discussion we will focus on gap junctions. Gap junctions are intercellular plasma membrane structures that allow the direct exchange of ions and small molecules through channels between adjacent cells. Gap junction channels are composed of connexin. Connexins play very important roles in the regulation of cell growth and differentiation. It has been shown that cancer cells usually have a downregulated expression of gap junctions, and evidence supports the claim that loss of gap junctional intercellular communication is an important step in the process of carcinogenesis.

Supporting this hypothesis is the observation that the renewed expression of connexins in cancer cells causes a reduction in the rate of tumor growth and a return to a pattern of normal cell growth control. Various mechanisms are involved in the loss of gap junctions in neoplastic cells, ranging from the loss of connexin gene transcription to the loss or alteration of function of connexin proteins. The details of actual mechanisms involved in the downregulation of connexins in carcinogenesis remain to be elucidated, and an understanding of their actions will aid a

deeper understanding of the role of gap junctions in the process of carcinogenesis.

In the experimental evaluation of substances for carcinogenesis based on experimental results of studies in a non-human species at some relatively high dose or exposure level, an attempt is made to predict the occurrence and level of tumorigenesis in humans at much lower levels. In this chapter we will examine the assumptions involved in this undertaking and review the aspects of design and interpretation of traditional long-term (lifetime) animal carcinogenicity studies as well as some alternative short-term models. It should be noted that these are required of the majority of marketed drugs but are only performed on the minority of drugs which reach a stage of development where a marketing application is likely and to occur within three or so years.

At least in a general way, we now understand what appear to be most of the mechanisms underlying chemical- and radiation-induced carcinogenesis. The most recent regulatory summary on identified carcinogens (NIH, 2004) lists 48 agents classified as “known to be human carcinogens.” Several hundred other compounds are also described as having lesser degrees of proof. A review of these mechanisms is not germane to this chapter (readers are referred to Miller and Miller (1981a) for a good short review), but it is now clear that cancer as seen in humans is the result of a multifocal set of causes.

The mechanisms and theories of chemical carcinogenesis are:

1. Genetic (all due to some genotoxic event)
2. Epigenetic (no mutagenic event)
3. Oncogene activation
4. Two-step (induction/promotion)
5. Multistep (combination of above)

Looked at another way, the four major carcinogenic mechanisms are DNA damage, cell toxicity, cell proliferation, and oncogene activation. Any effective program to identify those drugs which have the potential to cause or increase the incidence of neoplasia in humans must effectively screen for these mechanisms (Kitchin, 1999; Williams and Iatropoulos, 2001; McGregor, 2009).

The single most important statistical consideration in the design of bioassays in the past was based on the point of view that what was being observed and evaluated was a simple quantal response (cancer occurred or it didn't) and that a sufficient number of animals needed to be used to have reasonable expectations of detecting such an effect. Though the single fact of whether or not the simple incidence of neoplastic tumors is increased due to an agent of concern is of interest, a much more complex model must now be considered. The time to tumor, patterns of tumor incidence, effects on survival rate, and age of first tumor all must now

be captured in a bioassay and included in an evaluation of the relevant risk to humans.

The task of interpreting the results of any of the animal-based bioassays must be considered from three different potential perspectives as to organ responsiveness:

- I Those organs with high animal and low human neoplasia rates
- II Those organs with high neoplasia rates in both animals and humans
- III Those organs with low animal but high human neoplasia rates

Note that not considered is the potential other case—where the neoplasia rates are low for both animals and humans. This is a very rare case and one for which our current bioassay designs probably lack sufficient power to be effective.

In group I, the use of animal cancer data obtained in the liver, kidney, forestomach, and thyroid gland is perceived by some as being hyperresponsive, too sensitive, and of limited value and utility in the animal cancer data obtained in group I organs. The liver is such a responsive and important organ in the interpretation of carcinogenesis data that the discussion of this subject area has been broken up into three chapters for human, rat, and mouse data. Peroxisome proliferation in the liver, particularly in mice, is an area of interpretive battle as in many cases the metabolism and mechanisms involved are not relevant to man.

Group II organs (mammary gland, hematopoietic, urinary bladder, oral cavity, and skin) are less of an interpretive battleground than group I organs. For group II organs, all four major mechanisms of carcinogenesis (electrophile generation, oxidation of DNA, receptor–protein interactions, and cell proliferation) are known to be important. The high cancer rates for group II organs in both experimental animals and humans may at first give us a false sense of security about how well the experimental animal models are working. As we are better able to understand the probable carcinogenic mechanism(s) in the same organ in the three species, we may find that the important differences between the three species are more numerous than we suspect. This is particularly true for receptor-based and cell-proliferation-based carcinogenic mechanism.

Animal cancer data of group III organs are the opposite of group I organs. Group III organs have low animal cancer rates and high human cancer rates. In contrast to the continuing clamor and interpretive battleground associated with group I organs, there is little debate over group III organs. Few voices have questioned the adequacy of the present-day animal bioassay to protect the public health from possible cancer risks in these group III organs. Improved efforts must be made toward the development of cancer-predictive systems or short-term tests for cancer of the prostate gland, pancreas, colon/rectum, and cervix/uterus.

Carcinogenicity bioassays are the longest and most expensive of the extensive battery of toxicology studies

required for the registration of pharmaceutical products in the United States and in other major countries. In addition, they are often the most controversial with respect to interpretation of their results. These studies are important because, as noted by the International Agency for Research on Cancer (IARC) (1987), “in the absence of adequate data on humans, it is biologically plausible and prudent to regard agents for which there is sufficient evidence of carcinogenicity in experimental animals as if they presented a carcinogenic risk to humans.”

In this chapter, we consider the major factors involved in the design, conduct, analysis, and interpretation of carcinogenicity studies as they are performed in the pharmaceutical industry.

14.5 REGULATORY REQUIREMENTS AND TIMING

The need for carcinogenicity testing of pharmaceuticals is a dual criteria: that such studies were required to support registration of a drug that was to be administered for a period of 3 months or more for FDA (in Japan and Europe this was stated to be 6 months or more) and such testing had to be completed before filing for registration in such cases. ICH guidelines (ICH, 1995, 2012) now fix this human exposure triggering period at 6 months, excluding agents given infrequently through a lifetime or for shorter periods of exposure unless there is reason for concern (such as positive findings in genotoxicity studies, structure–activity relationships suggesting carcinogenic risk, evidence of occurrence of preneoplastic lesions in repeat-dose studies, or previous demonstration of carcinogenic potential in the product class that is considered relevant to humans). Such studies are still only required to be completed before filing for registration. Most developers conduct carcinogenicity studies in parallel with phase III clinical studies.

Endogenous peptides and protein substances and their analogs are generally not required to be evaluated for carcinogenicity. There are three conditions which call the need into question however:

1. Products where there are significant differences in biological effects to the natural counterparts
2. Products where modifications lead to significant changes in structure compared to the natural substance
3. Products resulting in humans having a significant increase over the existing local or systemic concentration

ICH has also given guidance on the design, dose selection, statistical analysis, and interpretation for such studies (ICH, 1995, 1997, 2008, 2012). The FDA has also offered guidance, the most recent from (FDA, 2001) a 44-page document available online.

There has been extensive debate and consideration on the relevance and value of the traditional long-term rodent bioassays. The FDA has looked at rat and mouse studies for 282 human pharmaceuticals, resulting in the conclusion that “sufficient evidence is now available for some alternative *in vivo* carcinogenicity models to support their application as complimentary studies in combination with a single 2-year carcinogenicity study to identify trans-species tumorigens (Contrera et al., 1997).” FDA is considering a range of changes to the S1 guidance governing the conduct of rodent carcinogen studies (FDA, 2012, 2013).

The Europeans, meanwhile, have focused on the need for better care in study design, conduct, and interpretation (Spindler et al., 2000), aiming to incorporate these in the revision of the Center for Proprietary Medicinal Products (CPMP) carcinogenicity guidelines (Weaver and Brunden, 1998).

The conduct of bioassays in compliance with regulatory expectations requires that an adequate 90-day range finder study be conducted in the intended test species to set doses. Additionally, it is strongly suggested that the planned study protocols (and data supporting species/models and intended doses) be reviewed and approved by the Carcinogenicity Advisory Committee (CAC).

14.6 SPECIES AND STRAIN

Two rodent species are routinely used for carcinogenicity testing in the pharmaceutical industry—the mouse and the rat. Sprague–Dawley-derived rats are most commonly used in American pharmaceutical toxicology laboratories. However, the Hans Wistar and Fischer 344 (F344) strains are favored by some companies, while the Long Evans and CFE (Carworth) strains are rarely used (Pharmaceutical Manufacturers Association (PMA), 1988).

With respect to mice, the CD-1 is by far the most commonly used strain in the pharmaceutical industry. Other strains used less frequently are the B6C3F1, CF-1, NMRI, C57B1, Balb/c, and Swiss (Pharmaceutical Manufacturers Association (PMA), 1988; Rao et al., 1988). “Swiss” is the generic term since most currently used inbred and outbred strains were originally derived from the “Swiss” mouse.

If either the mouse or the rat is considered to be an inappropriate species for a carcinogenicity study, the hamster is usually chosen as the second species.

The use of two species in carcinogenicity studies is based on the traditional wisdom that no single species can be considered an adequate predictor of carcinogenic effects in humans. Absence of carcinogenic activity in two different species is thought to provide a greater level of confidence that a compound is “safe” for humans than data derived from a single species.

One may question this reasoning on the basis that data from two “poor predictors” may not be significantly better

than data from a single species. It is also reasonable to expect that the ability of one rodent species to predict a carcinogenic effect in a second rodent species should be at least equal to, if not better than, its ability to predict carcinogenicity in humans. The concordance between mouse and rat carcinogenicity data has been investigated, and a summary of the results is presented in the next paragraph.

A review of data from 250 chemicals found 82% concordance between results of carcinogenicity testing in the mouse and the rat (Purchase, 1980). Haseman et al. (1984a) reported a concordance of 73% for 60 compounds studies in both species. However, 30–40% of 186 National Cancer Institute (NCI) chemicals were found to be positive in one species and negative in the other (Gold et al., 1984). It is reasonable to conclude that neither rodent species will always predict the results in the other rodent species or in humans and that the use of two species will continue until we have a much better understanding of the mechanisms of carcinogenesis.

The choice of species and strain to be used in a carcinogenicity study is based on various criteria including susceptibility to tumor induction, incidence of spontaneous tumors, survival, existence of an adequate historical database, and availability.

Susceptibility to tumor induction is an important criterion. There would be little justification for doing carcinogenicity studies in an animal model that did not respond when treated with a "true" carcinogen. Ideally, the perfect species/strain would have the same susceptibility to tumor induction as the human. Unfortunately, this information is usually unavailable, and the tendency has been to choose animal models that are highly sensitive to tumor induction to minimize the probability of false negatives.

The incidence of spontaneous tumors is also an important issue. Rodent species and strains differ greatly in the incidence of various types of spontaneous tumors. The Sprague–Dawley stock, although preferred by most pharmaceutical companies, has a very high incidence of mammary tumors in aging females, which results in substantial morbidity during the second year of a carcinogenicity study. If one chooses the F344 strain, the female mammary tumor incidence will be lower, but the incidence of testicular tumors will be higher (close to 100%), than that in Sprague–Dawley rats.

A high spontaneous tumor incidence can compromise the results of a carcinogenicity study in two ways. If a compound induces tumors at a site that already has a high spontaneous tumor incidence, it may be impossible to detect an increase above the high background "noise." Conversely, if a significant increase above control levels is demonstrated, one may question the relevance of this finding to humans on the basis that the species is "highly susceptible" to tumors of this type.

The ability of a species/strain to survive for an adequate period is essential for a valid assessment of carcinogenicity. Poor survival has caused regulatory problems for

pharmaceutical companies and is, therefore, an important issue (PMA, 1988). The underlying concept is that animals should be exposed to the drug for the greater part of their normal life span to make a valid assessment of carcinogenicity. If animals on study die from causes other than drug-induced tumors, they may not have been at risk long enough for tumors to have developed. The sensitivity of the bioassay would be reduced and the probability of a false-negative result would be increased.

The availability of an adequate historical database is often cited as an important criterion for species/strain selection. Historical control data can sometimes be useful in evaluating the results of a study. Although such data are not considered equal in value to concurrent control data, they can be helpful if there is reason to believe that the concurrent control data are "atypical" for the species/strain.

Although outbred stocks (e.g., Sprague–Dawley rats and CD-1 mice) are generally favored in the pharmaceutical industry, inbred strains are also used (e.g., F344 rats and B6C3F1 mice). Inbred strains may offer greater uniformity of response, more predictable tumor incidence, and better reproducibility than outbred strains. However, their genetic homogeneity may also result in a narrower range of sensitivity to potential carcinogens than exists in random-bred animals. In addition, extrapolation of animal data to humans is the ultimate goal of carcinogenicity studies, and the human population is anything but genetically homogeneous.

The ideal species for carcinogenicity bioassays should absorb, metabolize, and excrete the compound under study exactly as humans do. Unfortunately, because of the small number of species that meet the other criteria for selection, there is limited practical utility to this important scientific concept, as applied to carcinogenicity studies.

Before concluding this discussion of species/strain selection, it may be worthwhile to take a closer look at the animals preferred by pharmaceutical companies to determine to what extent they meet the conditions described earlier. Advantages of the CD-1 mouse are (i) a good historical database including various routes of exposure, (ii) demonstrated susceptibility to induction of tumors, and (iii) relatively low spontaneous incidence of certain tumors to which other strains are highly susceptible, especially mammary and hepatic tumors. Disadvantages are (i) lack of homogeneity, (ii) relatively low survival, (iii) moderate to high incidence of spontaneous pulmonary tumors and leukemias, and (iv) high incidence of amyloidosis in important organs, including the liver, kidney, spleen, thyroid, and adrenals (Sher et al., 1982).

There has recently been a reduction in survival of Sprague–Dawley rats and rats of other strains (Food and Drug Administration (FDA), 1993). This reduction may be the result of *ad libitum* feeding, as preliminary results suggest that caloric restriction may improve survival. Leukemia appears to be the major cause of decreasing survival in the

F344 rat. The problem of reduced survival may necessitate a reevaluation of the survival requirements for carcinogenicity studies by regulatory agencies.

14.7 ANIMAL HUSBANDRY

Because of the long duration and expense of carcinogenicity studies, the care of animals used in these studies is of paramount importance. Various physical and biological factors can affect the outcome of these studies. Some important physical factors include light, temperature, relative humidity, ventilation, atmospheric conditions, noise, diet, housing, and bedding (Rao and Huff, 1990). Biological factors include bacteria and viruses that may cause infections and diseases.

The duration, intensity, and quality of light can influence many physiological responses, including tumor incidence (Greenman et al., 1984; Wiskemann et al., 1986). High light intensity may cause eye lesions, including retinal atrophy and opacities (Bellhorn, 1980; Greenman et al., 1982). Rats housed in the top row and the side columns of a rack may be the most severely affected.

The influence of light on the health of animals may be managed in several ways. The animals may be randomly assigned to their cages on a rack such that each column contains animals of a single-dose group. The location of the columns on the rack may also be randomized so that the effect of light is approximately equal for all dose groups. In addition, the cages of each column of the rack may be rotated from top to bottom when the racks are changed.

Room temperature has been shown to influence the incidence of skin tumors in mice (Weisbrode and Weiss, 1981). Changes in relative humidity may alter food and water intake (Fox, 1977). Low humidity may cause "ringtail," especially if animals are housed in wire mesh cages (Flynn, 1960).

Diets for rodents in carcinogenesis studies should ideally be nutritionally adequate while avoiding excesses of nutrients that may have adverse effects.

Types of caging and bedding have been shown to affect the incidence and latency of skin tumors in mice. In a study by DePass et al. (1986), benzo[*a*]pyrene-treated mice were housed either in stainless steel cages or polycarbonate shoebox cages with hardwood bedding. The mice housed in shoebox cages developed tumors earlier and with higher frequency than those housed in steel cages.

Housing of rats in stainless steel cages with wire mesh floors may result in decubitus ulcers on the plantar surfaces. This condition may be a significant clinical problem associated with high morbidity and may affect survival of the animals if euthanasia is performed for humane reasons. Ulcers are particularly frequent and severe in older male Sprague-Dawley rats, perhaps because of their large size and weight compared with females and rats of other strains.

Common viral infections may affect the outcome of carcinogenicity studies by altering survival or tumor incidence. Nevertheless, viral infections did not cause consistent adverse effects on survival or tumor prevalence in control F344 rats from 28 NCI-National Toxicology Program (NTP) studies, though body weights were reduced by Sendai and pneumonia viruses of mice (Rao et al., 1989). The probability of such infections can be minimized by using viral-antibody-free animals, which are readily available.

14.8 DOSE SELECTION

14.8.1 Number of Dose Levels

In the pharmaceutical industry, most carcinogenicity studies have employed at least three dose levels in addition to the controls, but four levels have occasionally been used (PMA, 1988). The use of three or four dose levels satisfies regulatory requirements (Speid et al., 1990) as well as scientific and practical considerations. If a carcinogenic response is observed, information on the nature of the dose-response relationship will be available. If excessive mortality occurs at the highest dose level, a valid assessment of carcinogenicity is still possible when there is adequate survival at the lower dose levels.

14.8.2 Number of Control Groups

Pharmaceutical companies have most frequently favored the use of two control groups of equal size (PMA, 1988). A single control group of the same size as the treated groups is also used, and, less frequently, one double-sized control group may be used. The diversity of study designs reflects the breadth of opinion among toxicologists and statisticians on this issue.

Use of two control groups has the advantage of providing an estimate of the variation in tumor incidence between two groups of animals in the absence of a drug effect. If there are no significant differences between the control groups, the data can be pooled, and the analysis is identical to that using a single, double-sized group. When significant differences occur between the control groups, one must compare the data from the drug-treated groups separately with each control group.

There will be situations in which the incidence of a tumor in one or more drug-treated groups is significantly higher than that of one control group but similar to that of the other control group. In such a situation, it is often helpful to compare the tumor incidences in the control groups to appropriate historical control data. One may often conclude that, for this tumor, one of the control groups is more "typical" than the other and should, therefore, be given more weight in interpreting the differences in tumor incidence.

In spite of its current popularity in the pharmaceutical industry, the use of two control groups is opposed by some statisticians on the grounds that a significant difference between the two groups may indicate that the study was compromised by excessive, uncontrolled variation. Haseman et al. (1986), however, analyzed tumor incidence data from 18 color additives tested in rats and mice and found that the frequency of significant pairwise differences between the two concurrent control groups did not exceed that which would be expected by chance alone.

The use of one double-sized group is sometimes preferred because it may provide a better estimate of the true control tumor incidence than that provided by a smaller group. Nevertheless, more statistical power would be obtained by assigning the additional animals equally to all dose groups rather than to the control group only, if power is a primary consideration.

14.8.3 Criteria for Dose Selection

Dose selection is one of the most important activities in the design of a toxicology study. It is especially critical in carcinogenicity studies because of their long duration. Whereas faulty dose selection in an acute or subchronic toxicity study can easily be corrected by repeating the study, this situation is much less desirable in a carcinogenicity study, especially since such problems may not become evident until the last stages of the study.

The information used for dose selection usually comes from subchronic toxicity studies, but other information about the pharmacological effects of a drug and its metabolism and pharmacokinetics may also be considered. The maximum recommended human dose (MRHD) of the drug may be an additional criterion, if this is known when the carcinogenicity studies are being designed.

For most pharmaceutical companies, the doses selected are as follows. The highest dose is selected to be the estimated maximum tolerated dose (MTD). The lowest dose is usually a small multiple (1–5 times) of the MRHD, and the middose approximates the geometric mean of the other two doses (PMA, 1988; McGregor, 2009).

The MTD is commonly estimated to be the maximum dose that can be administered for the duration of the study that will not compromise the survival of the animals by causes other than carcinogenicity. It should be defined separately for males and females. ICH (1995) states that the MTD is “that dose which is predicted to produce a minimum toxic effect over the course of the carcinogenicity study, usually predicted from the results of a 90-day study.” Factors used to define minimum toxicity include no more than a 10% decrease in body weight gain relative to controls, target organ toxicity, and/or significant alterations in clinical pathology parameters. If the MTD has been chosen appropriately, there should be no adverse effect on survival, only a

modest decrement in body weight gain and minimal overt signs of toxicity. The procedures for dose selection described earlier are generally consistent with major regulatory guidelines for carcinogenicity studies (Speid et al., 1990; Food and Drug Administration (FDA), 1993¹). There are, however, exceptions to the general approach described previously. For example, for nontoxic drugs, the difference between the high and the low doses may be many orders of magnitude, if the high dose is set at the estimated MTD and the low dose is a small multiple of the clinical dose. Some guidelines request that the low dose be no less than 10% of the high dose (Speid et al., 1990). In this situation, it may be acceptable to set the high dose at 100 times the MRHD, even if the MTD is not achieved (Speid et al., 1990). Similarly, when a drug is administered in the diet, the highest concentration should not exceed 5% of the total diet, whether or not the MTD is achieved (Japanese Ministry of Health and Welfare (MHW), 1989).

Metabolism and/or pharmacokinetic data, when available, should also be considered in the dose selection process. It is desirable that a drug not be administered at such a high dose that it is excreted in a different manner than at lower doses, such as the MRHD. Similarly, the high dose should not lead to the formation of metabolites other than those formed at lower (clinical) doses. If data show that a given dosage produces maximum plasma levels, administration of higher doses should be unnecessary. These considerations may be very useful when interpreting the results of the study or attempting to extrapolate the results to humans.

14.9 GROUP SIZE

The minimum number of animals assigned to each dose group in pharmaceutical carcinogenicity studies is 50 of each sex (PMA, 1988). Most companies, however, use more than the minimum number, and some use up to 80 animals per sex per group. The most important factor in determining group size is the need to have an adequate number of animals for a valid assessment of carcinogenic activity at the end of the study. For this reason, larger group sizes are used when the drug is administered by daily gavage because this procedure may result in accidental deaths by perforation of the esophagus or aspiration into the lungs. Larger group sizes are also used when the carcinogenicity study is combined with a chronic toxicity study in the rat. In this case, serial sacrifices are performed at 6 and 12 months to evaluate potential toxic effects of the drug.

¹Note that the *FDA Redbook* applies, strictly speaking, only to food additives. It is cited here because it is a well-known toxicology guideline routinely applied to animal pharmaceuticals to which humans may be exposed. The *Redbook* has recently been updated by the FDA (Food and Drug Administration, 2007).

In the final analysis, the sensitivity of the bioassay for detecting carcinogens is directly related to the sample size. Use of the MTD has often been justified based on the small number of animals at risk compared to the potential human population, in spite of the difficulties inherent in extrapolating effects at high doses to those expected at much lower clinical doses. A reasonable compromise may be the use of doses lower than the MTD combined with a larger group size than the 50 per sex minimum accepted by regulatory agencies.

14.10 ROUTE OF ADMINISTRATION

In the pharmaceutical industry, the two most common routes of administration are via diet and gavage (PMA, 1988). Some compounds are given by drinking water, topical (dermal) application, or injection, depending on the expected clinical exposure route, which is the primary criterion for determining the route of administration in carcinogenicity studies. When more than one clinical route is anticipated for a drug, the dietary route is often chosen for practical reasons.

Dietary administration is often preferred over gavage because it is far less labor intensive. Another advantage is that the MTD has rarely been overestimated in dietary studies, whereas it has often been overestimated in gavage studies, according to data from the NTP (Haseman, 1985). The dietary route is unsuitable for drugs that are unstable in rodent chow or unpalatable. The dietary route is also disadvantaged by the fact that dosage can only be estimated based on body weight and food intake data, in contrast with gavage by which an exact dose can be given. Disadvantages of gavage testing are the likelihood of gavage-related trauma, such as puncture of the trachea or esophagus, and possible vehicle (e.g., corn oil) effects.

When doing studies by the dietary route, the drug may be administered as a constant concentration at each dose level, or the concentration may be increased as body weight increases to maintain a constant dose on a milligram per kilogram basis. The later method allows greater control of the administered dose and avoids age- and sex-related variations in the dose received, which occur with the former method. Both methods are acceptable to regulatory agencies.

14.11 STUDY DURATION

The duration of carcinogenicity studies for both rats and mice is 2 years in most pharmaceutical laboratories (PMA, 1988). Occasionally, rat studies are extended to 30 months, while some companies terminate mouse studies at 18 months. The difference in duration between mouse and rat studies is based on the belief that rats have a longer natural life span than mice. Recent data indicate, however, that this is not the

case. The most commonly used strains, the Sprague–Dawley rat and the CD-1 mouse, have approximately equal survival at 2 years, based on industry data (PMA, 1988). The same is true for the most popular inbred strains, the F344 rat and the B6C3F1 mouse (PMA, 1988). Data from NCI studies confirm that the 2-year survival of the B6C3F1 mouse is at least equal to, if not greater than, that of the F344 rat (Cameron et al., 1985).

14.12 SURVIVAL

As stated earlier, adequate survival is of primary importance in carcinogenicity studies because animals must be exposed to a drug for the greater part of their life span to increase the probability that late-occurring tumors can be detected. Early mortality, resulting from causes other than tumors, can jeopardize the validity of a study because dead animals cannot get tumors.

In general, the sensitivity of a carcinogenicity bioassay is increased when animals survive to the end of their natural life span, because weak carcinogens may induce late-occurring tumors. The potency of a carcinogen is often inversely related to the time-to-tumor development. By analogy, as the dose of a carcinogen is reduced, the time-to-tumor occurrence is increased (Littlefield et al., 1979; DePass et al., 1986).

Why do we not allow all animals on a carcinogenicity study to live until they die a natural death if by so doing we could identify more drugs as carcinogens? In fact, the sensitivity of a bioassay may not be improved by allowing the animals to live out their natural life span because the incidence of spontaneous tumors tends to increase with age. Thus, depending on the tumor type, the ability of the bioassay to detect a drug-related increase in tumor incidence may actually decrease, rather than increase, with time. Therefore, the optimum duration of a carcinogenicity study is that which allows late-occurring tumors to be detected but does not allow the incidence of spontaneous tumors to become excessive.

Reduced survival in a carcinogenicity study may or may not be drug related. Sometimes, the MTD is exceeded, and increased mortality occurs at the highest dose level and occasionally at the middose level as well. This situation may not necessarily invalidate a study; in fact, the protocol may be amended to minimize the impact of the drug-induced mortality. For example, cessation of drug treatment may enhance the survival of the animals in the affected groups and allow previously initiated tumors to develop. As shown by Littlefield et al. (1979) in the CNTR ED01 study, liver tumors induced by 2-AAF, which appeared very late in the study, were shown to have been induced much earlier and not to require the continuous presence of the carcinogen to develop. By contrast, bladder tumors that occurred in the same study were dependent on the continued presence of the carcinogen.

Whether drug treatment is terminated or not, drug-related toxicity may also be managed by performing complete histopathology on animals in the lower-dose groups rather than on high-dose and control animals only. If there is no increase in tumor incidence at a lower dose level that is not compromised by reduced survival, the study may still be considered valid as an assessment of carcinogenicity.

When reduced survival is related to factors other than excessive toxicity, the number of animals at risk for tumor development may be inadequate, and the validity of the study may be compromised even in the absence of a drug effect on survival. Obviously, the adjustments described earlier for excessive drug-related toxicity are not relevant to this situation.

There is no unanimity of opinion among regulatory agencies as to the minimum survival required to produce a valid carcinogenicity study or as to the best approach for dealing with survival problems. Even within a single agency such as the FDA, different opinions exist on these issues. For example, the recently issued *FDA Redbook II Draft Guideline* requires that rats, mice, or hamsters be treated for 24 months. Early termination due to decreased survival is not recommended. The EEC guidelines differ in that they suggest termination of the study when survival in the control group reaches 20%, while Japanese guideline suggest termination at 25% survival in the control or low-dose groups (Speid et al., 1990). These provisions make good sense in that they do not request termination of the study when drug-related mortality may be present only at the highest dose.

14.13 END POINTS MEASURED

A carcinogenicity study is more focused than a chronic toxicity study—fewer end points are evaluated, and as such it is a simpler study. The key end points are actually few:

Pathology (limited to neoplastic and preneoplastic tissue transformations)

Body weight (to ensure that toxicity is not so great as to invalidate the assays and also that it is just sufficient to validate the assay)

Survival (key to determining when to terminate the study)

Clinical pathology (limited to evaluating the morphology of white blood cells, and usually this is actually deferred until there are indications that such data is needed)

Food consumption (actually measured to ensure that dietary administration doses are accurate)

Only pathology will be considered in detail.

The primary information for carcinogenicity evaluation is generated by pathologists. Table 14.2 lists the tissues normally collected, processed, and evaluated. These professionals, like any other group of professionals, vary in their training and experience, and these are characteristics which

TABLE 14.2 Standard Tissue List

| | | |
|---------------------|--------------------------|------------------|
| Kidney | Urinary bladder | Aorta |
| Heart | Trachea | Lungs |
| Liver | Gallbladder | Pancreas |
| Fat | Salivary gland | Spleen |
| Cervical lymph node | Mesenteric lymph node | Thymus |
| Tongue | Esophagus | Stomach |
| Duodenum | Jejunum | Ileum |
| Cecum | Colon | Mammary gland |
| Skin | Skeletal muscle | Sciatic nerve |
| Parathyroid | Thyroid | Adrenal |
| Pituitary | Prostate | Seminal vesicles |
| Testes | Epididymides | Ovaries |
| Oviducts | Uterine horns | Uterine body |
| Cervix | Vagina | Brain |
| Spinal cord | Sternum | Rib/bone |
| Eyes | Harderian glands | BM smear |
| Nares | Clitoral/preputial gland | Zymbal's gland |
| Gross lesions | | |

may influence the evaluation in a number of ways. Some of these are listed as follows:

1. Differences in terminology may be important when considering controversial lesions.
2. Lack of consistency throughout a study is likely when a pathologist has only recently become involved with rodent carcinogenicity. Training is often in a clinical situation (especially in Europe), where each animal or person is unique and there is in a rodent carcinogenicity study consisting of 500 animals.
3. Unfamiliarity with the observed lesion in a particular species may cause problems in interpretation.

Possible bias introduced by knowledge of treatment can be corrected in several ways, but the use of a two-stage process would seem to be most efficient:

1. An initial evaluation is performed with full knowledge of the animal's history, including treatment.
2. A second evaluation of specific lesions is then carried out. This should be done blind, either by the same pathologist or, preferably, by the same and a second pathologist.

Differences in evaluation between pathologists should always be discussed by them to resolve the differences; they may be due to subtle differences in diagnosis and do not indicate incompetence in one of the pathologists. It is unacceptable for study sponsors to shop around until they find a pathologist who gives—for whatever reason—the result the pathologist is looking for without giving an opportunity for

interaction with all of the other evaluators. Sometimes these diagnoses are given years apart, during which time understanding of the pathogenesis of lesions may change, and even the first pathologist may not arrive at the same conclusion as the pathologist did some years ago.

Evaluation of the data is not purely a statistical exercise. A number of important factors should be considered: (i) dose–effect relationship, (ii) a shift toward more anaplastic tumors in organs where tumors are common, (iii) earlier appearance of tumors, and (iv) presence of preneoplastic lesions.

The language used to describe the carcinogenic response has masked its complexity and presents a stumbling block to its understanding among nonhistopathologists. Benign or malignant neoplasms do not arise without some precursor change within normal tissue. An important concept in carcinogenicity evaluation is that of neoplastic progression, which was derived from studies on skin tumors (Berenblum and Shubik, 1947) and expanded to a number of other tissues (Foulds, 1969, 1975). There is, on many occasions, a far from clear distinction between hyperplastic and “benign” neoplasias and between benign and malignant neoplasias.

Hyperplasia and benign and malignant neoplasias are convenient medical terms with prognostic significance. Hyperplasia can occur either as a regenerative response to injury, with no neoplastic connotations, or as a sustained response to a carcinogenic agent. It is an increase in the number of normal cells retaining normal intercellular relationships within a tissue. This normally may break down, resulting in altered growth patterns and altered cellular differentiation—a condition which may be described as atypical hyperplasia or presumptively as preneoplastic lesions. Possible sequelae to hyperplasia are (i) persistence without qualitative change in either structure or behavior; (ii) permanent regression; (iii) regression, with later reappearance; and (iv) progression to develop new characteristics

indicating increased probability of malignancy. The last of these is the least likely to occur in experimental multistage models, such as in mouse skin or rat liver, where large numbers of hyperplastic lesions may occur, but notably fewer carcinomas develop from them.

Benign neoplasms in most rodent tissues apparently arise in hyperplastic foci, for example, squamous cell papillomas of the skin and forestomach. Furthermore, these papillomas seldom demonstrate autonomous growth and even fewer progress to squamous cell carcinomas (Burns et al., 1976; Colburn, 1980). This decisive progression to carcinoma, when it occurs, provides powerful evidence for the multistage theory of carcinogenesis: the new, malignant cells arising as a focus within the papilloma or even in an area of hyperplasia, since the papilloma is not a necessary intermediate stage. In other organs, benign neoplasia is usually characterized by well-differentiated cell morphology, a fairly uniform growth pattern, clear demarcation from surrounding tissues, and no evidence of invasion. The progression toward malignancy involves anaplasia (loss of differentiation) and pleomorphism (variety of phenotypic characteristics within the neoplasm). These changes may be focal in an otherwise benign neoplasm and may vary in degree and extent. Evidence of invasion of the surrounding tissues or of metastasis is not an essential characteristic of malignancy, although their presence strengthens the diagnosis.

The grouping together of certain tumor types can aid statistical analysis, but it must be done carefully, with full appreciation of the biology and whatever is known of the pathogenesis of the lesions. Grouping for analysis of all animals showing neoplasia, irrespective of the tumor type, is inappropriate because the incidence in most treatment control groups can be very high and, in US NTP studies, approaches 100% in rats and 50–70% in mice (Table 14.3).

There may be similar incidences of tumors in aging people, but the real prevalence of tumors in human populations

TABLE 14.3 Tumor-Bearing Animals in Control Groups from Rodent Studies

| Control Animals for 2-Year NTP Bioassay | Number of Animals | % with Tumors | | |
|---|-------------------|---------------|--------|-------|
| | | Malignant | Benign | Total |
| B6C3F1 mice | | | | |
| Male | 1692 | 42 | 35 | 64 |
| Female | 1689 | 45 | 33 | 64 |
| F344 rats | | | | |
| Male | 1596 | 55 | 95 | 98 |
| Female | 1643 | 38 | 76 | 88 |
| Osborne–Mendel rats | | | | |
| Male | 50 | 26 | 68 | 78 |
| Female | 50 | 12 | 80 | 88 |
| Sprague–Dawley rats | | | | |
| Male | 56 | 9 | 36 | 39 |
| Female | 56 | 30 | 68 | 79 |

Haseman, unpublished summary of US NTP data.

is uncertain. In the United States, where autopsies are uncommon, over one-third reveal previously undiagnosed cancers when they are conducted (Silverberg, 1984). A single type of neoplasm, renal adenoma, is present in 15–20% of all adult kidneys (Holm-Nielsen and Olson, 1988), although it is unclear whether these 2–6 mm foci of proliferating tubular and papillary epithelium represent small carcinomas or benign precursors of renal cell carcinomas. Irrespective of the significance of these lesions in human pathology, the presence of similar foci in a rodent carcinogenicity experiment would trigger the recording of renal tumor-bearing animals and, hence, their consideration in the statistical and pathological evaluation processes. Evaluation is further complicated by the increased background incidences of tumors as animals get older.

The independent analysis of every different diagnosis in rodent studies would also mask significant effects in many cases while enhancing them in others. Benign and malignant neoplasms of a particular histogenesis are often grouped because the one is seen as a progression from the other. However, this grouping may result in a nonsignificant difference from the controls because there has been an acceleration of progression toward malignancy, the incidence of benign neoplasms decreasing while the malignant neoplasms increasing. Guidelines are available for “lumping” or “splitting” tumor types, but in using them, the basis for the classification of neoplastic lesions should be clarified, especially when data generated over several or many years are coupled, since diagnostic criteria and ideas regarding tumor histogenesis may have changed. Reliance on tabulated results alone can lead to serious misinterpretation by those not closely connected with a particular study. For this very important reason, the pathology and toxicology narrative should be full and clear. If it is not, then there will always be doubts about future interpretations, even if these doubts are not, in reality, justified.

14.14 TRANSGENIC MOUSE MODELS

Since the early 1970s, the standard for adequate evaluation of the carcinogenic potential of a candidate pharmaceutical has been the conduct of lifetime, high-dose assays in two species—almost always the rat and the mouse.

The relevance (and return on investment) for the bioassays preformed in mice have been questioned for some time. In 1997, ICH opened the possibility for the substitution of some form of short- or medium-term mouse test as an alternative for the traditional lifetime mouse bioassay. FDA has subsequently stated that it would accept “validated” forms of a set of medium-term mouse studies based on transgenic models, and significant effort has since gone into such validation.

The huge advances made in molecular biology since the late 1980s have provided the possibility of approaches to evaluating chemicals and potential drugs for carcinogenic potential in approaches which are different and less expensive and which take a shorter period of time than traditional long-term bioassays. This work has also been stimulated by dissatisfaction with the performance of traditional test systems.

The traditional long-term bioassays use highly inbred animals, developed with the goal of reducing the variability in background tumor incidences as a means of increasing the statistical sensitivity of the bioassays. This inbreeding has led to narrowing of the allelic pool in the strains of animals that are currently used for testing, as opposed to the wild-type populations (of humans) that the tests are intended to protect (Festing, 1979). Transgenic models should serve to improve the identification of carcinogens by providing the gene-specific mechanistic data, minimizing the influence of spontaneous tumors and strain-specific effects, and reducing time required, as well as cost and animal usage (Eastin et al., 1998).

As it has become possible to transfer new or engineered genes to the germlines of mammals, the results have been transgenic mice that can be used in shorter-term *in vivo* assays for carcinogenicity and that are also useful for research into the characterization of genotoxic events and mechanisms in carcinogenesis. By coupling reporter phenotypes (such as papillomas in the Tg.AC mouse), the task of “reading” results in test animals is made much less complex.

There are four transgenic mouse models that have been broadly evaluated—the TSPp53^{+/−}, the Tg.AC, the Hras2, and the XPA^{−/−}. Each of these has its own characteristics. Each of these merits some consideration. They are each made by either zygote injection or specific gene targeting in embryonic cells (French et al., 1999; McAnulty, 2000).

14.14.1 The Tg.AC Mouse Model

This was the earliest of the models to be developed, and its use in mouse skin carcinogenicity studies was first reported in 1990. The mice have four copies of the v-H-ras oncogene in tandem on chromosome 11, and the transgene is fused with a fetal ζ -globin gene which acts as a promoter. The transgene codes for a switch protein which is permanently “on,” and this results in the mice having genetically initiated skin. The application of tumor promoters to the surface of the skin causes the rapid induction of pedunculate papillomas that arise from the follicular epithelium. This is despite the fact that the transgene is not expressed in the skin, although it is present in the papillomas that form, and also in the focal follicular hyperplastic areas that are the precursors to the papillomas. In about 40% of the mice, the papillomas become malignant skin tumors—mainly squamous cell carcinomas and sarcomas.

The first assessments of this model as an alternative to traditional carcinogenicity studies were performed by the US NIEHS and NTP, and the results with over 40 chemicals have been published. The majority of studies were performed by skin painting, regardless of whether the product was a dermal or systemic carcinogen. However, a good correlation was found with the known carcinogenicity of the test compounds, and both mutagenic and nonmutagenic were identified. It was found that great care had to be taken with the skin, because damage could also induce papillomas, which means that these animals cannot be identified using transponder chips. This sensitivity may also explain some of the false-positive results that have occurred with resorcinol and rotenone. Of more concern is that there have been false negatives with known carcinogens, namely, ethyl acrylate and *N*-methyl-*o*-acrylamide. The model was designed for use in the context of the two-stage model of carcinogenesis with the underlying mechanistic pathway involving specific transcription factors, hypomethylation, and cell-specific expression of the results; along with the p53, this model has seen the widest use and evaluation (in terms of number of agents evaluated) so far. The carrier mouse strain employed, the FVB/N, is not commonly employed in toxicology and is prone to sound-induced seizures. It may be that the dermal route is not suitable for all systemic carcinogens, and this is the reason that in the ILSI program, both the dermal and systemic routes are being investigated in this model.

Another problem with this model was the occurrence of a nonresponder genotype to positive control agents. This was found to be attributable to a rearrangement of the ζ -globin promoter region, but it is claimed that this problem has been resolved. However, this has considerably delayed the ILSI studies with this model, but all data should be available in time for the November meeting. It is already clear that the model gives a robust response to the positive control agent, 12-*o*-tetradecanoylphorbol-13-acetate (TPA).

14.14.2 The Tg.rasH2 Mouse Model

This model was developed at CIEA in Japan, and the first information about the mouse was published in 1990. The mice have five or six copies of the human H-ras proto-oncogene inserted in tandem into their genome surrounded by their own promoter and enhancer regions. This transgene has been very stable, with no loss of responsiveness since the model was developed. The transgene codes for a molecular switch protein in the same way as the previous model, but the transgene is expressed in all organs and tissues. Thus the response end point is not primarily dermal.

The initial studies with this model revealed a rapid appearance of forestomach papillomas with *N*-methyl-*N*-nitrosourea (MNU), and this compound has already been used as the positive control agent in subsequent studies with this strain. A study duration of 6 months is sufficient to

obtain a positive response, and longer periods should be avoided because the mice start to develop various spontaneous tumors, such as splenic hemangiosarcomas, forestomach and skin papillomas, lung and Harderian gland adenocarcinomas, and lymphomas. It has a high level of constitutive expression and some spontaneous tumors even when the animals are younger. It is, however, very responsive to carcinogens—one gets a rapid onset after exposure and a higher response incidence than with the other models. The underlying mechanism is still not certain.

A large number of studies have been run in this strain in Japan in advance of the ILSI program. The model is sensitive to both mutagenic and nonmutagenic carcinogens, although cyclophosphamide and furfural have given equivocal results in each category, respectively. The majority of noncarcinogens have also been identified correctly, although again, there are a small number of compounds that have given equivocal results. In the ILSI program, 24 of the 25 studies will have been completed in time for the November meeting, and the final studies will be completed during 2001.

14.14.3 The P53^{+/-} Mouse Model

The TSP p53^{+/-} (hereafter referred to as the p53—the designation of the tumor suppressor gene involved) is a heterozygous knockout with (up to seven or so months of age) a low spontaneous tumor incidence. It is responsive to the genotoxic carcinogens by a mechanism based on the fact that many (but not all) tumors show a loss of the wild-type allele. The p53 has been extensively worked on by Tennant's group at NIEHS (Tennant et al., 1995, 1999). This model was developed in the United States and carries a hemizygous knockout of the p53 gene which was developed by integrating a mutated copy of the gene into the genome of mice. The p53 gene is known as a tumor suppressor gene, and it is the most commonly mutated gene in human malignancies. It searches for a protein transcription factor which activates multiple genes when damage to DNA strands occurs, and this in turn leads to either the arrest of the cell cycle while DNA repair occurs or to apoptosis (programmed cell death), which removes the damaged cell. The heterozygote is used because homozygotes show a very high incidence of spontaneous tumors within a few months of birth. The heterozygotes have a low background incidence of tumors up to 12 months, but during this time there is a high chance of a second mutagenic event occurring—following exposure to a carcinogen, for example—and this would result in a loss of suppressor function or an increase in transforming activity.

The initial studies with this model as an alternative in traditional carcinogenicity testing were performed at the US NIEHS, and these suggested that it was sensitive to mutagenic carcinogens such as benzene and *p*-cresidine within 6 months. Nonmutagenic carcinogens were negative in the assay, as were mutagenic noncarcinogens. However,

subsequent studies and some parts of the ILSI program have shown clear indications that a 6-month duration may be insufficient. In particular, benzene has given negative or equivocal results within 6 months, although positive results have been obtained by extending the study to 9 months. It will be very important to assess the results of the ILSI program when deciding the best study duration for this model. This is the most popular model in the United States.

14.14.4 The XPA^{-/-} Mouse Model

This was the last of the models to be developed and was created using a knockout technique after the XPA gene had been cloned. The first data were published by RIVM in the Netherlands in 1995. Both alleles of the XPA gene have been inactivated by a homologous recombination in ES cells, resulting in a homozygous deletion of the gene spanning exons three and four. The protein coded by this gene is essential for the detection and repair of DNA damage, using the NER pathway. This model only has between 2 and 5% of residual NER activity.

The initial studies at RIVM demonstrated that exposure of these mice to IV-B radiation or 7,12-dimethylbenz[a]anthracene resulted in the rapid induction of skin tumors. It was also shown that various internal tumors could be induced following oral administration of mutagenic carcinogens such as benzo[a]pyrene (B[a]P) and 2-acetylaminofluorene (2-AAF). The early studies suggested that this response could occur within six months, but further experience has indicated that a 9-month treatment period is essential in order to obtain a response with positive control agents such as B[a]P, 2-AAF, and p-cresidine.

All of the 13 studies that have been undertaken with this model were available for review at the November 2000 meeting. The model is sensitive to both UV and genotoxic carcinogens and also the some nonmutagenic carcinogens, such as DES, Wy-14,643 and cyclosporin A. There have been no false positives with noncarcinogens. Some laboratories have also investigated a double transgenic XPA^{-/-} p53^{+/-} model, and this seems to increase the sensitivity of the assay. For example, in a DES study, seven animals with metastasizing osteosarcomas were found in the double transgenic group, compared with one in the XPA group and none among the wild-type animals. There remains concern (as with any new model) that these models may be overly sensitive or (put another way) that the relevance of positive findings to risk in humans may not be clear. The results of the ILSI/HESI workshop seem to minimize these concerns.

It is generally proposed that while such models can improve the identification of carcinogens in three ways (providing gene-specific mechanistic data, minimizing the influence of spontaneous tumors and strain-specific effects, and reducing the time, cost, and animal usage involved),

they have two potential uses in pharmaceutical development. These are either in lieu of the mouse 2-year cancer bioassay or in subchronic toxicity assessments prior to making a decision to commit to a pair of 2-year carcinogenicity bioassays.

As performance data has become available on these strains, ICH (1997) has incorporated their use into pharmaceutical testing guidelines in lieu of the second rodent species tests (i.e., to replace the long-term mouse bioassay when the traditional rat study has been performed). The FDA has stated that they would accept such studies when “performed in a validated code.” In fact, CBER has accepted such studies as a sole carcinogenicity bioassay in some cases where there were negative traditional genotoxicity data and strong evidence of a lack of a mechanistic basis for concern.

A joint ILSI and HESI validation program has been completed, looking at the results of the four prime candidate models in identifying carcinogens as compared to the results of traditional long-term rodent bioassays. This validation program involved 51 different laboratories and imposed protocol standards to allow comparison of results. Three dose levels were studied per chemical, with 15 males and 15 females being used for each dose group. A vehicle and high-dose control in wild-type animals was also included, with information from NTP bioassays and 4-week range-finding assays being used to help set doses. Animals were dosed for 26 weeks. The issues in and coming out of these validation programs bear consideration (Tennant et al., 1999):

- Is the proper comparator data for evaluating performance human or rodent bioassay data? It should be kept in mind that there are sets of rodent bioassay data (particularly those involving liver tumors in mice) that are widely accepted as irrelevant in the prediction of human risk.
- How will the data from these assays be incorporated into any weight-of-evidence (WOE) approach to assessing human health risk?
- What additional mechanistic research needs to be undertaken to improve our understanding of the proper incorporation and best use of the data from these assays?
- How can the results of current validation efforts be best utilized in the timely evaluation of the next generation of assays?
- Given that, at least under some conditions, assays using these models tend to “blow up” (have high spontaneous tumor rates) once the animals are more than 8 or 9 months of age, how critical are age and other not currently apprehended factors to optimizing both sensitivity and specificity?
- How wide and unconditional will FDA (and other regulatory bodies) acceptance be of these models in lieu of the traditional 2-year mouse bioassay?

14.15 INTERPRETATION OF RESULTS: CRITERIA FOR A POSITIVE RESULT

There are three generally accepted criteria for a positive result in a carcinogenicity study. The first two are derived directly from the results of the statistical analysis: (i) a statistically significant increase in the incidence of a common tumor and (ii) a statistically significant reduction in the time-to-tumor development. The third criterion is the occurrence of very rare tumors, that is, those not normally seen in control animals, even if the incidence is not statistically significant. Table 14.4 presents an evaluation matrix for these factors.

14.16 STATISTICAL ANALYSIS

Irrespective of the specific protocols used, all carcinogenicity studies end with a statistical comparison of tumor proportions between treated and control groups. This analysis is necessary because the control incidence of most tumor types is rarely zero. In the unlikely case that a type of tumor is found in treated animals but not in concurrent or appropriate historical controls, it is reasonable to conclude that the tumor is drug related without statistical analysis.

Most pharmaceutical companies analyze tumor data using mortality-adjusted methods (PMA, 1988). Peto/IARC methodology is most commonly used, perhaps because this method is currently favored by the FDA (Peto et al., 1980). The use of life-table methods is most appropriate for "lethal" tumors, that is, those that cause the death of the animals.

Various statistical methods are available for analyzing the incidence of the lethal and nonlethal tumors (e.g., Gart et al., 1979, 1986; Chu et al., 1981; Dinse and Lagakos, 1983; McKnight, 1988; Portier and Bailer, 1989; Gaylor and Kodell, 2001). These methods are especially useful when there are drug-related differences in mortality rates. When there is no drug effect on survival, unadjusted methods will generally give the same results.

As a general approach, most pharmaceutical statisticians begin by testing for the presence of a dose-related trend in tumor proportions. If the trend test is significant, that is, the *p* value is less than or equal to 0.05, pairwise comparisons are performed between the treated and control groups. Trend and pairwise analyses may be adjusted for mortality as stated earlier or performed without mortality adjustment using such simple methods as chi-square or Fisher's exact tests.

Although in most cases the use of trend tests is appropriate since most biological responses are dose related, there are exceptions to this rule. Certain drugs, especially those with hormonal activity, may not produce classical dose responses and may even induce inverse dose-response phenomena. In these cases, a pairwise comparison may be appropriate in the absence of a significant positive trend.

Most (70%) pharmaceutical companies use one-tailed comparisons, and a substantial number use two-tailed methods (PMA, 1988). Since regulatory agencies are primarily interested in identifying carcinogenic drugs, as opposed to those that inhibit carcinogenesis, the use of one-tailed tests is generally considered more appropriate. Some companies prefer two-tailed comparisons because, in the absence of a true carcinogenic effect, there is an equal

TABLE 14.4 Interpretation of the Analysis of Tumor Incidence and Survival Analysis (Life Table)

| Outcome Type | Tumor Association with Treatment ^a | Mortality Association with Treatment | Interpretation |
|--------------|---|--------------------------------------|---|
| A | – | + | Unadjusted test ^b may underestimate tumorigenicity of treatment |
| B | + | + | Unadjusted test gives valid picture of tumorigenicity of treatment |
| C | + | – | Tumors found in treated groups may reflect longer survival of treated groups. Time-adjusted analysis is indicated |
| D | – | + | Apparent negative findings on tumors may be due to the shorter survival in treated groups. Time-adjusted analysis and/or a retest at lower doses is indicated |
| E | – | 0 | Unadjusted test gives a valid picture of the possible tumor-preventive capacity of the treatment |
| F | – | – | Unadjusted test may underestimate the possible tumor-preventive capacity of the treatment |
| G | 0 | + | High mortality in treated groups may lead to unadjusted test missing a possible tumorigen. Adjusted analysis and/or retest at lower doses is indicated |
| H | 0 | 0 | Unadjusted test gives a valid picture of lack of association with treatment |
| I | 0 | – | Longer survival in treated groups may mask tumor-preventive capacity of treatment |

^a+ = Yes, – = no, and 0 = no bearing on discussion.

^bThe unadjusted test referred to here is a contingency table type of analysis of incidence, such as Fisher's exact test.

probability of seeing significant decreases as well as significant increases by chance alone.

One of the most important statistical issues in the analysis of carcinogenicity data is the frequency of “false positives,” or type I errors. Because of the multiplicity of tumor sites examined and the number of tests employed, there is concern that noncarcinogenic drugs may be erroneously declared carcinogens. If a $p < 0.05$ increase in tumor incidence is automatically regarded as a biologically meaningful result, then the false-positive rate may be as high as 47–50% (Haseman et al., 1986).

Several statistical procedures designed to correct for the multiplicity of significance tests have been published (and reviewed by Haseman (1990)). One approach to the problem of multiple tumor site/type testing is a procedure attributed to Tukey by Mantel (1980). This method is used to adjust a calculated p value based on the number of tumor types/sites for which there are a minimum number of tumors in the particular study. The reasoning here is that, for tumor sites, the number of tumors found is so small that it is impossible to obtain a significant result for that tumor site no matter how the tumors might have been distributed among the dose groups. Only those sites for which a minimum number of tumors are present can contribute to the false-positive rate for a particular study.

A method proposed by Schweder and Spjotvoll (1982) is based on a plot of the cumulative distribution of observed p values. Farrar and Crump (1988) have published a statistical procedure designed not only to control the probability of false-positive findings but also to combine the probabilities of a carcinogenic effect across tumor sites, sexes, and species.

Another approach to controlling the false-positive rate in carcinogenicity studies was proposed by Haseman (1983). Under this “rule,” a compound would be declared a carcinogen if it produced an increase significant at the 1% level in a common tumor or an increase significant at the 5% level in a rear tumor. A rare neoplasm was defined as a neoplasm that occurred with a frequency of less than 1% in control animals. The overall false-positive rate associated with this decision rule was found to be not more than 7–8%, based on control tumor incidences from NTP studies in rats and mice. This false-positive rate compares favorably with the expected rate of 5%, which is the probability at which one would erroneously conclude that a compound was a carcinogen. The method is notable for its simplicity and deserves serious consideration by pharmaceutical statisticians and toxicologists. Without resorting to sophisticated mathematics, this method recognizes the fact that tumors differ in their spontaneous frequencies and, therefore, in their contribution to the overall false-positive rates in the carcinogenicity studies. False-positive results are much less likely to occur at tissue sites with low spontaneous tumor incidences than at those with high frequencies.

As a final point that has special relevance to pharmaceutical carcinogenicity studies, one may question whether the corrections for multiple comparisons and their effect on the overall false-positive rate are appropriate for all tumor types. For example, if a compound is known to bind to receptors and produce pharmacological effects in a certain organ, is it justified to arbitrarily correct the calculated p value for the incidence of tumors in that organ, using the methods described previously? It is difficult to justify such a correction considering that the basis for correcting the calculated p value is that the true probability of observing an increased incidence of tumors at any site by chance alone may be much higher than the nominal alpha level (usually 0.05). It is reasonable to expect that when a drug has known pharmacological effects on a given organ, the probability of observing an increased tumor incidence in that organ by chance alone is unlikely to be higher than the nominal 5% alpha level.

Although most pharmaceutical statisticians and toxicologists agree on the need to control the probability of false-positive results, there is no consensus as to which method is most appropriate or most acceptable to regulatory agencies. The FDA and other such agencies will accept a variety of statistical procedures but will often reanalyze the data and draw their own conclusions based on their analyses.

The actual statistical techniques used to evaluate the results of carcinogenicity bioassays basically utilize four sets of techniques, three of which have been presented earlier in this book. These methods are:

1. Exact tests
2. Trend tests
3. Life tables (such as log-rank techniques)
4. Peto analysis

These are then integrated into the decision-making schemes discussed earlier in this chapter. The methods themselves and alternatives are discussed elsewhere in detail (Chow and Liu, 1998; Gad, 2005).

14.16.1 Exact Tests

The basic forms of these (the Fisher exact test and chi-square) have previously been presented, and the reader should review these. Carcinogenicity assays are, of course, conducted at doses that are at least near those that will compromise mortality. As a consequence, one generally encounters competing toxicity producing differential mortality during such a study. Also, often, particularly with certain agricultural chemicals, latency of spontaneous tumors in rodents may shorten as a confounded effect of treatment with toxicity. Because of such happenings, simple tests on proportions, such as χ^2 and Fisher–Irwin exact tests on contingency tables, may not produce optimal evaluation of

the incidence data. In many cases however, statisticians still use some of these tests as methods of preliminary evaluation. These are unadjusted methods without regard for the mortality patterns in a study. Failure to take into account mortality patterns in a study sometimes causes serious flaws in interpretation of the results. The numbers at risk are generally the numbers of animals histopathologically examined for specific tissues.

Some gross adjustments on the numbers at risk can be made by eliminating early deaths or sacrifices by justifying that those animals were not at risk to have developed the particular tumor in question. Unless there is dramatic change in tumor prevalence distribution over time, the gross adjusted method provides fairly reliable evidence of treatment effect, at least for nonpalpable tissue masses.

14.16.2 Trend Tests

Basic forms of the trend tests (such as that of Tarone) have previously been presented in this text.

Group comparison tests for proportions notoriously lack power. Trend tests, because of their use of prior information (dose levels), are much more powerful. Also, it is generally believed that the nature of true carcinogenicity (or toxicity for that matter) manifests itself as dose-response. Because of the previous facts, evaluation of trend takes precedence over group comparisons. In order to achieve optimal test statistics, many people use ordinal dose levels (0, 1, 2, ...) instead of

the true arithmetic dose levels to test for trend. However, such a decision should be made a priori. The example in Table 14.5 demonstrates the weakness of homogeneity tests.

As is evident from this example, often group comparison tests will fail to identify significant treatment while trend test will. Same arguments apply to survival-adjusted tests on proportions as well. In an experiment with more than one dose group ($K > 1$), the most convincing evidence for carcinogenicity is given by tumor incidence rates that increase with increasing dose. A test designed specifically to detect such dose-related trends is Tarone's (1975) trend test.

Letting $\mathbf{d} = (O, d_1, d_2, \dots, d_k)^T$ be the vector of dose levels in all $K + 1$ groups and letting $(\mathbf{O} - \mathbf{E}) = (O_0 - E_0, \dots, O_k - E_k)^T$

and $\mathbf{V} = \begin{pmatrix} V_{00} & \dots & V_{0K} \\ \vdots & & \vdots \\ V_{K0} & \dots & V_{KK} \end{pmatrix}$ contain elements as described in

the previous section but for all $K + 1$ groups, the trend sta-

tistic is given by $\chi_T^2 = \frac{[d^T (\mathbf{O} - \mathbf{E})]^2}{d^T \mathbf{V} d}$.

The statistic χ_T^2 will be large when there is evidence of a dose-related increase or decrease in the tumor incidence rates and small when there is little difference in the tumor incidence between groups or when group differences are not dose related. Under the null hypothesis of no differences between groups, χ_T^2 has approximately a chi-squared distribution with one degree of freedom.

TABLE 14.5 Trend versus Heterogeneity

| Number at Risk | Number with Tumor | Dose Level | | |
|---|--------------------------------|----------------------|--------|----------|
| 50 | 2 | 0 | | |
| 50 | 4 | 1 | | |
| 50 | 6 | 2 | | |
| 50 | 7 | 3 | | |
| <i>Cochran–Armitage Test for Trend</i> | | | | |
| | Calculated Chi-Square Subgroup | DF | Alpha | 2-Tail p |
| Trend | 3.3446 | 1 | 0.0500 | 0.0674 |
| Departure | 0.0694 | 2 | 0.0500 | 0.9659 |
| Homogeneity | 3.4141 | 3 | 0.0500 | 0.3321 |
| <i>One-Tail Tests for Trend</i> | | | | |
| Type | Probability | | | |
| Uncorrected | 0.0337 ^a | | | |
| Continuity corrected | 0.0426 ^a | | | |
| Exact | 0.0423 ^a | | | |
| <i>Multiple Pairwise Group Comparisons by Fisher–Irwin Exact Test</i> | | | | |
| Groups Compared | Alpha | One-Tail Probability | | |
| 1 versus 2 | 0.0500 | 0.33887 | | |
| 2 versus 3 | 0.0500 | 0.13433 | | |
| 1 versus 4 | 0.0500 | 0.07975 | | |

^aDirection = +.

Tarone's trend test is most powerful at detecting dose-related trends when tumor onset hazard functions are proportional to each other. For more power against other dose-related group differences, weighted versions of the statistic are also available; see Breslow (1984) or Crowley and Breslow (1984) for details.

These tests are based on the generalized logistic function (Cox, 1972). Specifically one can use Cochran–Armitage test (or its parallel, Mantel–Haenszel version) for monotonic trend as heterogeneity test.

14.16.3 Life Table and Survival Analysis

These methods are essential when there is any significant degree of mortality in a bioassay. They seek to adjust for the differences in periods of risk individual animals undergo. Life-table techniques can be used for those data where there are observable or palpable tumors. Specifically, one should use Kaplan–Meier product limit estimates from censored data graphically, Cox–Tarone binary regression (log-rank test), and Gehan–Breslow modification of Kruskal–Wallis tests (Thomas et al., 1977) on censored data.

The Kaplan–Meier estimates produce a step function for each group and are plotted over the lifetime of the animals. Planned, accidentally killed, and lost animals are censored. Moribund deaths are considered to be treatment related. A graphical representation of Kaplan–Meier estimates provides an excellent interpretation of survival-adjusted data except in the cases where the curves cross between two or more groups. When the curves cross and change direction, no meaningful interpretation of the data can be made by any statistical method because proportional odds characteristic is totally lost over time. This would be a rare case where treatment initially produces more tumor or death and then, due to repair or other mechanisms, becomes beneficial.

In Cox–Tarone binary regression (Tarone, 1975; Thomas et al., 1977), censored survival and tumor incidence data are expressed in a logistic model in dose over time. The log-rank test (Peto, 1974), tests based on the Weibull distribution, and Mantel–Haenszel (Mantel and Haenszel, 1952) test are very similar to this test when there are no covariates or stratifying variables in the design. The logistic regression-based Cox–Tarone test is preferable because one can easily incorporate covariates and stratifying variables which one cannot in the IARC methods.

The Gehan–Breslow modification of the Kruskal–Wallis test is a nonparametric test on censored observations. It assigns more weight to early incidences compared to the Cox–Tarone test.

For survival-adjusted tests on proportions, as mentioned earlier, in the case of survival-adjusted analyses, instead of having a single $2 \times k$ table, one has a series of

such $2 \times k$ tables across the entire lifetime of the study. The numbers at risk for such analyses will depend on the type of tumor one is dealing with. These are shown as follows:

1. Palpable or lethal tumors: Number at risk at time t = number of animals surviving at the end of time $t - 1$.
2. Incidental tumors: The number at risk at time t = number of animals that either died or sacrificed whose particular tissue was examined histopathologically.

The methods of analyzing the incidences, once the appropriate numbers at risk are assigned for these tumors, are rather similar, either binary regression based or by pooling evidence from individual tables (Gart et al., 1986).

14.16.4 Peto Analysis

The Peto method of analysis of bioassay tumor data is based on careful classification of tumors into five different categories, as defined by IARC:

1. Definitely incidental
2. Probably incidental

Comment Combine (1) and (2)

3. Probably lethal
4. Definitely lethal

Comment These categories may be combined into one (otherwise it requires a careful cause of death determination).

5. Mortality independent (such as mammary, skin, and other observable or superficial tumors)

14.16.4.1 Interval Selection for Occult (Internal Organ) Tumors

1. FDA: 0–50, 51–80, 81–104 weeks, interim sacrifice, terminal sacrifice
2. NTP: 0–52, 53–78, 79–92, 93–104 weeks, interim sacrifice, terminal sacrifice
3. IARC: *Ad hoc* selection method (Peto et al., 1980)

Comment Any of the above may be used. Problems with IARC selection method include two sexes and two or more strains will have different intervals for the same compound. Different interval selection methods will produce different statistical significance levels. This may produce bias and requires an isotonic tumor prevalence for ready analysis.

14.16.4.2 Logistic Regression Method for Occult (Internal Organ) Tumors (Dinse, 1985) Tumor prevalence is modeled as logistic function of dose and polynomial in age.

Comment Logistic tumor prevalence method is unbiased. It requires maximum likelihood estimation and allows for covariates and stratifying variables. It may be time consuming and have convergence problem with sparse tables (low tumor incidences) and clustering of tumors.

14.16.5 Methods to Be Avoided

The following methods and practices should be avoided in evaluation of carcinogenicity:

1. Use of only the animals surviving after 1 year in the study.
2. Use of a two-strata approach: Separate analyses for animals killed during the first year of the study and the ones thereafter.
3. Exclusion of all animals in the study that died on test and analyze only the animals that are sacrificed at the end of the study.
4. Exclusion of interim sacrifice animals from statistical analyses.
5. Evaluation of number of tumors of all sites as opposed to the number of animals with tumors for specific sites of specific organs.

Another issue is subjectivity in slide reading by most pathologists who do not want to read them in a coded fashion whereby they will not know the dose group an animal is coming from. This is not under statisticians' control, but they should be aware of that in any case.

Often a chemical being tested is both toxic and potentially carcinogenic. When competing toxicity causes extreme differences in mortality or there is clustering effect in tumor prevalence in a very short interval of time, none of the adjusted methods works. One then must use biological intuition to evaluate the tumor data.

Use of historical control incidence data for statistical evaluation is controversial. There are too many sources of variation in these data. For example, different pathologists use different criteria for categorizing tumors (in fact, the same pathologist may change opinion over time); there is laboratory-to-laboratory variation; there may be genetic drift over time; location of suppliers may make a difference; and finally, these data are not part of the randomized concurrent control. Regulatory agencies and pathologists generally use these data for qualitative evaluation. My personal view is that is where they belong.

14.16.6 Use of Historical Controls

When the study is over, the data analyzed, and the *p* values corrected, as appropriate, one may find that one or more

tumor types increased in drug-treated groups relative to concurrent controls. Although the FDA and other regulatory agencies play down the importance of historical control data, it is common practice in the pharmaceutical industry to use historical data in the interpretation of tumor findings. The first and most appropriate comparison of a treated group is with concurrent control group(s), but it is of interest to see how tumor incidences in the treated groups compare with the historical incidence, and such a comparison is an accepted practice in toxicology and biostatistics (Gart et al., 1979; Hajian, 1983; Haseman et al., 1984b). A treated group may have a tumor incidence significantly higher than that of the concurrent control group(s), but comparable to or lower than the historical incidence. Occasionally, a small number of tumors may be found in a treated group, and the incidence may be significant because of the absence of this tumor in the concurrent controls. Review of appropriate historical control data may reveal that the low tumor incidence in the treated group is within the "expected" range for this tumor.

The role of historical control data in interpreting carcinogenicity findings depends on the "quality" of the historical data. Ideally, the data should be derived from animals of the same age, sex, strain, and supplier and housed in the same facility, and the pathology examinations should have been performed by the same pathologist or using the same pathological criteria for diagnosis. Since genetic drift occurs even in animals of a given strain and supplier, recent data are more useful than older data. The value of historical control data is directly proportional to the extent to which these conditions are fulfilled.

Although methods are available for including historical control data in the formal statistical analysis (Tarone, 1982; Dempster et al., 1983), this is usually not done and for good reason. The heterogeneity of historical data requires that they be used qualitatively and selectively to aid in the final interpretation of the data, after completion of the formal statistical analysis. Table 14.6 presents a summary of background tumor incidences for the most commonly employed rodent strains.

14.16.7 Relevance to Humans

After statistical analyses have been performed and historical data consulted, the final interpretation may be that a drug appears to cause tumors at one or more tissue sites in the mouse or the rat. But what does this mean for the species to which the drug will be administered, namely, the human? Extrapolation of rodent carcinogenicity data to humans remains one of the greatest challenges of modern toxicology. There is no simple formula, and each case must be evaluated on its own merits. Very generally speaking, the FDA and other major regulatory agencies consider compounds that are tumorigenic in one or more animal species to be "suspect" tumorigens in humans. The actual impact of this

TABLE 14.6 Comparative Percent Incidence of Pertinent Neoplasia in Different Strains of Rats and Mice (104 Weeks Old)

| Types of Neoplasia | F344 Rats | | S-D Rats | | Wistar Rats | | B6C3F ₁ Mice | | CD-1 Mice | |
|----------------------------|-----------|---------|----------|---------|-------------|---------|-------------------------|---------|-----------|---------|
| | Males | Females | Males | Females | Males | Females | Males | Females | Males | Females |
| Hepatocellular adenoma | 4 | <1 | 5 | <1 | 1 | 2 | 29 | 30 | 26 | 5 |
| Hepatocellular carcinoma | 2 | 0 | 2 | 0 | <1 | <1 | 26 | 16 | 10 | 1 |
| Pancreatic islet adenoma | 12 | 2 | 8 | 9 | 4 | 2 | 2 | 0 | <1 | <1 |
| Pancreatic islet carcinoma | 3 | 0 | <1 | 5 | <1 | <1 | 0 | 0 | 0 | 0 |
| Pancreatic acinar adenoma | 6 | 0 | 1 | 0 | 13 | <1 | 2 | 0 | <1 | 0 |
| Pheochromocytoma | 21 | 4 | 23 | 5 | 10 | 2 | 0 | 2 | <1 | <1 |
| Adrenocortical adenoma | 0 | 2 | 3 | 0 | 8 | 9 | <1 | 0 | 1 | <1 |
| Pituitary adenoma | 49 | 42 | 62 | 85 | 34 | 55 | 2 | 8 | 0 | 5 |
| Thyroid C-cell adenoma | 17 | 8 | 7 | 6 | 6 | 8 | 0 | 0 | 0 | 0 |
| Thyroid follicular adenoma | 0 | 0 | 4 | 2 | 2 | 1 | 2 | 6 | 1 | <1 |
| Mammary gland fibroadenoma | 4 | 57 | 2 | 54 | 3 | 36 | 0 | 0 | <1 | 1 |
| Mammary gland carcinoma | 0 | 4 | <1 | 26 | 1 | 13 | 0 | 0 | 0 | 6 |
| Skin fibroma | 10 | 2 | 2 | <1 | 5 | 1 | 1 | 2 | <1 | <1 |
| Skin papilloma | 6 | 0 | 2 | 0 | 2 | <1 | 0 | 0 | <1 | 0 |
| Pulmonary adenoma | 4 | 4 | <1 | <1 | <1 | 0 | 22 | 6 | 15 | 15 |
| Preputial gland neoplasia | 10 | NA | >1 | NA | <1 | NA | <1 | NA | <1 | NA |
| Leydig cell neoplasia | 89 | NA | 7 | NA | 11 | NA | 0 | NA | 1 | NA |
| Clitoral gland neoplasia | NA | 14 | NA | <1 | NA | <1 | NA | <1 | NA | 0 |
| Uterine polyps | NA | 14 | NA | 6 | NA | 16 | NA | 1 | NA | <1 |
| Ovarian neoplasia | NA | 6 | NA | 1 | NA | 8 | NA | 6 | NA | 1 |
| Mononuclear cell leukemia | 62 | 42 | 0 | 0 | <1 | <1 | 0 | 0 | 2 | 2 |
| Lymphoma | 0 | 0 | 2 | 1 | 3 | 5 | 14 | 24 | 8 | 22 |
| Forestomach papilloma | 0 | 2 | <1 | <1 | 0 | <1 | 4 | 2 | <1 | <1 |
| Scrotal mesothelioma | 5 | NA | 1 | NA | 2 | NA | 0 | NA | 0 | NA |

Note: F344, Fischer 244 rats; S-D, Sprague–Dawley rats; B6C3F₁, mice (C57BL/6N+C3H/HeN)F₁; CD-1, 1CRCr: CD-1 mice; NA, nonapplicable; the average number used by species/strain/gender was in excess of 750 animals.

conclusion on the approval of a drug depends on the target population and the indication. For example, even a suspicion of carcinogenic activity may be fatal for a potential contraceptive drug intended for use in a very large population of healthy people. By contrast, clear evidence of carcinogenic activity may be overlooked in a drug being considered for use in a restricted population with a life-threatening disease.

Regardless of the target population and indication, the FDA and other agencies have, in recent years, attempted to consider the mechanism of tumor induction in rodents and its relevance for humans. If a drug is known to cause tumors in a rodent via a mechanism that does not exist in humans, the importance of the tumor findings may be markedly reduced. For example, drugs that cause tumors by a secondary hormonal mechanism shown to be inapplicable to humans may be given special consideration. It is the sponsor's responsibility to provide pertinent data on the mechanism of tumor induction and its relevance, or irrelevance, for humans. If the sponsor can show that an apparently drug-related tumor is species specific, the importance of the tumor in the overall evaluation of the drug will be greatly minimized. Table 14.7 presents a list of neoplastic/tumorigenic responses seen in rodents which have limited relevance to human safety. Part of the consideration must also be

recognition of the main characteristics of nongenotoxic carcinogens. These are recognized to be dose-dependent responses with operative thresholds. The major characteristics are (Spindler et al., 2000) as follows:

- Specificity (of species, sex, and organ).
- A threshold is operative and must be exceeded for cell proliferation and tumor development to occur.
- There is a stepwise dose–response curve/relationship between exposure, cell proliferation, and tumor development.
- The response is reversible with a cessation of dosing unless a point of no return has been passed.

Regulatory agencies are very aware of these challenges and deserve credit for attempting to respond to changes in the state of knowledge while still discharging their responsibility to protect the public health. For example, the latest version of the Japanese guidelines (Speid et al., 1990) acknowledges that the highest dose in a carcinogenicity study may be set at 100 times the clinical dose, instead of requiring that the MTD be achieved. It is also noteworthy that the FDA Center for Drug Evaluation established the Carcinogenicity Assessment Committee (CAC) representing

TABLE 14.7 Examples of Neoplastic Effects in Rodents with Limited Significance for Human Safety

| Neoplastic Effect | Pathogenesis (Agents) |
|--|--|
| Renal tubular neoplasia in male rats | $\alpha_2\mu$ -Globulin nephropathy/hydrocarbons (<i>d</i> -limonene, <i>p</i> -dichlorobenzene) |
| Hepatocellular neoplasia in rats and mice | Peroxisome proliferation (clofibrate, phthalate esters, phenoxy agents) Phenobarbital-like promotion |
| Urinary bladder neoplasia in rats | Crystalluria, carbonic anhydrase inhibition, urine pH extremes, melamine, saccharine, carbonic anhydrase inhibitors, dietary phosphates |
| Hepatocellular neoplasia in mice | Enzymatic-metabolic activation (in part unknown)/phenobarbital-like promotion |
| Thyroid follicular cell neoplasia in rats | Hepatic enzyme induction, thyroid enzyme inhibition/oxazepam, amobarbital, sulfonamides, thioureas |
| Gastric neuroendocrine cell neoplasia mainly in rats | Gastric secretory suppression, gastric atrophy induction (cimetidine, omeprazole, butachlor) |
| Adenohypophysis neoplasia in rats | Feedback interference/neuroleptics (dopamine inhibitors) |
| Mammary gland neoplasia in female rats | Feedback interference/neuroleptics, antiemetics, antihypertensives (calcium channel blockers), serotonin agonists, anticholinergics, exogenous estrogens |
| Pancreatic islet cell neoplasia in rats | Feedback interference/neuroleptics |
| Harderian gland neoplasia in mice | Feedback interference/misoprostol (PGE_1), nalidixic acid, aniline dyes |
| Adrenal medullary neoplasia in rats | Feedback interference (lactose, sugar alcohols) |
| Forestomach neoplasia in rats and mice | Stimulation of proliferation/butylated hydroxyanisole, phthalate esters, propionic acid |
| Lymphomas in mice | Immunosuppression/cyclosporin |
| Mononuclear cell leukemia in rats (mainly F344) | Immunosuppression (in part unknown)/furan, iodinated glycerol |
| Splenic sarcomas in rats | Methemoglobinemia (in part unknown)/dapson |
| Osteomas in mice | Feedback interference/lactose, sugar alcohols, H_2 antagonists, carbamazepine, vidarabine, isradipine, dopaminergics, finasteride |
| Leydig cell testicular neoplasia in mice | Feedback interference (proestrogens, finasteride, methoxychlor, cadmium) |
| Endometrial neoplasia in rats | Feedback interference (proestrogens, dopamine agonists) |
| Uterine leiomyoma in mice | Feedback interference (β_1 -antagonists) |
| Mesovarial leiomyoma in rats (occasionally in mice) | Feedback interference (β_2 -agonists) |
| Ovarian tubulostromal neoplasia in mice | Feedback interference (cytotoxic agents, nitrofurantoin) |

all drug review divisions. This group advises all the divisions on issues related to carcinogenicity and reviews and advises on proposed drug carcinogenicity testing programs, protocols, and test doses.

14.17 WEIGHT-OF-EVIDENCE FACTORS FOR CONSIDERATION IN A CARCINOGENICITY ASSESSMENT DOCUMENT (CAD)

Each of the factors listed in the following should be considered in formulating a prediction in the outcome and value of conducting a 2-year rat carcinogenicity study and an overall integrated assessment of the carcinogenic risk for humans. Some factors can be appropriate for both, other more appropriate for one or the other purpose. This guidance can be used as a guide by sponsors for writing a carcinogenicity assessment document (CAD):

- Knowledge of intended drug target and pathway pharmacology, secondary pharmacology, and drug target distribution in rats and humans

Target- and pathway-related mechanistic/pharmacologic and understood secondary pharmacologic characteristics can contribute to the prediction of outcomes of carcinogenicity studies and can improve prediction of potential human carcinogens. The CAD is expected to convey a thorough and critical assessment of the sponsor's knowledge of all such characteristics including a comprehensive literature review specifically addressing carcinogenicity risk. Examples of such data sources include the following:

- Prior experience with other molecules in the drug class
- Experience with human genetic polymorphisms in the target or pathway
- Clinical trial data
- Genetically engineered rodent models
- Animal disease models
- Unintended pharmacology
- Hormonal perturbation
- Targeted tissue genomic biomarker measurements

- *Genetic toxicology study results*

The criteria in ICH S2(R1) will be used to evaluate genetic toxicology data using a WOE approach.

- *Histopathologic evaluation of repeat-dose rat toxicology studies*

Histopathologic risk factors of neoplasia should be evaluated in the 6-month chronic rat study. Findings seen only in shorter-term repeat-dose rat toxicity studies are generally considered of less value for 2-year rat study outcome prediction but should be addressed. Histopathologic findings of particular interest include cellular hypertrophy, diffuse and/or focal cellular hyperplasia, persistent tissue injury and/or chronic inflammation, preneoplastic changes, and tumors. It is important to address the human relevance of such findings. For example, liver tumors are observed at relatively high frequency in the rat, sometimes with Leydig cell and thyroid follicular cell tumors. Hepatocellular hypertrophy associated with increased liver weight often results from hepatic enzyme induction, the latter being a well-understood mechanism of rodent specific tumorigenesis at these sites with little relevance to humans (McClain, 1989; Cook et al., 1999). The CAD should review the data supporting such mechanisms in assessing the risk for humans.

- *Exposure margins in chronic rat toxicology studies*

A high exposure margin in a chronic rat toxicology study absent of any carcinogenic risk factors can provide additional support for a carcinogenicity study waiver. Additionally, risk factors for neoplasia occurring only at high multiples of anticipated human exposure may provide additional support for considering a carcinogenicity study waiver. The inability to achieve high exposure margins in a chronic rat toxicology study due to limitations of tolerability, pharmacology, or absorption would not preclude a rat carcinogenicity study waiver.

- *Metabolic profile*

As per ICH S1C(R2), a comparison of the metabolic profile between rats and humans should also be taken into account when assessing the potential carcinogenicity of small molecules. Therefore, the adequacy of the metabolic profile in rats and exposure to human metabolites should also be discussed in the CAD.

- *Evidence of hormonal perturbation*

Evidence of hormonal perturbation should be considered from both repeat-dose and reproductive toxicology studies. Such evidence can come from weight, gross, and/or microscopic changes in endocrine organs or parameters from reproductive toxicology studies. Serum hormone levels can be useful to address findings but are not always essential.

- *Immunosuppression*

Immunosuppression can be a causative factor for tumorigenesis in humans. Effects on the immune system can alter tumor surveillance or result in tumors secondary to recrudescence of oncogenic viruses. As such, an assessment of potential impact to the immune system should be evaluated according to the ICH S8 guideline and factored into the CAD.

- *Special studies and end points*

Data from special stains, new biomarkers, emerging technologies, and alternative test systems can be submitted with scientific rationale to help explain or predict animal and/or human carcinogenic pathways and mechanisms when they would contribute meaningfully.

- *Results of nonrodent chronic study*

Assessment of carcinogenic risk factors in the nonrodent toxicology studies should be considered for human risk assessment regardless of results in the chronic rat study.

- *Transgenic mouse study*

A transgenic mouse carcinogenicity study (usually *raH2* or *p53^{+/-}* mouse) is not required for the WOE argument. However, if conducted on a case-by-case basis, a transgenic mouse carcinogenicity study can contribute to the WOE.

14.18 CONCLUSIONS

The design, conduct, and interpretation of carcinogenicity studies are the major challenges for the pharmaceutical toxicologist, pathologist, biostatistician, and regulator. This is a rapidly changing field generating more questions than answers. The largest question continues to be the extrapolation of rodent data to humans, especially when data on mechanisms of tumor induction are unavailable or controversial. Much has been written on the difficulties inherent in extrapolating results from rodents treated with MTDs of a compound to humans who will be exposed to much lower doses and often for shorter periods. A discussion of these and other aspects of carcinogenic risk assessment is beyond the scope of this chapter.

REFERENCES

- Abuja PM, Albertini R. (2001) Methods for monitoring oxidative stress, lipid peroxidation, and oxidation resistance of lipoproteins. *Clin. Chim. Acta.* 306:1–17.
- American Cancer Society. (2013) *Cancer Facts & Figures 2013*. American Cancer Society, Atlanta, GA.
- Ames BN, Shigenaga MK. (1993) Oxidants are a major contributor to cancer and aging. In: *DNA and Free Radicals* (Halliwell B and Aruoma O, Eds.). Ellis Horwood Ltd, London, pp. 1–15.

- Andersen M, Brusick D, Cohen S, Dragon Y, Frederick C, Goodman JI, Hard G, Meek B, Melnick RL, O'Flaherty EJ. (1998) U.S. Environmental Protection Agency's revised cancer guidelines for carcinogen risk assessment. *Toxicol. Appl. Pharmacol.* 153:133–136.
- Asagoshi K, Wilson SH. (2011) Base-excision repair: role of DNA polymerase β in late-stage base excision repair. In: *Chemical Carcinogenesis* (Penning TM, Ed.). Humana Press, New York, pp. 297–320.
- Balmain A, Brown K. (1988) Oncogene activation in chemical carcinogenesis. *Adv. Cancer Res.* 51:147–182.
- Basu AK. (2011) Mutagenesis: the outcome of faulty replication of DNA. In: *Chemical Carcinogenesis* (Penning TM, Ed.). Humana Press, New York, pp. 375–400.
- Baylin SB. (1997) Tying it all together: epigenetics, genetics, cell cycle, and cancer. *Science.* 277:1948–1949.
- Becker RA, Shank RC. (1985) Kinetics of formation and persistence of ethylguanines in DNA of rats and hamsters treated with diethylnitrosamine. *Cancer Res.* 45:2076–2084.
- Bellhorn RW. (1980) Lighting in the animal environment. *Lab. Anim. Sci.* 30:440–450.
- Berenblum I, Shubik P. (1947) A new, quantitative, approach to the study of the stages of chemical carcinogenesis in the mouse's skin. *Br. J. Cancer.* 1(4):383–391.
- Bhandari R, Baghi Bhandari S, Sharma D. (2012) *Molecular Basis of Carcinogenesis, Molecular Basis of Cancer*. LAP Lambert Academic Publishing, Saarbrücken.
- Bishop JM. (1991) Molecular themes in oncogenesis. *Cell.* 64:235–248.
- Breslow N. (1984) Comparison of survival curves. In: *Cancer Clinical Trials: Methods and Practice* (Buyse ME, Staquet MJ and Sylvester RJ, Eds.). Oxford University Press, Oxford, pp. 381–406.
- Burkhardt S, Mellert W, Reinacher M, Bahnenmann R. (2001) Zonal evaluation of proliferation and apoptosis in mice reveals new mechanistic data for PB, WY 14, 643 and CH. *Toxicologist.* 60:286.
- Burns FJ, Vanderlan M, Snyder E, Albert RE. (1976) Induction and progression kinetics of mouse skin papillomas. In: (T.J. Slage, A. Sivak, and R.K. Boutwell), *Carcinogenesis, Vol. 2, Modifiers of Chemical Carcinogenesis*, Raven Press, New York, pp. 91–96.
- Callegari AJ, Kelly TJ. (2007) Shedding light on DNA damage checkpoint. *Cell Cycle.* 6:660–666.
- Cameron TP, Hickman RL, Kornreich MR, Tarone RE. (1985) History, survival, and growth patterns of B6C3F1 mice and F344 rats in the National Cancer Institute Carcinogenesis Testing Program. *Fundam. Appl. Toxicol.* 5:526–538.
- Capen CC, Dayan AD, Green S. (1995) Receptor-mediated mechanisms in carcinogenesis: an overview. *Mutat. Res.* 333(1–2):215–224.
- di Caprio L, Cox BS. (1981) DNA synthesis in UV-irradiated yeast. *Mutat. Res.* 82:69–85.
- Chow SC, Liu J-P. (1998) *Design and Analysis of Animal Studies in Pharmaceutical Development*. Marcel Dekker, New York.
- Chu K, Cueto C, Ward J. (1981) Factors in the evaluation of 20 NCI carcinogenicity bioassays. *J. Toxicol. Environ. Health.* 8:251–280.
- Clarkson PM, Thompson HS. (2000) Antioxidants: what role do they play in physical activity and health? *Am. J. Clin. Nutr.* 72:637–646.
- Colburn NH. (1980) Tumor promotion and preneoplastic progression. In: *Carcinogenesis: Modifiers of Carcinogenesis* (Slaga TJ, Ed.), Vol. 5. Raven Press, New York, pp. 33–56.
- Colombo BM, Benedetti S, Ottolenghi S, Mora M, Pollo B, Poli G, Finocchiaro G. (1995) The “bystander effect”: association of U-87 cell death with ganciclovir-mediated apoptosis of nearby cells and lack of effect in athymic mice. *Hum. Gene Ther.* 6(6):763–772.
- Conney AH. (1982) Introduction of microsomal enzymes by foreign chemicals and carcinogenesis by polycyclic aromatic hydrocarbon. *Cancer Res.* 42:4875–4917.
- Contrera JF, Jacobs AC, DeGeorge JJ. (1997) Carcinogenicity testing and the evaluation of regulatory requirements for pharmaceuticals. *Regul. Toxicol. Pharmacol.* 25:130–145.
- Cook JC, Kelinefelter GR, Hardisty JF, Sharpe RM, Foster PMD. (1999) Rodent Leydig cell tumorigenesis: a review of the physiology, pathology, mechanisms and relevance to humans. *Crit. Rev. Toxicol.* 29:169–261.
- Cox DR. (1972) Regression models and life-tables. *J. R. Stat. Soc.* 34B:187–220.
- Creech HJ. (2000) A brief history of the American Association for Cancer Research, Inc.: 1907–1995. In: *Directory of the American Association for Cancer Research*. American Association for Cancer Research, Philadelphia, PA, pp. viii–xi.
- Crowley J, Breslow N. (1984) Statistical analysis of survival data. *Annu. Rev. Public Health.* 5:385–411.
- Crum Brown A, Fraser TR. (1869) On the connection between chemical constitution and physiological action. Part II. On the physiological action of the ammonium bases derived from atropia and conia. *Trans. R. Soc. Edinb.* 25:693–739.
- Darwanto A, Van Ornam JD, Valinluck Lao V, Sowers LC. (2011) Chemical carcinogenesis and epigenetics. In: *Chemical Carcinogenesis* (Penning TM, Ed.). Humana Press, New York, pp. 245–266.
- Dempster AP, Selivyn MR, Weeks BJ. (1983) Combining historical and randomized controls for assessing trends in proportions. *J. Am. Stat. Assoc.* 78:221–227.
- DePass LR, Weil CS, Ballantyne B, Lewis SC, Losco PE, Reid JB, Simon GS. (1986) Influence of housing conditions for mice on the results of a dermal oncogenicity bioassay. *Fundam. Appl. Toxicol.* 7:601–608.
- Diamandopoulos GT. (1996) Cancer: a historical perspective. *Anticancer Res.* 16:1595–1602.
- Ding M, Shi X, Castranova V, Vallythan V. (2000) Predisposing factors in occupational lung cancer: inorganic minerals and chromium. *J. Environ. Pathol. Toxicol. Oncol.* 19:129–138.
- Dinse, GE (1985) *Estimating tumor prevalence, lethality and mortality*, Presented at the Symposium on Long-term Animal Carcinogenicity Studies: A Statistical Perspective, March 4–6, 1985, Bethesda, MD.
- Dinse GE, Lagakos SW. (1983) Regression analysis of tumor prevalence data. *Appl. Stat.* 32:236–248.
- Eastin WC, Haseman JK, Mahler JF, Bucher JR. (1998) The National Toxicology Program evaluation of genetically altered

- mice as predictive models for identifying carcinogens. *Toxicol. Pathol.* 26:461–584.
- Eyford JE, Bodvarsdottir SK. (2005) Genomic stability and cancer: networks involved in response to DNA damage. *Mutat. Res.* 592:18–28.
- Farrar DB, Crump KS. (1988) Exact statistical tests for any carcinogenic effect in animal bioassays. *Fundam. Appl. Toxicol.* 11:652–663.
- Festing MW. (1979) Properties of inbred strains and outbred stocks, with special reference to toxicity testing. *J. Toxicol. Environ. Health.* 5:53–68.
- Flynn RJ. (1960) Studies on the aetiology of ringtail in rats. *Proc. Anim. Care Panel.* 9:155–160.
- Food and Drug Administration (FDA). (2007) Toxicological principles for the safety assessment of direct food additives and color additives used in food. In: *Redbook 2000*, revised July, 2007. Food and Drug Administration, Washington, DC, pp. 111–115.
- Food and Drug Administration (FDA). (2001) *Guidance for Industry: Statistical Aspects of the Design, Analysis and Interpretation of Chronic Rodent Carcinogenicity Studies of Pharmaceuticals*. USDHEW, Washington, DC.
- Foulds L. (1969) *Neoplastic Development*, Vol. 1. Academic Press, New York.
- Foulds L. (1975) *Neoplastic Development*, Vol. 2. Academic Press, New York.
- Fox JG. (1977) Clinical assessment of laboratory rodents on long term bioassay studies. *J. Environ. Pathol. Toxicol.* 1:199–226.
- French JE, Spalding JW, Dunnick JK, Tice RR, Furedi-Marchacck M, Tennant RW. (1999) The use of transgenic animals in cancer testing. *Inhal. Toxicol.* 11:541–544.
- Gad SC. (2005) *Statistics and Experimental Design for Toxicologists*, 4th Ed. CRC Press, Boca Raton, FL.
- Gad SC, Spainhour CB, Gad SE. (2015) Carcinogenicity and genotoxicity. In: *Mammalian Toxicology* (Abou-Donia MB, Ed.). John Wiley & Sons, Inc., Hoboken, NJ, pp. 341–393.
- Gart JJ, Chu KC, Tarone RE. (1979) Statistical issues in interpretation of chronic bioassay tests for carcinogenicity. *J. Natl. Cancer Inst.* 62:957–974.
- Gart JJ, Krewski D, Lee PN, Tarone RE, Wahrendorf J. (1986) The design and analysis of long-term animal experiment. In: *Statistical Methods in Cancer Research*, Vol. III IARC Scientific Publication No. 79. International Agency for Research on Cancer, Lyon.
- Gaylor DW, Kodell RL. (2001) Dose-response trend tests for tumorigenesis adjusted for differences in survival and body weight across doses. *Toxicol. Sci.* 59:219–225.
- Gold LS, Sawyer CB, Magaw R, Blackman G, deVeciana M, Levenson R, Hooper NK, Havender WR, Bernstein L, Peto R, Pike MC, Ames BN. (1984) A carcinogenic potency data base of the standardized results of animal bioassays. *Environ. Health Perspect.* 58:9–319.
- Greenman DL, Bryant P, Kodell RL, Sheldon W. (1982) Influence of cage shelf level on retinal atrophy in mice. *Lab. Anim. Sci.* 32:353–356.
- Greenman DL, Kodell RL, Sheldon WG. (1984) Association between cage shelf level and spontaneous induced neoplasms in mice. *J. Natl. Cancer Inst.* 73:107–113.
- Groopman JD, Wogan GN. (2011) Aflatoxin and hepatocellular carcinoma. In: *Chemical Carcinogenesis* (Penning TM, Ed.). Humana Press, New York, pp. 113–134.
- Hajian G. (1983) Statistical issues in the design and analysis of carcinogenicity bioassays. *Toxicol. Pathol.* 11:83–89.
- Haseman JK. (1983) A reexamination of false-positive rates for carcinogenicity studies. *Fundam. Appl. Toxicol.* 3:334–339.
- Haseman JK. (1985) Issues in carcinogenicity testing: dose selection. *Fundam. Appl. Toxicol.* 5:66–78.
- Haseman JK. (1990) Use of statistical decision rules for evaluating laboratory animal carcinogenicity studies. *Fundam. Appl. Toxicol.* 14:637–648.
- Haseman JK, Crawford DD, Huff JE, Boorman GA, McConnell EE. (1984a) Results for 86 two-year carcinogenicity studies conducted by the National Toxicology Program. *J. Toxicol. Environ. Health.* 14:621–639.
- Haseman JK, Huff J, Boorman GA. (1984b) Use of historical control data in carcinogenicity studies in rodents. *Toxicol. Pathol.* 12:126–135.
- Haseman JK, Winbush JS, O'Donnell MW. (1986) Use of dual control groups to estimate false positive rates in laboratory animal carcinogenicity studies. *Fundam. Appl. Toxicol.* 7:573–584.
- Herzig M, Christofori G. (2002) Recent advances in cancer research: mouse models of tumorigenesis. *Biochim. Biophys. Acta.* 1602:97–113.
- Holliday R. (1990) Mechanisms for the control of gene activity during development. *Biol. Rev.* 65:431–471.
- Holm-Nielsen P, Olson TS. (1988) Ultrastructure of renal adenoma. *Ultrastruct. Pathol.* 12:27–39.
- Hood A, Liu J, Klaassen CD. (1999) Effects of phenobarbital, pregnenolone-16 α -carbonitrile, and propylthiouracil on thyroid follicular cell proliferation. *Toxicol. Sci.* 50:45–53.
- Hsieh LL, Hsieh TT. (1993) Detection of aflatoxin B1-DNA adducts in human placenta and cord blood. *Cancer Res.* 53:1278–1280.
- Hussain SP, Harris CC. (1999) Mutation spectrum and load: the generation of hypotheses linking the exposure of endogenous or exogenous carcinogens to human cancer. *Mutat. Res.* 428(1–2):23–32.
- ICH. (1995) *S1A Need for Carcinogenicity Studies of Pharmaceuticals*. International Conference on Harmonisation, Geneva, Switzerland.
- ICH. (1997) *S1B Testing for Carcinogenicity of Pharmaceuticals*. International Conference on Harmonisation, Geneva, Switzerland.
- ICH. (2008) *S1C(R2) Dose Selection for Carcinogenicity Studies of Pharmaceuticals*. International Conference on Harmonisation, Geneva, Switzerland.
- FDA. (2012) *S1 Rodent Carcinogenicity Studies for Human Pharmaceuticals*. Concept Paper. International Conference on Harmonisation, Geneva, Switzerland.
- FDA. (2013) *Proposed Change to Rodent Carcinogenicity Testing of Pharmaceuticals*. Regulatory Notice Document. International Conference on Harmonisation, Geneva, Switzerland. Revised March 10, 2015.
- International Agency for Research on Cancer (1987) *IARC Monographs on the Evaluation of Carcinogenic Risks to Humans—Preamble*. IARC Internal Technical Report 87/001, IARC, Lyon.

- Japanese Ministry of Health and Welfare (MHW). (1989) *Revised Guidelines for Toxicity Studies Required for Application for Approval of Manufacturing/Importing Drugs*. Ministry of Health and Welfare, Tokyo, pp. 37–48.
- Kelsell DP, Dunlop J, Hodgins MB. (2001) Human diseases: clues to cracking the connexin code? *Trends Cell Biol.* 11(1):2–6.
- Kiaris H. (2006) *Understanding Carcinogenesis: An Introduction to the Molecular Basis of Cancer*. Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim.
- Kitchin KT. (1999) *Carcinogenicity*. Marcel Dekker, New York.
- Klaunig JE. (2013) Chemical carcinogens. In: *Casarett and Doull's Toxicology: The Basic Science of Poisons* (Klaassen CD, Ed.), 8th Ed. McGraw-Hill, New York, pp. 393–444.
- Klaunig JE, Kamendulis LM. (2004) The role of oxidative stress in carcinogenesis. *Annu. Rev. Pharmacol. Toxicol.* 44:239–267.
- Klaunig JE, Kamendulis LM. (2008) Chemical carcinogens. In: *Casarett and Doull's Toxicology: The Basic Science of Poisons* (Klaassen CD, Ed.), 7th Ed. McGraw Hill, New York, pp. 329–380.
- Klaunig JE, Babich MA, Baetcke KP, Cook JC, Corton JC, David RM, Deluca JG. (2003) PPAR α agonist-induced rodent tumors: modes of action and human relevance. *Crit. Rev. Toxicol.* 33:655–780.
- Kodama S, Koike C, Negishi M, Yamamoto Y. (2004) Nuclear receptors CAR and PXR cross talk with FOXO1 to regulate genes that encode drug-metabolizing and gluconeogenic enzymes. *Mol. Cell. Biol.* 24:7931–7940.
- Lehmann AR. (1972) Post-replication repair of DNA in ultraviolet-irradiated mammalian cells. No gaps in DNA synthesized late after ultraviolet irradiation. *Eur. J. Biochem.* 31:438–445.
- Leithe E, Sirnes S, Omori Y, Rivedal E. (2006) Downregulation of gap junctions in cancer cells. *Crit. Rev. Oncog.* 12(3–4):225–256.
- Li R, Yerganian G, Duesberg P, Kraemer A, Willer D, Rausche C, Hehlmann R. (1997) Aneuploidy correlated 100% with chemical transformation of Chinese hamster cells. *Proc. Natl. Acad. Sci.* 94:14506–14511.
- Littlefield NA, Farmer JH, Gaylor DW, Sheldon WG. (1979) Effects of dose and time in a long-term, low-dose carcinogenic study. In: *Innovations in Cancer Risk Assessment (ED₀₁ Study)* (Staffa JA and Mehlman MA, Eds.). Pathotox Publishers, Park Forest South, IL, pp. 17–34.
- Loeb KR, Loeb LA. (2000) Significance of multiple mutations in cancer. *Carcinogenesis*. 21:379–385.
- Lopes M, Foiani M, Sogo JM. (2006) Multiple mechanisms control chromosome integrity after replication fork uncoupling and restart at irreparable UV lesions. *Mol. Cell.* 21:15–27.
- Mantel N. (1980) Assessing laboratory evidence for neoplastic activity. *Biometrics*. 36:381–399.
- Mantel N, Haenszel W. (1952) Statistical aspects of the analysis of data from the retrospective studies of disease. *J. Natl. Cancer Inst.* 22:719–748.
- Marselos M, Vainio H. (1991) Carcinogenic properties of pharmaceutical agents evaluated in the IARC Monographs programme. *Carcinogenesis*. 12:1751–1766.
- Martindale JL, Holbrook NJ. (2002) Cellular response to oxidative stress: signaling for suicide and survival. *J. Cell. Physiol.* 192:1–15.
- Masutani C, Kusumoto R, Yamada A, Dohmae N, Yokoi M, Yuasa M, Araki M, Iwai S, Takio K, Hanaoka F. (1999) The XPV (xeroderma pigmentosum variant) gene encodes human DNA polymerase η . *Nature*. 399:700–704.
- McAnulty PA. (2000) Transgenic mouse models in carcinogenicity testing. *Eur. Pharm. Contract.*
- McClain RM. (1989) The significance of hepatic microsomal enzyme induction and altered thyroid function in rats: implications for thyroid gland neoplasia. *Toxicol. Pathol.* 17:294–306.
- McGregor D. (2009) Carcinogenesis and carcinogens that are also genotoxic. In: *General and Applied Toxicology* (Ballantyne B, Marrs T and Syversen T, Eds.), 2nd Ed. MacMillan, London, pp. 1733–1756.
- McKinnell RG. (2006) Genetics and heredity. In: *The Biological Basis of Cancer* (McKinnell RG, Parchment RE, Perantoni AO, Damjanov I and Pierce GB, Eds.), 2nd Ed. Cambridge University Press, Cambridge, NY, pp. 126–144.
- McKnight B. (1988) A guide to the statistical analysis of long-term carcinogenicity assays. *Fundam. Appl. Toxicol.* 10:355–364.
- Melnick RL, Kohn MC, Portier CJ. (1996) Implications for risk assessment of suggested nongenotoxic mechanisms of chemical carcinogenesis. *Environ. Health Perspect.* 104(S1):123–133.
- Miller EC, Miller JA. (1981a) Mechanisms of chemical carcinogenesis. *Cancer*. 47:1055–1064.
- Miller EC, Miller JA. (1981b) Searches for ultimate chemical carcinogens and their reactions with cellular macromolecules. *Cancer*. 47(10):2327–2345.
- Miller EC, Swanson AB, Phillips DH, Fletcher TL, Liem A, Miller JA. (1983) Structure-activity studies of the carcinogenicities in the mouse and rat of some naturally occurring and synthetic alkenylbenzene derivatives related to safrole and estragole. *Cancer Res.* 43:1124–1134.
- Moennikes O, Loeppen S, Buchmann A, Andersson P, Ittrich C, Poellinger L, Schwarz M. (2004) Constitutively active dioxin/aryl hydrocarbon receptor promotes hepatocarcinogenesis in mice. *Cancer Res.* 64(14):4707–4710.
- Moggs JG, Goodman JJ, Trosko JE, Roberts RA. (2004) Epigenetics and cancer: implications for drug discovery and safety assessment. *Toxicol. Appl. Pharmacol.* 196:422–430.
- Morrison A, Christensen RB, Alley J, Beck AK, Bernstine EG, Lemontt JF, Lawrence CW. (1989) REV3, a *Saccharomyces cerevisiae* gene whose function is required for induced mutagenesis, is predicted to encode a nonessential DNA polymerase. *J. Bacteriol.* 171:5659–5667.
- Nebert DW, Roe AL, Dieter MZ, Solis WA, Yang Y, Dalton TP. (2000) Role of the aromatic hydrocarbon receptor and [Ah] gene battery in the oxidative stress response, cell cycle control, and apoptosis. *Biochem. Pharmacol.* 59:65–85.
- Neumann HG. (1983) Role of extent and persistence of DNA modifications in chemical carcinogenesis by aromatic amines. *Recent Results Cancer Res.* 84:77–89.
- NIH (2004) *13th Report on Carcinogens*. U.S. Department of Health and Human Services, Washington, DC.
- Noller KL, Decker DG, Lanier AP, Kurkland LT. (1972) Clear-cell adenocarcinoma of the cervix after maternal treatment with synthetic estrogens. *Mayo Clin. Proc.* 47:629–630.

- Pegg AE. (1984) Methylation of the O6 position of guanine in DNA is the most likely initiating event in carcinogenesis by methylating agents. *Cancer Invest.* 2(3):223–231.
- Peng Y, Wang H, Santana-Santos L, Kisker C, Van Houten B. (2011) Nucleotide excision repair from bacteria to humans: structure-function studies. In: *Chemical Carcinogenesis* (Penning TM, Ed.). Humana Press, New York, pp. 267–296.
- Penning TM. (2011) Metabolic activation of chemical carcinogens. In: *Chemical Carcinogenesis* (Penning TM, Ed.). Humana Press, New York, pp. 135–158.
- Perantoni AO. (2006) Cancer-associated genes. In: *The Biological Basis of Cancer* (McKinnell RG, Parchment RE, Perantoni AO, Damjanov I and Pierce GB, Eds.), 2nd Ed. Cambridge University Press, Cambridge, NY, pp. 80–125 & 145–194.
- Perera F, Mayer J, Santella RM, Brenner D, Jeffrey A, Latriano L, Smith S, Warburton D, Young TL, Tsai WY, Hemminki K, Brandt-Rauf P. (1991) Biological markers in risk assessment for environmental carcinogens. *Environ. Health Perspect.* 90:247–254.
- Peters JM, Aoyama T, Cattley RC, Nobumitsu U, Hashimoto T, Gonzalez FJ. (1998) Role of peroxisome proliferator-activated receptor alpha in altered cell cycle regulation in mouse liver. *Carcinogenesis*. 19:1989–1994.
- Peto R. (1974) Guidelines on the analysis of tumor rates and death rates in experimental animals. *Br. J. Cancer*. 29:101–105.
- Peto R, Pike MC, Day NE, Gray RG, Lee PN, Parish S, Peto J, Richards S, Wahrendorf J. (1980) Guidelines for simple, sensitive significance tests for carcinogenic effects in long-term animal experiments. In: *IARC Monographs on the Evaluation of the Carcinogenic Risk of Chemicals to Humans*. International Agency for Research on Cancer, Lyon.
- Pharmaceutical Manufacturers Association (PMA). (1988) *Results of a Questionnaire Involving the Design of and Experience with Carcinogenicity Studies*. This document has not been published in the open literature but is widely available within the pharmaceutical industry. It may be obtained by writing to the Pharmaceutical Manufacturers Association.
- Pierce GB, Damjanov I. (2006) The pathology of cancer. In: *The Biological Basis of Cancer* (McKinnell RG, Parchment RE, Perantoni AO, Damjanov I and Pierce GB, Eds.), 2nd Ed. Cambridge University Press, Cambridge, NY, pp. 14–49.
- Pitts JD. (1994) Cancer gene therapy: a bystander effect using the gap junctional pathway. *Mol. Carcinog.* 11(3):127–130.
- Pollard M, Luckert PH. (1985) Tumorigenic effects of direct- and indirect-acting chemical carcinogens in rats on a restricted diet. *J. Natl. Cancer Inst.* 74(6):1347–1349.
- Portier CJ, Bailer AJ. (1989) Testing for increased carcinogenicity using a survival-adjusted quantal response test. *Fundam. Appl. Toxicol.* 12:731–737.
- Powell CJ, Berry CL. (2000) Non-genotoxic or epigenetic carcinogenesis. In: *General and Applied Toxicology* (Ballantyne B, Marrs T and Syversen T, Eds.), 2nd Ed. MacMillan, London.
- Prakash L. (1981) Characterization of postreplication repair in *Saccharomyces cerevisiae* and effects of rad6, rad18, rev3, rad52 mutations. *Mol. Gen. Genet.* 184:471–478.
- Preston RJ, Hoffmann GR. (2013) Genetic toxicology. In: *Casarett and Doull's Toxicology: The Basic Science of Poisons* (Klaassen CD, Ed.), 7th Ed. McGraw-Hill, New York, pp. 445–480.
- Pugh G Jr, Isenberg JS, Kamendulis LM, Ackley DC, Clare LJ, Brown R, Lington AW, Smith JH, Klaunig JE. (2000) Effects of di-isononyl phthalate, di-2-ethylhexyl phthalate, and clofibrate in cynomolgus monkeys. *Toxicol. Sci.* 56:181–188.
- Purchase IFH. (1980) Interspecies comparisons of carcinogenicity. *Br. J. Cancer*. 41:454–468.
- Rao GN, Huff J. (1990) Refinement of long-term toxicity and carcinogenesis studies. *Fundam. Appl. Toxicol.* 15:33–43.
- Rao MS, Reddy JK. (1991) An overview of peroxisome proliferator-induced hepatocarcinogenesis. *Environ. Health Perspect.* 93:205–209.
- Rao GN, Birnbaum LS, Collins JJ, Tennant RW, Skow LC. (1988) Mouse strains for chemical carcinogenicity studies: overview of a workshop. *Fundam. Appl. Toxicol.* 10:385–394.
- Rao GN, Haseman JK, Edmondson J. (1989) Influence of viral infections on body weight, survival, and tumor prevalence in Fisher 344 rats on two-year studies. *Lab. Anim. Sci.* 39:389–393.
- Rupp WD, Howard-Flanders P. (1968) Discontinuities in the DNA synthesized in an excision-defective strain of *Escherichia coli* following ultraviolet irradiation. *J. Mol. Biol.* 31:291–304.
- Sancar A. (2003) Structure and function of DNA photolyase and cryptochrome blue-light photoreceptors. *Chem. Rev.* 103(6):2203–2237.
- Schweder T, Spjotvoll E. (1982) Plots of *p*-values to evaluate many tests simultaneously. *Biometrika*. 69:493–502.
- Semenza JC, Weasel LH. (1997) Molecular Epidemiology in environmental health: the potential of tumor suppressor gene as a biomarker. *Environ. Health Perspect.* 105(Suppl. 1):155–163.
- Sher SP, Jensen RD, Bokelman DL. (1982) Spontaneous tumors in control F344 and Charles River CD rats and Charles River CD-1 and B6C3F1 mice. *Toxicol. Lett.* 11:103–110.
- Silverberg SG. (1984) The autopsy and cancer. *Arch. Pathol. Lab. Med.* 108:476–478.
- Simeonova PP, Luster MI. (2000) Mechanisms of arsenic carcinogenicity: genetic or epigenetic mechanisms? *J. Environ. Pathol. Toxicol. Oncol.* 19:281–286.
- Sistare F, Morton D, Alden C, Christensen J, Keller D, Jonghe SD, Storer RD, Reddy MV, Kraynak A, Trela B, Bienvenu JG, Bjurström S, Bosmans V, Brewster D, Colman K, Dominick M, Evans J, Hailey JR, Kinter L, Liu M, Mahrt C, Marien D, Myer J, Perry R, Potenta D, Roth A, Sherratt P, Singer T, Slim R, Soper K, Fransson-Steen R, Stoltz J, Turner O, Turnquist S, van Heerden M, Woicke J, DeGeorge JJ. (2011) An analysis of pharmaceutical experience with decades of rat carcinogenicity testing: support for a proposal to modify current regulatory guidelines. *Toxicol. Pathol.* 39:716–744.
- Sontag J. (1981) Carcinogenicity of substituted-benzenediamines (phenylenediamines) in rats and mice. *J. Natl. Cancer Inst.* 66:591–602.
- Speid LH, Lumley CE, Walker SR. (1990) Harmonization of guidelines for toxicity testing of pharmaceuticals by 1992. *Regul. Toxicol. Pharmacol.* 12:179–211.
- Spindler P, Loan J, Ceuppens P, Harling R, Eittlin R, Lima BS. (2000) Carcinogenicity testing of pharmaceuticals in the European Union: a workshop report. *Drug Inf. J.* 34:821–828.
- Stary A, Sarasin A. (2002) The genetics of the hereditary xeroderma pigmentosum syndrome. *Biochimie*. 84:49–60.

- Swenberg JA, Dyroff MC, Bedell MA, Popp JA, Huh N, Kirstein U, Rajewsky MF. (1984) O4-ethyldeoxythymidine, but not O6-ethyldeoxyguanosine, accumulates in hepatocyte DNA of rats exposed continuously to diethylnitrosamine. *Proc. Natl. Acad. Sci.* 81(6):1692–1695.
- Swenberg JA, Richardson FC, Boucheron JA, Dyroff MC. (1985) Relationships between DNA adduct formation and carcinogenesis. *Environ. Health Perspect.* 62:177–183.
- Tarone RE. (1975) Tests for trend in life table analysis. *Biometrika.* 62:679–682.
- Tarone RE. (1982) The use of historical control information in testing for a trend in proportions. *Biometrics.* 38:215–220.
- Tavani A, Negri E, Parazzini F, Franceschi S, La Vecchia C. (1993) Female hormone utilisation and risk of hepatocellular carcinoma. *Br. J. Cancer.* 67(3):635–637.
- Tennant RW, French JE, Spalding JW. (1995) Identification of chemical carcinogens and assessing potential risks in short-term bioassays using transgenic mouse models. *Environ. Health Perspect.* 103:942–950.
- Tennant RW, Stasiewicz S, Mennear J, French JE, Spalding JW. (1999) Genetically altered mouse models for identifying carcinogens. In: *The Use of Short and Medium-Term Tests for Carcinogens and Data on Genetic Effects in Carcinogenic Hazard Evaluation* IARC Science Publications No. 146 (McGregor DB, Rice JM and Venith S, Eds.). International Agency for Research on Cancer, Lyon, pp. 23–148.
- Thomas DG, Breslow N, Gart JJ. (1977) Trend and homogeneity analyses of proportions and life table data. *Comput. Biomed. Res.* 10:373–381.
- Trosko JE, Chang CC, Upham BL, Wilson MR. (1998) Epigenetic toxicology as toxicant-induced changes in intracellular signaling leading to altered gap junctional intercellular communication. *Toxicol. Lett.* 102–103:71–78.
- Ueda A, Hamadeh HK, Webb HK, Yamamoto Y, Sueyoshi T, Afshari CA, Lehmann JM, Negishi M. (2002) Diverse roles of the nuclear orphan receptor CAR in regulating hepatic genes in response to phenobarbital. *Mol. Pharmacol.* 61:1–6.
- Volkert MR. (1998) Adaptive response of *Escherichia coli* to alkylation damage. *Environ. Mol. Mutagen.* 11(2):241–255.
- Watson JD, Baker TA, Bell SP, Gann A, Levine M, Losick R. (2004) *Molecular Biology of the Gene*, 5th Ed. Pearson Benjamin Cummings & Cold Spring Harbor Laboratory Press, Woodbury, NY.
- Weaver RJ, Brunden MN. (1998) The design of long-term carcinogenicity studies. In: *Design and Analysis of Animal Studies in Pharmaceutical Development* (Chow S and Lide J, Eds.). Marcel Dekker, Inc., New York.
- Weisbrode SE, Weiss HS. (1981) Effect of temperature on benzo(a)pyrene induced hyperplastic and neoplastic skin lesions in mice. *J. Natl. Cancer Inst.* 66:978–981.
- White TW, Paul DL. (1999) Genetic diseases and gene knockouts reveal diverse connexin functions. *Annu. Rev. Physiol.* 61:283–310.
- Whysner J, Ross PM, Williams GM. (1996) Phenobarbital mechanistic data and risk assessment: Enzyme induction, enhanced cell proliferation, and tumor promotion. *Pharmacol. Ther.* 71:153–191.
- Willecke K, Eiberger J, Degen J, Eckardt D, Romualdi A, Güldenagel M, Deutsch U, Söhl G. (2002) Structural and functional diversity of connexin genes in the mouse and human genome. *Biol. Chem.* 383(5):725–737.
- Williams GM, Iatropoulos MJ. (2001) Principles of testing for carcinogenic activity. In: *Principles and Methods of Toxicology* (Hayes AW, Ed.), 5th Ed. Taylor & Francis, Philadelphia, PA.
- Wiskemann A, Sturm E, Klehr NW. (1986) Fluorescent lighting enhances chemically induced papilloma formation and increases susceptibility to tumor challenge in mice. *J. Cancer Res. Clin. Oncol.* 112:141–143.
- Zhang H, Lawrence CW. (2005) The error-free component of the RAD6/RAD18 DNA damage tolerance pathway of budding yeast employs sister-strand recombination. *Proc. Natl. Acad. Sci.* 102:15954–15959.

HISTOPATHOLOGY IN NONCLINICAL PHARMACEUTICAL SAFETY ASSESSMENT

15.1 INTRODUCTION

Toxicologic pathology is the study of the molecular, cellular, tissue, and/or organ responses of a living organism when exposed to injurious chemical or physical agents. These responses represent a spectrum of cellular changes ranging from cell death to malignant transformations, tissue and organ responses (including regeneration inflammation), and organization and overall response as identified by clinical changes and alterations in body fluids (Arnold et al., 1990; MHLW, 1990; EEC, 1992; EMEA, 2000; ICH, 2000, 2005; Haschek et al., 2013b; McCance and Huether, 2015). It starts from recognition of the fact that the cell constitutes the basic unit of life. Accordingly, morphologic changes in organs and tissues that arise as a result of injury begin with the responses of underlying cells to the toxic insult. A proper evaluation and understanding of related pathology must start at the cellular level. Some cellular components whose alterations have been reported to be critically associated with cell injury include the plasma membrane, site of osmotic, electrolyte, and water regulation, as well as signal transduction; the mitochondrion, site of energy storage and aerobic respiration; the endoplasmic reticulum, site of much protein synthesis; and the nucleus, which contains the genetic information and where transcription of the genetic code takes place (Ham, 1974; Wallig, 2002).

Biochemical changes such as enzyme induction and gene expression occur at the earlier stages of the exposure-disease continuum. The degrees of cellular injury in different target tissue depend on the metabolic rate. Cells with high metabolic rates such as neurons, myocardial cells, and renal proximal convoluted tubule epithelial cells frequently suffer from injury more quickly than low-metabolizing ones. These high-metabolism cells depend on a continuous flow of oxygen to

conduct the aerobic metabolism necessary to provide required energy in the form of ATP for the maintenance of membrane polarity and membrane integrity (neurons), for the continual muscular contraction/relaxation and calcium transport (myocardium), and for the transport of fluids, electrolytes, and metabolites (renal PTCs). Hence, any depletion in oxygen supply is likely to have a significant impact on their survival.

In contrast, cells with low metabolic activity such as fibroblasts and adenocytes are less affected by the low supply of oxygen, and they have a prominent role in regeneration and scarring. Homeostasis is one of the most remarkable and most typical properties of highly complex biologic systems. It defines the ultimate environment under which cells maintain the physiochemical conditions (intracellular pH, cytosolic osmolarity, ion gradients) necessary to perform their biologic functions. In biologic systems, homeostasis is maintained by means of a multiplicity of dynamic equilibria rigorously controlled by interdependent regulation mechanisms. Hence, the homeostatic system reacts to changes or disturbances in response to various insults by exerting a series of modifications or adjustments to maintain the internal balances and conditions within tolerable limits.

Pathology (including all the aspects of anatomic, aka histopathology, clinical chemistry and clinical pathology, aka hematology), is generally considered the single most significant portion of data to come out of a systemic toxicity studies (particularly the repeat dose with versions going from 14 days to 2 years in duration) (Roy and Andrews, 2004; Perez and Barthold, 2007; Tehounwan and Centeno, 2008; Frame et al., 2014). Anatomic pathology evaluations actually consist of three related sets of data (gross pathology observations, organ weights, and microscopic pathology) that are collected during termination of the study animals.

TABLE 15.1 Tissues for Histopathology

| Adrenals ^a | Mainstream Bronchi |
|--------------------------------------|--|
| Body and cervix | Major salivary gland |
| Brain, all three levels ^a | Mesenteric lymph nodes |
| Cervical lymph nodes | Ovaries and tubes |
| Cervical spinal cord | Pancreas |
| Duodenum | Pituitary |
| Esophagogastric junction | Prostate |
| Esophagus | Skeletal muscle from proximal hind limb |
| Eyes with optic nerves | Spleen ^a |
| Femur with marrow | Sternebrae with marrow |
| Heart ^a | Stomach |
| Ileum | Testes with epididymides ^a |
| Kidneys ^a | Thymus and mediastinal contents ^a |
| Large bowel | Thyroid with parathyroid ^a |
| Larynx with thyroid and parathyroid | Trachea |
| Liver ^a | Urinary bladder |
| Lungs ^a | Uterus including horns |

^aOrgans to be weighed.

At the end of the study, a number of tissues are collected during termination of all surviving animals (test and control). Organ weights and terminal body weights are recorded at study termination so that absolute and relative (to body weight) values can be statistically evaluated. Michael et al. (2007) have provided a review of practices of how such organ weight information is evaluated and utilized in the overall evaluation of pathology and adverse effects. In general with the exception of brain weights relative to body weight, changes are considered more relevant to identifying target organ toxicities.

These “standard list” tissues, along with the organs for which weights are determined, are listed in Table 15.1 (OECD, n.d.). All tissues collected are typically processed for microscopic observation, but only those from the high-dose and control groups are necessarily evaluated microscopically. If a target organ is discovered in the high-dose group, then the effect is “followed” in successively lower-dose groups until a NOEL (devoid of effect) is determined.

In theory, all microscopic evaluations should be performed in a blind manner (without the pathologist knowing from which dose group a particular animal came), but this is difficult to do in practice, and such an approach frequently actually limits and degrades the quality of the evaluation. Like all the other portions of data in the study, proper evaluation benefits from having access to all data that addresses the relevance, severity, timing, and potential mechanisms of a specific toxicity. Blind examination is best applied in peer review or consultations on specific findings after a primary evaluation.

In addition to the “standard” set of tissues specified in Table 15.1, observations during the course of the study or

TABLE 15.2 Principles of Drug Testing Before Trials in Humans as Defined in 1938 by Geiling and Cannon

1. Exact composition of drug should be known; if not, method of preparation
2. Acute toxicity studies in animals of different species
3. Chronic toxicity experiments at varying doses in different species for cumulative effects
4. Careful and frequent observations of animals to develop a composite picture of clinical effects
5. Careful pathological examination of tissues with appropriate stains
6. Effects of drugs on excretory or detoxifying organs, especially kidney and liver
7. Rate of absorption and elimination, path and manner of excretion, and concentration in blood and tissues at varying times
8. Possible influence of other drugs and foodstuffs
9. Careful examination for any synergies or untoward reactions

in other or previous studies may dictate that special examination or tissue preparation procedures (such as polarized light or electron microscopy, immunocytochemistry, or quantitative morphometry) be undertaken to evaluate the relevance of such findings and help or understand the mechanisms underlying certain observations.

The evaluation of the pathological alterations induced in laboratory animals by new drugs represents the cornerstone of their safety assessment before they can be first tried in patients. This preliminary assessment, which is based largely on conventional histopathological techniques, represents a major contribution to the development of new treatments for both human and animal diseases.

Although there have been many changes over the past few decades in the details of study design and conduct, the principles of drug testing prior to trial in humans are the same as those expounded by Geiling and Cannon after they studied the pathological effects and causes of death of patients treated with toxic elixir of sulfanilamide over 60 years ago (Geiling and Cannon, 1938; Table 15.2). The basic paradigm of dosing laboratory animals with various doses of new drug for increasing periods of time accompanied by careful clinical observations and biochemical and hematological monitoring followed by histopathological examination of the tissue remains essentially unaltered. The pathologist is required to not only evaluate alterations to organs and tissues and any relationship that they might have to drug treatment but also to assess the likely relevance any treatment-related findings might have for patients (see Table 15.3).

Statistical analysis can also be helpful in assessing whether findings are chance or not, and the causality and relevance of such changes (Gad and Rousseaux, 2013), but care must be taken in their application.

The use of animals to study the pathological effects of chemicals and therapeutic agents has a long history.

TABLE 15.3 Discriminating Factors for Assessing Cause–Effect Relationships and Adversity of Pathology Findings*Discriminating Factors for Assessing Cause–Effect Relationship*

There is no obvious dose–response
 The group change is due to an outlier in one or more animals
 The measurement of the end point is inherently imprecise
 The change is within normal biological variation (historical control or reference values)
 There is a lack of biological plausibility (e.g., the difference is inconsistent with class effects, mode of action, or what is known or expected of the test material)

Discriminating Factors for Assessing Adversity

The effect causes no alteration in the general function of the test organism or the organs/tissues affected
 The effect is adaptive
 The effect is transient (i.e., resolves in the course of treatment vs. reversibility, which refers to resolution with cessation of treatment)
 The severity of the effect is limited (below threshold of concern)
 The effect is not a precursor (i.e., not part of a continuum of changes known to progress with time to an established adverse effect)
 The effect is secondary to other adverse effects
 The effect is a consequence of the experimental model (e.g., stress associated with restraint or reactions to physical properties of the test substance, such as taste or odor).

Source: Adapted from Lewis et al. (2002, pp. 66–74).

In the eighteenth century, Morgagni reported his attempts to compare pathological changes produced by accidental ingestion by people of chemicals such as arsenic. A thorough and systematic review of pathology induced by toxins in humans and animals was published by Orfila as long ago as 1815 (Orfila, 1815). Although in the modern era drug safety evaluation has been practiced in rodent and nonrodent species widely since before World War II, there have been very few critical comparisons of the effects of drugs in man and those seen in laboratory animals. Much potentially useful information still awaits data mining in the archives of pharmaceutical companies and government agencies. Nevertheless, the available data suggests that the traditional approach using experimental pharmacology alongside conventional toxicology studies with pathology is usually sufficient to predict important adverse effects and to support the safe conduct of the first clinical studies in humans (Turton and Hoasen, 1998; Greaves et al., 2004). Such a degree of concordance varies between different organs and tissues. Therefore each observed drug-induced pathological finding needs to be assessed on a case-by-case basis for its likely clinical relevance. For some systems, histopathology remains critical for some organ systems but of lesser importance for others. Traditional animal studies are poor predictors of subjective neurological symptoms, but histopathological examination of the nervous system in laboratory

animals treated with cancer drugs does well at potential serious clinical neurotoxic effects. Results from nonclinical studies frequently fail to predict renal and hepatic toxicity (largely because of a “formula” approach to evaluation), but there is generally a good correlation for gastrointestinal effects, and histopathology still seems to represent one of the most sensitive techniques to detect effects on the reproductive system; though the relevance of such findings can be confounded by general systemic toxicity, the pathologist also needs to be aware that some minor inflammatory alterations in certain organs, such as the liver, may have greater significance for the use of a drug in humans than other types of severe damage such as subendocardial necrosis in the myocardium mediated by exaggerated hemodynamic effects.

Treatment-induced findings in conventional toxicity studies found in different laboratory animal species also seem to possess different degrees of relevance for humans. Although the data is fragmentary, findings in beagle dog studies are not often better predictors of human adverse effects than data from rodents or, surprisingly, from primates (Greaves et al., 2004). Dog gastrointestinal and cardiovascular physiology appears to model particularly well for humans, though the pig is generally better yet.

Another long-standing issue most recently recognized again due to findings with the cyclooxygenase 2 (COX-2) inhibitors is the importance of evaluating the adverse effects of some therapies with specific human diseases. COX-2 inhibitors were used for inflammatory disorders because of their perceived lower adverse effect profile on the gastrointestinal tract compared with conventional drugs, but this benefit is outweighed by an increased incidence of cardiovascular disease in some patients. Such effects are difficult if not impossible to predict from nonclinical safety studies in “normal” healthy animals. Unfortunately the detection of an increased incidence of a common event such as heart attack or stroke is difficult in patients, for it requires careful collection and analysis of data, even though it may have a big impact on public health (Dragen, 2005). Such interactions usually require randomized controlled trials specifically designed to look for such risks. It has to be remembered that aspirin was in use for over 100 years before it became generally acknowledged about 30 years ago to be associated with Reye’s syndrome, a devastating toxicity in children (Monto, 1999). While the actual mechanism involved in Reye’s syndrome is still unknown, it is often preceded by a viral infection and displays a strong correlation with the subsequent ingestion of aspirin.

Histopathology testing is a terminal procedure, and, therefore, sampling of any single animal is a one-time event (except in the case of a tissue collected by biopsy). Because it is a regulatory requirement that the tissues from a specific minimum number of animals be examined at the stated end of the study, any evaluation of effects in tissues at another time course (most commonly, to investigate recovery from an effects at study termination) requires that additional

numbers of animals be incorporated into that study at start-up. Such animals are randomly assigned at the beginning of the study and otherwise treated exactly the same as the rest of their group/cohort animals.

Anatomic pathology evaluation occurs only after the in-life portion of a study is complete and is typically the rate-limiting step in producing a report on the study. At the end of the study the animals are euthanized, final blood and urine samples are taken, and tissues are collected, with a prespecified set being weighed while still wet, and evaluated grossly (see Table 15.1) as to whether they are other than normative in appearance then preserved in appropriate manners (Greaves, 2009, 2012; Haschek et al., 2013a) and processed so as to optimize evaluation (Gray, 1964). Experimental design may call for an interim necropsy (to allow for evaluation of progression of lesions or observation of indications of adaptive change by animals) and/or a recovery group (usually additional numbers of high-dose and control animals in which treatment is discontinued at the time of the main necropsy). Such recovery animals are maintained without further manipulation or treatment for a period of time after the termination of the main study animals (usually a month), allowing for an assessment of treatment-free regression or progression of conditions seen at the end of the main study. Once gathered, tissues must be processed, mounted, stained, and examined with great care. The entire set of steps involved in anatomic pathology represent a significant portion of the time required to complete a study and add from 30 to 50% to the cost of the shorter (14, 28, and 90 days) studies.

Nevertheless, the available data suggests that the traditional approach using experimental pharmacology alongside conventional toxicology studies with pathology is usually sufficient to predict important adverse effects and to support the safe conduct of the first clinical studies in humans (Greaves et al., 2004). Indeed, the dosing a rodent and non-rodent species with a new drug up to one month identifies over 90% of adverse effects that will ever be detected in the usual nonclinical safety assessment studies. However, more generally, these studies do not detect all adverse drug effects that can occur in clinical practice, and there remains significant over- and underprediction of human toxicity. Overall, the true positive concordance rate (sensitivity) is of the order of 70% with perhaps 30% of human adverse effects not predicted by safety pharmacology or conventional toxicity studies (Olsen et al., 2000). Moreover, this concordance varies between different organs and tissues. Therefore each drug-induced pathological finding needs to be assessed on a case-by-case basis for its likely clinical relevance.

15.1.1 Pathological Techniques

Over the past few years, a number of excellent reviews of standardized techniques for use in the histopathology evaluation of toxicology studies have been produced covering

tissue selection, blocking and sectioning procedures, immunocytochemical stains for laboratory animals, and other basic techniques (Bregman et al., 2003; Mikaelian et al., 2004).

In addition, the scientific literature is full of interesting techniques and novel reagents that can be applied to tissue sections. Some of these can be very useful in the analysis of pathological alterations in toxicity studies; some fail to work in routinely fixed material. However, it is important that these techniques are used in a judicious manner with clear aims following careful analysis of conventional hematoxylin- and eosin-stained sections. This is particularly true for the application of microarray and bioinformatics technology. While undoubtedly useful in toxicology, these techniques should not be applied in isolation but in combination with other information, particularly pathology.

15.1.2 Organ Weights

Regulatory guidelines indicate that certain organs should be weighed during the course of the necropsy in repeat-dose toxicity studies (Alder and Zbinden, 1988). The extent to which organs are weighed varies between laboratories, but organ weighing is a useful adjunct to macroscopic assessment. Therefore, the selection of organs for weighing is the primary responsibility of the study pathologist. Weighing helps to focus the histopathological examination on key target organs, such as the liver and kidney, the weights of which are frequently altered upon administration of xenobiotics (Peters and Boyd, 1996).

Heart weight is a guide to potential cardiac alterations and especially important in the assessment of cardiovascular drugs. Likewise, the lungs are weighed in inhalation studies as this can provide a useful indication of the extent of edema or accumulation of exudate. Brain weight is employed as a stable reference point in adult animals as it is fairly independent of body weight changes. The weights of endocrine organs are useful guides to alterations in the endocrine status of laboratory animals (Pfeiffer, 1998). However, weighing a small and firmly attached organ such as the thyroid can severely disrupt its quality and orientation in the sections and thus offset any apparent advantage (Michael et al., 2007).

Testicular weights correlate with testicular toxicity, and weights can be compared with in-life measurement of testicular size (Heywood and James, 1978; Creasy, 2003). Weighing the testis is a useful precaution at the early phase of development of a novel drug prior to any assessment of male fertility. By contrast, ovarian weight is highly variable as a consequence of cyclical ovarian development and is therefore a less sensitive indicator of treatment-induced changes in the female reproductive system (Long et al., 2001).

TABLE 15.4 Examples of Basic Tests Applicable to Most Rat, Dog, and Monkey Studies

| Hematology and Coagulation | Clinical Chemistry | Urinalysis |
|--|----------------------------|---|
| RBC count | Glucose | Color and clarity |
| Hemoglobin | Urea nitrogen (or urea) | Overnight volume (e.g., 16h) |
| Hematocrit | Creatinine | Urine specific gravity |
| Mean corpuscular volume | Total protein | Reagent strip test, pH, protein, glucose, ketones, bilirubin, urobilinogen, blood |
| Mean corpuscular hemoglobin | Albumin | Microscopic examination of sediment: cells, casts, crystal, bacteria, sperm |
| Mean corpuscular hemoglobin concentration | Globulin (calculated) | |
| RBC morphology | A/G ratio (calculated) | |
| WBC count | Cholesterol | |
| WBC differential count | Total bilirubin | |
| Platelet count | Alanine aminotransferase | |
| Blood and bone marrow smears | Aspartate aminotransferase | |
| Prothrombin time (PT) | Alkaline phosphatase | |
| Activated partial thromboplastin time (APTT) | Gamma-glutamyltransferase | |
| | Creatine kinase | |
| | Calcium | |
| | Inorganic phosphorus | |
| | Sodium | |
| | Potassium | |
| | Chloride | |

15.2 CLINICAL PATHOLOGY

Clinical pathology is the evaluation of changes (or lack of changes) in the formed blood elements and their characteristics—the most common of these parameters are listed in Table 15.4. These parameters reflect the homeostasis and function of both the hematopoietic system and associated metabolic systems. Beutler et al. (1995) provide an extensive and detailed overview of these systems, though primarily from the perspective of the human system. These measures have the advantage that samples can be taken (and therefore evaluations made) at multiple points over the course of drug administration and at points subsequent to the discontinuation of such administration (i.e., during “recovery”).

At the same time, there is the disadvantage that these are frequently indirect measures of what is happening at specific target sites (primarily the bone marrow). The actual target organ effects can generally only be evaluated after termination of the test animals.

Actual evaluation of meaning and relevance and mechanism of cause of observed changes requires of course consideration of not individual parameters in isolation but rather of the entire set of measures and the relationships/correlations of these changes. This evaluation is discussed in the earlier chapter on “Repeat-Dose Toxicity Studies.”

15.2.1 Clinical Chemistry

One of the portions of the information employed into a pathological evaluation which is not terminally collected comes from samples of blood and urine. These can be analyzed for

the presence and activity of enzyme and endogenous physiologic components (such as electrolytes), including those listed in Table 15.4. Which are collected and how they are analyzed is largely simply following what is done in human beings during clinical evaluation. Some adaptations have been made, and indeed the measurement methods (and their validation) and interpretation are essentially modified for the specific species in question.

The interpretation of these parameters is primarily addressed in the chapter on the repeat-dose studies, though there are some excellent references on the field (Loeb and Quimby, 1999; Lewandrowski, 2003; Burtis et al., 2012). The last of these best addresses modern biomarkers for organ damage. The issue of sampling in nonclinical safety studies is a multifaceted one. First, each species presents limitations on how much (and how often) sample can be drawn. Mice are the most limiting, with large nonrodent species (dogs, primates, and pigs) the least. The proximity of sample collection of the time of drug to dosing or organ damage means that there is a strong emphasis on a need for frequent sample collection. One must not lose sight of the fact that we only see where we look, that is, only where we sacrifice.

15.2.2 Target Organ Toxicity Biomarkers

As this is written, FDA/EMA continues to work on qualifying new biomarkers for target organ toxicity. These serve to significantly improve the performance of nonclinical safety studies in identifying potential drug-related toxicities.

The prototype set is for nephrotoxicity, where the two agencies have qualified a set of seven biomarkers—Kim-1,

albumin, total protein, cystatin C, B2-microglobulin, urinary clustering, and urinary trefoil factor. Two of these (albumin and total protein) have been part of the clinical chemistry parameter set for decades. As with traditional clinical chemistry and clinical pathology measures, these biomarkers should be considered as sets not in isolation.

REFERENCES

- Alder SA, Zbinden G. (1988) *National and International Drug Safety Guidelines*. M.T.C. Verlag Zollikon, Zollikon.
- Arnold DL, Dice HC, Krewski DR. (1990) *Handbook of In-Vivo Toxicity Testing*. Academic Press, San Diego, CA.
- Beutler E, Lichtman MA, Collier BS, Kipps TJ. (1995) *Williams Hematology*, 5th Ed. McGraw Hill, New York.
- Bregman CI, Adler RR, Morton DG, Regan KS, Yano BL. (2003) Recommended tissue list for histopathologic examination in repeat-dose toxicity and carcinogenicity studies: a proposal of the Society of Toxicologic Pathology (STP). *Toxicol. Pathol.* 31(2):252–253.
- Burtis CA, Ashwood ER, Brunis DE. (2012) *Tietz Textbook of Clinical Chemistry and Molecular Diagnostics*, 5th Ed. Elsevier Saunders, Philadelphia.
- Creasy DM. (2003) Evaluation of testicular toxicology: a synopsis and discussion of the recommendations proposed by the society of toxicologic pathology. *Birth Defects Res. Part B Develop. Reprod. Toxicol.* 68:408–415.
- Dragen JM. (2005) COX-2 inhibitors—a lesson in unexpected problems. *N. Engl. J. Med.* 352:1131–1132.
- EMA. (2000) *The European Agency for the Evaluation of Medicinal Products Committee for Proprietary Medicinal Products, Note for Guidance on Repeated Dose Toxicity*. Available at: http://www.ema.europa.eu/docs/en_GB/document_library/Scientific_guideline/2009/09/WC500003102.pdf (accessed May 4, 2016).
- European Economic Community (EEC). (1992) *Toxicology and Chemical Substances Guidelines*. European Chemicals Bureau.
- Frame SR, Mann PC, Caverly Rae JM. (2014) Principles of pathology for toxicology studies. In: *Principles and Methods of Toxicology* (Hayes AW, Ed.), 6th Ed. CRC Press, Boca Raton, FL, pp. 571–596.
- Gad SC, Rousseaux CG. (2013) Use and misuse of statistics in toxicologic pathology. In: *Handbook of Toxicologic Pathology* (Haschek WM and Rousseaux CG, Eds.), 3rd Ed. Academic Press, San Diego, CA, pp. 327–418.
- Geiling EMK, Cannon PR. (1938) Pathological effects of elixir and sulphanilamide (diethylene, glycol) poisoning. *JAMA.* 111:919–926.
- Gray P. (1964) *Handbook of Basic Microtechnique*. McGraw Hill Book Company, New York.
- Greaves P. (2009) Pathological techniques in toxicology. In: *General and Applied Toxicology* (Ballantyne B, Marrs T and Synersen T, Eds.), 3rd Ed. John Wiley & Sons, Ltd, Chichester, pp. 691–706.
- Greaves P. (2012) *Histopathology of Preclinical Toxicity Studies*, 4th Ed. Academic Press, San Diego, CA.
- Greaves P, Williams A, Ever M. (2004) First dose of potential medicines to humans; how animals help. *Nat. Rev. Drug Discov.* 3:226–236.
- Ham AW. (1974) *Histology*, 7th Ed. J.B. Lippincott Co., Philadelphia, PA.
- Haschek WM, Rousseaux CG, Wallig MA. (Eds.) (2013a) *Handbook of Toxicologic Pathology*, 3rd Ed. Academic Press, San Diego, CA.
- Haschek WM, Rousseaux CG, Wallig MA. (2013b) Toxicologic pathology: an introduction. In: *Haschek and Rousseaux's Handbook of Toxicologic Pathology* (Haschek WM, Rousseaux CG and Wallig MA, Eds.), 3rd Ed. Academic Press, San Diego, CA.
- Heywood R, James RW. (1978) Assessment of testicular toxicity in lab animals. *Environ. Health Persp.* 24:73–80.
- ICH. (2000) *International Conference of Harmonization of Technical Requirements for Registration of Pharmaceuticals for Human Use (ICH) Topic S5(R2), Detection of Toxicity to Reproduction for Medicinal Products and Toxicity to Male Fertility*. Available at: http://www.ich.org/fileadmin/Public_Web_Site/ICH_Products/Guidelines/Safety/S5/Step4/S5_R2_Guideline.pdf (accessed May 4, 2016).
- ICH. (2005) *International Conference of Harmonization of Technical Requirements for Registration of Pharmaceuticals for Human Use (ICH) Topic S8 Immunotoxicity Studies for Human Pharmaceuticals, Note for Guidance on Immunotoxicity Studies for Human Pharmaceuticals*. Available at: http://www.ich.org/fileadmin/Public_Web_Site/ICH_Products/Guidelines/Safety/S8/Step4/S8_Guideline.pdf (accessed May 4, 2016).
- Lewandowski K. (2003) *Clinical Chemistry*. Lippincott, Williams and Wilkins, Philadelphia, PA.
- Lewis RW, Billington R, Debryune E, Gamer A, Lang B, Carpanini F. (2002) Recognition of adverse and nonadverse effects in toxicity studies. *Toxicol. Pathol.* 30(1):66–74.
- Loeb WF, Quimby FW. (1999) *The Clinical Chemistry of Laboratory Animals*, 2nd Ed. Taylor & Francis, Philadelphia, PA.
- Long GG, Cohen IR, Gries CL, Young JK, Francis PC, Capen CC. (2001) Proliferative lesions of ovarian granulosa cells and reversible hormonal changes induced in rats by a selective estrogen receptor modulator. *Toxicol. Pathol.* 29(4):403–410.
- McCance KL, Huether SE. (2015) *Pathophysiology*, 7th Ed. Elsevier, San Diego, CA.
- Michael B, Yano B, Sellers RS, Perry R, Morton D, Roome N, Johnson JK, Schafer K. (2007) Evaluation of organ weights for rodent and non-rodent toxicity studies: a review of regulatory guidelines and a survey of current practices. *Toxicol. Pathol.* 35:742–750.
- Mikaelian I, Nanney LB, Parman KS, Kusewitt DF, Ward JM, Näf D, Krupke DM, Eppig JT, Bult CJ, Seymour R, Ichiki T, Sundberg JP. (2004) Antibodies that label paraffin-embedded mouse tissues: a collaborative endeavor. *Toxicol. Pathol.* 32:181–191.
- Ministry of Health, Labour and Welfare, Pharmaceutical Affairs Bureau, New Drugs Division. (1990) *Guidelines for Toxicity*

- Studies of Drugs Manual*. Ministry of Health, Labour and Welfare, Tokyo.
- Monto AS. (1999) The disappearance of Reyes syndrome—a public health triumph. *N. Engl. J. Med.* 340:1423–1424.
- OECD. (n.d.) *Guideline for the Testing of Chemicals*. Organization for Economic Cooperation and Development Environment Directorate (OECD), Paris. Available at: <http://www.oecd.org/env/ehs/oecdguidelinesforthetestingofchemicalsandrelateddocuments.htm>.
- Olsen H, Betton G, Robinson D, Thomas K, Monro A, Kolaja G, Lilly P, Sanders J, Sipes G, Bracken W, Dorato M, Van Deun K, Smith P, Berger B, Heller A. (2000) Concordance of the toxicity of pharmaceuticals of humans and animals. *Regul. Toxicol. Pharmacol.* 32:56–67.
- Orfila MJB. (1815) *Traite des poisons Tires des regens mineral, vegetal et animal, on toxicologic generale, connidlhee sons les repports de la physiologic, de la pathologic et de la medicine legale*. Crochard, Paris.
- Perez DH, Barthold SW. (2007) *Pathology of Laboratory Rodents and Rabbits*, 3rd Ed. Blackwell Publishing, Ames, IA.
- Peters JM, Boyd EM. (1996) Organ weights and water levels of the rat following reduced food intake. *J. Nutr.* 90(4):345–360.
- Pfeiffer CJ. (1998) A mathematical evaluation of the thymic weight parameter. *Toxicol. Appl. Pharmacol.* 13(2):220–227.
- Roy D, Andrews PA. (2004) Nonclinical testing: from theory to practice. In: *Anticancer Drug Development Guide* (Teicher BA, and Andrews PA, Eds.). Humana Press, Totowa, NJ, p. 306.
- Tehounwan PB, Centeno JA. (2008) Toxicologic pathology. In: *Preclinical Development Handbook: Toxicology* (Gad SC, Ed.). John Wiley & Sons, Inc., Hoboken, NJ, pp. 551–580.
- Turton J, Hoasen J. (1998) *Target Organ Pathology*. Taylor & Francis, Philadelphia, PA.
- Wallig MA. (2002) Mophologic manifestations of toxic cell Injury. In: *Handbook of Toxicologic Pathology* (Haschek WM, Rousseaux CG and Wallig MA, Eds.). Academic Press, New York, pp. 39–65.

IRRITATION AND LOCAL TISSUE TOLERANCE IN PHARMACEUTICAL SAFETY ASSESSMENT

16.1 INTRODUCTION

Local tissue tolerance (and their subset, irritation) studies assess the short-term hazard of pharmaceutical agents in the immediate region of their application, administration, or installation. In particular, these studies are done (expected) to assess topically or parenterally administered drug formulations. Note that these are all hazard tests, properly performed using the intended clinical formulation.

Topical local tolerance effects are almost entirely limited to irritation. Though this usually means dermal irritation, it can also be intracutaneous, mucosal, penile, perivascular, vaginal, bladder, rectal, nasal, or ocular, depending on the route of drug administration. All but ocular irritation uses some version of a common subjective rating scale (see Table 16.1) to evaluate responses. The outcome of all of these tests primarily evaluates the response of the first region of tissue (which is exposed to the highest concentration) to an administered drug substance. In general, any factor which enhances absorption through the contacted tissue is likely to decrease tissues tolerance. Wilhelm et al. (2012) and Gad and Chengelis (1998) should be referred to for a more detailed coverage of the subject of topical tissue toxicology.

For the skin, this scale is used in the primary dermal irritation (PDI) test, which is performed for those agents that are to be administered to patients by application to the skin. As with all local tolerance tests, it is essential that the material be evaluated in “condition of use”—that is, in the final formulated form, applied to test animals in the same manner that the agent is to be used clinically.

16.2 FACTORS AFFECTING IRRITATION RESPONSES AND TEST OUTCOME

The results of local tissue irritation tests are subject to considerable variability due to relatively small differences in test design or technique. Weil and Scala (1971) arranged and reported on the best known of several intralaboratory studies to clearly establish this fact. Though the methods presented previously have proven to give reproducible results in the hands of the same technicians over a period of years (Gad et al., 1986) and contain some internal controls (the positive and vehicle controls in the PDI) against large variabilities in results or the occurrence of either false positives or negatives, it is still essential to be aware of those factors that may systematically alter test results. These factors are summarized in the following:

- A. In general, any factor that increases absorption through the stratum corneum or mucous membrane will also increase the severity of an intrinsic response. Unless this factor mirrors potential exposure conditions, it may, in turn, adversely affect the relevance of test results.
- B. The physical nature of solids must be carefully considered both before testing and interpreting results. Shape (sharp edges), size (small particles may abrade the skin due to being rubbed back and forth under the occlusive wrap), and rigidity (stiff fibers or very hard particles will be physically irritating) of solids may all enhance an irritation response.
- C. Solids frequently give different results when they are tested dry than if wetted for the test. As a general rule,

TABLE 16.1 Evaluation of Local Tissue Reactions in Tissue Irritation Studies

| Skin Reaction | Value |
|--|-------|
| Erythema and eschar formation | |
| No erythema | 0 |
| Very slight erythema (barely perceptible) | 1 |
| Well-defined erythema | 2 |
| Moderate to severe erythema | 3 |
| Severe erythema (beet redness) to slight eschar formation (injuries in depth) | 4 |
| Necrosis (death of tissue) | +N |
| Eschar (sloughing or scab formation) | +E |
| Edema formation | |
| No edema | 0 |
| Very slight edema (barely perceptible) | 1 |
| Slight edema (edges of area well defined by definite raising) | 2 |
| Moderate edema (raised ~1 mm) | 3 |
| Severe edema (raised more than 1 mm and extending beyond the area of exposure) | 4 |
| Total possible score for primary irritation | 8 |

solids are more irritating if moistened (going back to Item A, wetting is a factor that tends to enhance absorption). Care should also be taken as to moistening agent—some (few) batches of US Pharmacopeia physiological saline (used to simulate sweat) have proven to be mildly irritating to the skin and mucous membrane on their own. Liquids other than water or saline should not be used.

- D. If the treated region on potential human patients will be a compromised skin surface barrier (e.g., if it is cut or burned), some test animals should likewise have their application sites compromised. This procedure is based on the assumption that abraded skin is uniformly more sensitive to irritation. Experiments, however, have shown that this is not necessarily true; some materials produce more irritation on abraded skin, while others produce less (Guillot et al., 1982; Gad et al., 1986).
- E. The degree of occlusion (in fact, the tightness of the wrap over the test site) also alters percutaneous absorption and therefore irritation. One important quality control issue in the laboratory is achieving a reproducible degree of occlusion in dermal wrappings.
- F. Both the age of the test animal and the application site (saddle of the back vs. flank) can markedly alter test outcome. Both of these factors are also operative in humans, of course (Mathias, 1983), but in dermal irritation tests, the objective is to remove all such sources of variability. In general, as an animal ages, sensitivity to irritation decreases. For the dermal test, the skin middle of the back (other than directly over the spine)

tends to be thicker (and therefore less sensitive to irritations) than that on the flanks.

- G. The sex of the test animals can also alter study results because both regional skin thickness and surface blood flow vary between males and females.
- H. Finally, the single most important (yet also most frequently overlooked) factor that influences the results and outcome of these (and, in fact most) acute studies is the training of the staff. In determining how test materials are prepared and applied and in how results are “read” against a subjective scale, both accuracy and precision are extremely dependent on the technicians involved. To achieve the desired results, initial training must be careful and all-inclusive. As important, some form of regular refresher training must be exercised—particularly in the area of scoring results. The use of a set of color photographic standards as a training and reference tool is strongly recommended; such standards should clearly demonstrate each of the grades in the Draize dermal scale.
- I. It should be recognized that the dermal irritancy test is designed with a bias to preclude false negatives and, therefore, tends to exaggerate results in relation to what would happen in humans. Findings of negligible irritancy (or even in the very low mild irritant range) should therefore be of no concern unless the product under test is to have large-scale and prolonged dermal contact.

16.3 PRIMARY DERMAL IRRITATION (PDI) TEST

The prototypical test to use to illustrate the principles underlying all other irritation tests is the PDI. It looks at the potential for a single exposure of a chemical to the skin to cause inflammation (erythema and edema) or damage to that skin. It uses the albino white rabbit as a model and a subjective (but reproducible) classification style grading system to score results.

A. Rabbit Screening Procedure

1. A group of at least 4–6 New Zealand white rabbits are screened for the study.
2. All rabbits selected for the study must be in good health; any rabbit exhibiting sniffles, hair loss, loose stools, or apparent weight loss is rejected and replaced.
3. One day (at least 18 h) prior to application of the test substance, each rabbit is prepared by clipping the hair from the back and sides using a small animal clipper. A size No. 10 blade is used to remove long hair, and then a size No. 40 blade is used to remove the remaining hair.

4. Six animals with skin sites that are free from hyperemia or abrasion (due to shaving) are selected. Skin sites that are in the telogen phase (resting stage of hair growth) are used; those skin sites that are in the anagen phase (stage of active growth, indicated by the presence of a thick undercoat of hair) are not used.

B. Study Procedure

1. As many as four areas of skin, two on each side of the rabbit's back, can be utilized for sites for administration.
2. Separate animals are not required for an untreated control group. Each animal serves as its own control.
3. Besides the test substance, a positive control substance (a known skin irritant—1% sodium lauryl sulfate in distilled water) and a negative control (untreated patch) are applied to the skin. When a vehicle is used for diluting, suspending, or moistening the test substance, a vehicle control patch is required—especially if the vehicle is known to cause any toxic dermal reactions or if there is insufficient information about the dermal effects of the vehicle.
4. The intact (free of abrasion) sites of administration are assigned a code number. Up to four sites can be used, as follows:
 - #1. Test substance
 - #2. Negative control
 - #3. Positive control
 - #4. Vehicle control (if required)
5. Application sites should be rotated from one animal to the next to ensure that the test substance and controls are applied to each position at least once.
6. Each test or control substance is held in place with a 1 in. × 1 in. 12-ply surgical gauze patch. The gauze patch is applied to the appropriate skin site and secured with 1 in.-wide strips of surgical tape at the four edges, leaving the center of the gauze patch nonoccluded.
7. If the test substance is a solid or a semisolid, a 0.5 g portion is weighed and placed on the gauze patch. The test substance patch is placed on the appropriate skin site and secured. The patch is subsequently moistened with 0.5 mL of physiological saline.
8. When the test substance is in flake, granule, powder, or other particulate form, the weight of the test substance that has a volume of 0.5 mL (after compacting as much as possible without crushing or altering the individual particles, such as by tapping the measuring container) is used whenever this volume is less than 0.5 g. When applying powders, granules, and the like, the gauze patch designated for the test sample is secured to the appropriate skin site with one of the four strips of the tape at the most ventral position of the animal. With one hand, the appropriate amount of sample measuring 0.5 mL is carefully poured from a glycine weighing paper onto the gauze patch that is held in a horizontal (level) position with the other hand. The patch containing the test sample is then carefully placed into position onto the skin and the remaining three edges secured with tape. The patch is subsequently moistened with 0.5 mL of physiological saline.
9. If the test substance is a liquid, a patch is applied and secured to the appropriate skin site. A 1 mL tuberculin syringe is used to measure and apply 0.5 mL of test substance to the patch.
10. The negative control site is covered with an untreated 12-ply surgical gauze patch (1 in. × 1 in.).
11. The positive control substance and vehicle control substance are applied to a gauze patch in the same manner as a liquid test substance.
12. The entire trunk of the animal is covered with an impervious material (such as Saran Wrap) for a 24 h period of exposure. The Saran Wrap is secured by wrapping several long strips of athletic adhesive tape around the trunk of the animal. The impervious material aids in maintaining the position of the patches and retards evaporation of volatile test substances.
13. An Elizabethan collar is fitted and fastened around the neck of each test animal. The collar remains in place for the 24 h exposure period. The collars are utilized to prevent removal of wrappings and patches by the animals while allowing the animals' food and water *ad libitum*.
14. The wrapping is removed at the end of the 24 h exposure period. The test substance skin site is wiped to remove any test substance still remaining. When colored test substances (such as dyes) are used, it may be necessary to wash the test substance from the test site with an appropriate solvent or vehicle (one that is suitable for the substance being tested). This is done to facilitate accurate evaluation for skin irritation.
15. Immediately after removal of the patches, each 1 in. × 1 in. test or control site is outlined with indelible marker by dotting each of the four corners. This procedure delineates the site for identification.

C. Observations

1. Observations are made of the test and control skin sites 1 h after removal of the patches (25 h postinitiation of application). Erythema and edema are evaluated and scored on the basis of designated values presented earlier in Table 16.1.
2. Observations are again performed 48 and 72 h after application and scores are recorded.

3. If necrosis is present or the dermal reaction is unusual, the reaction should be described. Severe erythema should receive the maximum score (4), and +N should be used to designate the presence of necrosis and +E the presence of eschar.
4. When a test substance produces dermal irritation that persists 72 h postapplication, daily observations of test and control sites are continued on all animals until all irritation caused by the test substance resolves or until day 14 postapplication.

D. Evaluation of Results

1. A *subtotal irritation value* for erythema or eschar formation is determined for each rabbit by adding the values observed at 25, 48, and 72 h postapplication.
2. A *subtotal irritation value* for edema formation is determined for each rabbit by adding the values observed at 25, 48, and 72 h postapplication.
3. A *total irritation value* is calculated for each rabbit by adding the subtotal irritation value for erythema or eschar formation to the subtotal irritation value for edema formation.
4. The *primary dermal irritation index* (PDII) is calculated for the test substance or control substance by dividing the sum of the total irritation scores by the number of observations (3 days \times 3 animals = 9 observations).
5. The categories of the PDII are as follows (this categorization of dermal irritation is a modification of the original classification described by Draize et al. (1944)):

| | |
|----------|---------------------|
| PDII=0.0 | Nonirritant |
| >0.0–0.5 | Negligible irritant |
| >0.5–2.0 | Mild irritant |
| >2.0–5.0 | Moderate irritant |
| >5.0–8.0 | Severe irritant |

Other abnormalities, such as atonia or desquamation, should be noted and recorded.

16.4 OTHER NONPARENTERAL ROUTE IRRITATION TESTS

The design of vaginal, rectal, and nasal irritation studies is less formalized but follows the same basic pattern as the PDI test. The rabbit is the preferred species for vaginal and rectal irritation studies, but the monkey and dog have also been used for these (Eckstein et al., 1969; Gad and Chengelis, 1998). Both the rabbit and rat have commonly seen use for nasal irritation evaluations. Defined quantities (typically

1.0 mL) of test solutions or suspensions are instilled into the orifice in question. For the vagina or rectum inert, bungs are usually installed immediately thereafter to continue exposure for a defined period of time (usually the same period of hours as future human exposure). The orifice is then flushed clean, and 24 h after exposure, it is examined and evaluated (graded) for irritation using the scale in Table 16.1.

16.5 OCULAR IRRITATION TESTING

Ocular irritation is significantly different from the other local tissue irritation tests on a number of grounds. For the pharmaceutical industry, eye irritation testing is performed when the material is intended to be put into the eye as a means or route of application for ocular therapy (Chan and Hayes, 2014). There are a number of special tests applicable to pharmaceuticals or medical devices that are beyond the scope of this volume, since they are not intended to assess potential acute effects or irritation. In general, however, it is desired that an eye irritation test that is utilized by this group be both sensitive and accurate in predicting the potential to cause irritation in humans. Failing to identify human ocular irritants (lack of sensitivity) is to be avoided, but of equal concern is the occurrence of false positives.

The primary eye irritation test was originally intended to predict the potential for a single splash of chemical into the eye of a human being to cause reversible and/or permanent damage. Since the introduction of the original Draize test more than 60 years ago (Draize et al., 1944), ocular irritation testing in rabbits has both developed and diverged. Indeed, clearly there is no longer a single test design that is used, and different objectives are pursued by different groups using the same test. This lack of standardization has been recognized for some time, and attempts have been made to address standardization of at least the methodological aspects of the test, if not the design aspects.

One widely used study design, which begins with a screening procedure as an attempt to avoid testing severe irritants or corrosives in animals, goes as follows:

A. Test Article Screening Procedure

1. Each test substance will be screened in order to eliminate potentially corrosive or severely irritating materials from being studied for eye irritation in the rabbit.
2. If possible, the pH of the test substance will be measured.
3. A PDI test will be performed prior to the study.
4. The test substance will not be studied for eye irritation if it is a strong acid (pH of 2.0 or less) or strong alkali (pH of 12.0 or greater) and/or if the test substance is a severe dermal irritant (with a PDII of 5–8) or causes corrosion of the skin.

5. If it is predicted that the test substance does not have the potential to be severely irritating or corrosive to the eye, continue to Section B, Rabbit Screening Procedure.

B. Rabbit Screening Procedure

1. A group of at least six New Zealand white rabbits of either sex are screened for the study. The animals are removed from their cages and placed in rabbit restraints. Care should be taken to prevent mechanical damage to the eye during this procedure.
2. All rabbits selected for the study must be in good health; any rabbit exhibiting snuffles, hair loss, loose stools, or apparent weight loss is rejected and replaced.
3. One hour prior to installation of the test substance, both eyes of each rabbit are examined for signs of irritation and corneal defects with a handheld slit lamp. All eyes are stained with 2.0% sodium fluorescein and examined to confirm the absence of corneal lesions. *Fluorescein staining:* Cup the lower lid of the eye to be tested and instill one drop of 2% (in water) sodium fluorescein solution onto the surface of the cornea. After 15 s, thoroughly rinse the eye with physiological saline. Examine the eye, employing a handheld long-wave ultraviolet (UV) illuminator in a darkened room. Corneal lesions, if present, appear as bright yellowish-green fluorescent areas.
4. Only three of the six animals are selected for the study. The three rabbits must not show any signs of eye irritation and must show either a negative or minimum fluorescein reaction (due to normal epithelial desquamation).

C. Study Procedure

1. At least 1 h after fluorescein staining, the test substance is placed in one eye of each animal by gently pulling the lower lid away from the eyeball to form a cup (conjunctival cul-de-sac) into which the test material is dropped. The upper and lower lids are then gently held together for 1 s to prevent immediate loss of material.
2. The other eye remains untreated and serves as a control.
3. For testing liquids, 0.01 mL of the test substance is used.
4. For solids or pastes, 100 mg of the test substance is used.
5. When the test substance is in flake, granular, powder, or other particulate form, the amount that has a volume of 0.01 mL (after gently compacting the particles by tapping the measuring container in a way that will not alter their individual form) is used whenever this volume weighs less than 10 mg.

6. For aerosol products, the eye should be held open and the substance administered in a single 1-s burst at a distance of about 4 in. directly in front of the eye. The velocity of the ejected material should not traumatize the eye. The dose should be approximated by weighing the aerosol can before and after each treatment. For other liquids propelled under pressure, such as substances delivered by pump sprays, an aliquot of 0.01 mL should be collected and instilled in the eye as for liquids.
7. The treated eyes of the three rabbits are not washed following the instillation of the test substance.
8. The treated eyes of the remaining three rabbits are irrigated for 1 min with room-temperature tap water, starting 20 s after instillation.
9. To prevent self-inflicted trauma by the animals immediately after instillation of the test substance, the animals are not immediately returned to their cages. After the test and control eyes are examined and graded at 1 h postexposure, the animals are returned carefully to their respective cages.

D. Observations

1. The eyes are observed for any immediate signs of discomfort after instilling the test substance. Blepharospasm and/or excessive tearing are indicative of irritating sensations caused by the test substance, and their duration should be noted. Blepharospasm does not necessarily indicate that the eye will show signs of ocular irritation.
2. Grading and scoring of ocular irritation are performed in accordance with Table 16.2. The eyes are examined and grades of ocular reactions are recorded. To aid in the standardization of grading, the CPSC published an illustrated guide in 1976 (CPSC, 1976), which is reproduced in Gad and Chengelis (1998).
3. If signs of irritation persist at day 7, readings are continued on days 10 and 14 after exposure or until all signs of reversible toxicity are resolved.
4. In addition to the required observation of the cornea, iris, and conjunctiva, serious effects (such as panus, rupture of the globe, or blistering of the conjunctivae) indicative of a corrosive action are reported.
5. Whether or not toxic effects are reversible depends on the nature, extent, and intensity of damage. Most lesions, if reversible, will heal or clear within 21 days. Therefore, if ocular irritation is present at the 14-day reading, a 21-day reading is required to determine whether the ocular damage is reversible or nonreversible.

TABLE 16.2 Scale of Weighted Scores for Grading the Severity of Ocular Lesions^a

| Reaction Criteria | Score |
|--|-------|
| I. Cornea | |
| A. Opacity degree of density (area that is most dense is taken for reading) | |
| 1. Scattered or diffuse area, details of iris clearly visible | 1 |
| 2. Easily discernible translucent area, details of iris slightly obscured | 2 |
| 3. Opalescent areas, no details of iris visible, size of pupil barely discernible | 3 |
| B. Area of cornea involved | |
| 1. One-quarter (or less) but not zero | 1 |
| 2. Greater than one-quarter, less than one-half | 2 |
| 3. Greater than one-half, less than whole area | 3 |
| 4. Greater than three-quarters up to whole area | 4 |
| Scoring equals $A \times B \times 5$; total maximum = 80 ^b | |
| II. Iris | |
| A. Values | |
| 1. Folds above normal, congestion, swelling, circumcorneal ingestion (any one or all of these or combination of any thereof), iris still reacting to light (sluggish reaction is possible) | 1 |
| 2. No reaction to light, hemorrhage, gross destruction (any one or all of these) | 2 |
| Scoring equals $A \times B$ (where B is the area of the iris involved, graded as “under cornea”); total maximum = 10 | |
| III. Conjunctivae | |
| A. Redness (refers to palpebral conjunctivae only) | |
| 1. Vessels definitely injected above normal | 1 |
| 2. More diffuse, deeper crimson red, individual vessels not easily discernible | 2 |
| 3. Diffuse beefy red | 3 |
| B. Chemosis | |
| 1. Any swelling above normal (includes nictitating membrane) | 1 |
| 2. Obvious swelling with partial eversion of the lids | 2 |
| 3. Swelling with lids about half closed | 3 |
| 4. Swelling with lids about half closed to completely closed | 4 |
| C. Discharge | |
| 1. Any amount different from normal (does not include small amount observed in inner canthus of normal animals) | 1 |
| 2. Discharge with moistening of the lids and hair just adjacent to the lids | 2 |
| 3. Discharge with moistening of the lids and considerable area around the eye | 3 |
| Scoring $(A + B + C) \times 2$; total maximum = 20 | |

^a The maximum total score is the sum of all scores obtained for the cornea, iris, and conjunctivae.

^b All $A \times B = \Sigma (1 - 3) \times \Sigma (1 - 4)$ for three animals.

More than three decades of work have been done on efforts to develop, validate, and accept alternatives to the *in vivo* eye irritation test and to modify it to use fewer animals (successful here in reducing the number of rabbits from 9 to 3) and reduce potential discomfort to test animals (also successful—see Gad et al., 1986, which proposed the current OECD schema for using a sequential method to avoid the severe and marked ocular irritation outcome in rabbits without exposing humans to potential harm). Gad (2000, 2009) survey these efforts.

Starting in 2015, at least one division of the FDA has asked for the *in vitro* bovine corneal opacity and permeability (BOCP) test (Gautheron et al., 2004) rather than the traditional *in vivo* study.

16.6 VAGINAL IRRITATION

Few, if any, products are administered via the vagina that are intended for systemic absorption. Thus, this route has not been as widely studied and characterized as others. On the

other hand, large numbers of different products (douches, spermicides, antiyeast agents, etc.) have been developed that require introduction into the vagina in order to assert their localized effects. Increased research into different birth control and antiviral prophylaxis will result in more vaginal products in the future. All these must be assessed for vaginal irritation potential, and this serves as an example of the other tissue tolerance issues.

Considerable research (Eckstein et al., 1969; Auletta, 1994) has indicated that the rabbit is the best species for assessing vaginal irritation. There are those investigators, however, who consider the rabbit too sensitive and recommend the use of ovariectomized rats. Ovariectomy results in a uniformly thin, uncornified epithelium which is more responsive to localized effects. This model is used when the results from a study with rabbits are questionable (Auletta, 1994). The routine progression of studies consists of first doing an acute primary vaginal irritation study, then a 10-day repeated-dose study in rats. These protocols are summarized

later. Longer-term vaginal studies have been conducted in order to assess systemic toxicity of the active agents when administered by these routes (while the intended effects may be local, one cannot assume that there will be no systemic exposure).

16.7 ACUTE PRIMARY VAGINAL IRRITATION STUDY IN THE FEMALE RABBIT

This study is used to assess tolerance for drugs (or devices) placed in the vagina to achieve either local or systemic exposure.

1. *Overview of study design* One group of three adult rabbits received a single vaginal exposure and observed for 3 days with periodic examination (1, 24, 48 and 72 h postdosing) of the genitalia. Animals are then euthanized and the vagina is examined macroscopically.

2. *Administration*

Route The material is generally introduced directly into the vagina using a lubricated 18 French rubber catheter attached to a syringe for quantification and delivery of the test material. Gentle placement of the catheter is important because one needs to ensure complete delivery of the dose without mechanical trauma. For rabbits, the depth of insertion is about 7.5 cm, and the catheter should be marked to about that depth. After delivery is completed, the tube is slowly withdrawn. No attempt is made to close or seal the vaginal orifice. Alternative methods may be used to administer more viscous materials. The most common is to backload a lubricated 1 mL tuberculin syringe and then warm the material close to the body temperature. The syringe is then inserted into the vagina and the dose administered by depressing the syringe plunger.

Dosage The test material should be the one (concentration, vehicle, etc.) that is intended for human application.

Frequency Once.

Duration 1 day.

Volume 1 mL per rabbit.

3. *Test system*

Species, age, and weight range Sexually mature New Zealand white rabbits are generally used, weighing between 2 and 5 kg. The weight is not as important as the fact that the animals need to be sexually mature.

Selection Animals should be multiparous and non-pregnant. Animals should be healthy and free of external genital lesions.

TABLE 16.3 Scoring Criteria for Vaginal Irritation

| Value | |
|------------------|--|
| <i>Erythema</i> | |
| 0 | No erythema |
| 1 | Very slight erythema (barely perceptible) |
| 2 | Slight erythema (pale red in color) |
| 3 | Moderate to severe erythema (definite red in color) |
| 4 | Severe erythema (beet or crimson red) |
| <i>Edema</i> | |
| 0 | No edema |
| 1 | Very slight edema (barely perceptible) |
| 2 | Slight edema (edges of area well defined by definite raising) |
| 3 | Moderate edema (raised ~1 mm) |
| 4 | Severe edema (raised more than 1 mm and extending beyond area of exposure) |
| <i>Discharge</i> | |
| 0 | No discharge |
| 1 | Very slight discharge |
| 2 | Slight discharge |
| 3 | Moderate discharge |
| 4 | Severe discharge (moistening of considerable area around the vagina) |

Randomization Because there is only one group of animals, randomization is not a critical issue.

4. *In-life observations*

Daily observations At least once daily for clinical signs.

Detailed physical examination Once during the week prior to dosing.

Body weight Day of dosing.

Vaginal irritation Scored at 1, 24, 48, and 72 h post-dosing. Scoring criteria are shown in Table 16.3.

5. *Postmortem procedures* Rabbits are euthanized by lethal dose of a barbiturate soon after the last vaginal irritation scores are collected. The vagina is opened by longitudinal section and examined for evidence of mucosal damage such as erosion, localized hemorrhage, etc. No other tissues are examined. No tissues are collected. After the macroscopic description of the vagina is recorded, the animal is discarded.

16.7.1 Repeated-Dose Vaginal Irritation in the Female Rabbit

If there is to be a repeated exposure/administration by the vaginal route, a repeat-exposure study should be conducted to assess cumulative effects

1. *Overview of study design* Four groups of three adult rabbits each receive a single vaginal exposure daily for 10 days. The genitalia are examined daily. Animals are then euthanized and the vagina is examined macroscopically and microscopically.

2. *Administration*

Route The test materials are introduced directly into the vagina using a lubricated 18 French rubber catheter using the techniques described previously for acute studies.

Dosage The test material should be the one (concentration, vehicle, etc.) that is intended for human application. There will also be a sham-negative control (catheter in place but nothing administered), a vehicle control, and a positive control (generally 2% nonoxynol-9).

Frequency Once daily.

Duration 10 days.

Volume 1 mL per rabbit for each material.

3. *Test system*

Species, age, and weight range Sexually mature New Zealand white rabbits are generally used, weighing between 2 and 5 kg. The weight is not as important as the fact that the animals need to be sexually mature.

Selection Animals should be nulliparous and non-pregnant. Animals should be healthy and free of external genital lesions.

Randomization At least 14 animals should be on pre-test. Randomization to treatment groups is best done using a computerized blocking by body weight method or a random number generation method.

4. *In-life observations*

Daily observations At least once daily for clinical signs.

Detailed physical examination Once during the week prior to dosing and immediately prior to necropsy.

Body weight First, fifth, and last day of dosing.

Vaginal irritation Scored once daily. Scoring criteria shown in Table 16.3.

5. **Postmortem procedures** Rabbits are euthanized by lethal dose of a barbiturate soon after the last vaginal irritation scores are collected. The vagina is isolated using standard prosection techniques and then opened by longitudinal section and examined for evidence of mucosal damage such as erosion, localized hemorrhage, etc. No other tissues are examined. The vagina and cervix are collected and fixed in 10% neutral buffered formalin. Standard hematoxylin/eosin-stained, paraffin-embedded histologic glass slides are prepared by routing methods. Three levels of the vagina (low, mid, and upper) are examined and graded using the scoring system shown in Table 16.4. Each level is

TABLE 16.4 Microscopic Scoring Procedure for Vaginal Sections

| | Value |
|---|-------|
| <i>Epithelium</i> | |
| Intact—normal | 0 |
| Cell degeneration or flattening of the epithelium | 1 |
| Metaplasia | 2 |
| Focal erosion | 3 |
| Erosion or ulceration, generalized | 4 |
| <i>Leukocytes</i> | |
| Minimal—<25 per high-power field | 1 |
| Mild—25–50 per high-power field | 2 |
| Moderate—50–100 per high-power field | 3 |
| Marked—>100 per high-power field | 4 |
| <i>Injection</i> | |
| Absent | 0 |
| Minimal | 1 |
| Mild | 2 |
| Moderate | 3 |
| Marked with disruption of vessels | 4 |
| <i>Edema</i> | |
| Absent | 0 |
| Minimal | 1 |
| Mild | 2 |
| Moderate | 3 |
| Marked | 4 |

Source: Data from Eckstein et al. (1969).

cored separately and an average is calculated. Irritation is rated as follows:

| Score | Rating |
|-------|---------------------|
| 0 | Nonirritating |
| 1–4 | Minimal irritation |
| 5–8 | Mild irritation |
| 9–11 | Moderate irritation |
| 12–16 | Marked irritation |

The score for each rabbit is then averaged and acceptability ratings are given as follows:

| Average Score | Acceptability Ratings |
|---------------|-----------------------|
| 0–8 | Acceptable |
| 9–10 | Marginal |
| 11 or greater | Unacceptable |

16.7.2 Repeated-Dose Vaginal Irritation in the Ovariectomized Rats

This study is very similar in design to that described previously for rabbits, with the following (sometimes obvious) exceptions. Mature ovariectomized female rats can be obtained from a commercial breeder. A 15% surplus should be

obtained. Ten animals per group should be used (40 total for the study). The vaginal catheter is placed to a depth of approximately 2.5 cm, and the treatment volume should be 0.2 mL.

16.8 PARENTERAL IRRITATION/TOLERANCE

There are a number of special concerns about the safety of materials that are routinely injected (parenterally administered) into the body. By definition, these concerns are all associated with materials that are the products of the pharmaceutical and (in some minor cases) medical device industries. Such parenteral routes include three major ones—intravenous (IV), intramuscular (IM), and subcutaneous (SC)—and a number of minor routes (such as intra-arterial) that are not considered here.

16.8.1 Parenteral Routes

There are at least 13 different routes by which to inject material into the body, including the following:

1. Intravenous
2. Subcutaneous
3. Intramuscular
4. Intra-arterial
5. Intradermal
6. Intralesional
7. Epidural
8. Intrathecal
9. Intracisternal
10. Intracardiac
11. Intraventricular
12. Intraocular
13. Intraperitoneal

Only the first three are discussed in any detail here. Most of these routes of administration place a drug directly or indirectly into systemic circulation. There are a number of these routes, however, by which the drug exerts a local effect, in which case most of the drug does not enter systemic circulation (e.g., intrathecal, intraventricular, intraocular, intracisternal). Certain routes of administration may exert both local and systemic effects depending on the characteristics of the drug and excipients (e.g., SC).

The choice of a particular parenteral route will depend on the required time of onset of action, the required site of action, and the characteristics of the fluid to be injected, among other factors.

These unusual concerns include irritation (vascular, muscular, or SC), pyrogenicity, blood compatibility, and sterility (Avis, 1985). The background of each of these,

along with the underlying mechanisms and factors that influence the level of occurrence of such an effect, is briefly discussed later.

16.8.1.1 Irritation Tissue irritation upon injection, and the accompanying damage and pain, is a concern that must be addressed for the final formulation, which is to be either tested in humans or marketed, rather than for the active ingredient. This is because most irritation factors are either due to or influenced by aspects of formulation design (see Avis, 1985, for more information on parenteral preparations). These factors are not independent of the route (IV, IM, or SC) that will be used and, in fact (as discussed later), are part of the basis for selecting between the various routes.

The lack of irritation and tissue damage at the injection site is sometimes called *tolerance*. Some of the factors that affect tolerance are not fully under the control of an investigation and are also unrelated to the material being injected. These include body movement, temperature, and animal age. Factors that can be controlled, but that are not inherent to the active ingredient, include solubility, tonicity, and pH. And, finally, the active ingredient and vehicle can have inherent irritative effects and factors such as solubility (in the physiological milieu into which they are being injected), concentration, volume molecular size, and particle size. Gray (1978) and Ballard (1968) discuss these factors and the morphological consequences that may occur if they are not addressed.

16.8.1.2 Blood Compatibility It is important that cellular components of the blood are not disrupted and that serum- or plasma-based responses are not triggered by parental administration. Therefore, two mechanisms must be assessed regarding the blood compatibility of component materials. These include the material's effect on cellular components that cause membrane destruction and hemolysis and the activation of the clotting mechanism resulting in the formation of the thromboemboli.

Many of the inactive, ingredient-related physicochemical factors that influence irritation (e.g., tonicity, pH, and particle size) also act to determine blood compatibility. But the chemical features of a drug entity itself—its molecular size and reactivity—can also be of primary importance.

16.8.1.3 Sterility Sterility is largely a concern to be answered in the process of preparing a final clinical formulation, and it is not addressed in detail in this chapter. However, it should be clear that it is essential that no viable microorganisms are present in any material to be parenterally administered (except for vaccines).

Bolus versus Infusion Technically, for all the parenteral routes (but in practice only for the IV route), there are two options for injecting a material into the body. The bolus and

infusion methods differentiated on the single basis of rate of injection, but they actually differ on a wide range of characteristics.

The most commonly exercised option is the bolus, or “push,” injection, in which the injection device (syringe or catheter) is appropriately entered into the vein and a defined volume of material is introduced through the device. The device is then removed. In this operation, it is relatively easy to restrain an experimental animal, and the stress on the animal is limited. Though the person doing the injection must be skilled, it takes only a short amount of time to become so. And the one variable to be controlled in determining dosage is the total volume of material injected (assuming dosing solutions have been properly prepared).

There are limitations and disadvantages to the bolus approach, however. Only a limited volume may be injected, which may prohibit the use of bolus when volumes to be introduced are high (due to, e.g., low active compound solubility or a host of other reasons). Only two devices (syringe and catheter) are available for use in the bolus approach. If a multiple-day course of treatment is desired (say, every day for 15 days), separate injections must be made at discreet entry sites.

The infusion approach involves establishing a fixed entry point into the vein, and then slowly passing the desired test material through that point over a period of time (30 min is about minimum, while continuous around-the-clock treatment is at least therapeutically possible). There are a number of devices available for establishing their entry point; catheter, vascular port (Garrazone, 1986), or osmotic pump (Theeuwes and Yum, 1976). Each of these must, in turn, be coupled with a device to deliver the dosing solution at a desired rate. The osmotic pump, which is implanted, is also its own delivery device. Other options are gravity-driven “drips,” handheld syringes (not practical or accurate over any substantial period of time), or syringe pumps. Very large volumes can be introduced by fusion over a protracted period of time, and only a single site need be fitted with an entry device.

However, infusions also have their limitations. Skilled labor is required, and the setup must be monitored over the entire period of infusion. Larger animals must be restrained, while there are special devices that make this requirement unnecessary for smaller animals. Restraint and protracted manipulation are very stressful on animals. Over a period of time, one must regularly demonstrate patency of a device—that is, that entry into the vascular system continues to exist. Finally, one is faced with having to control two variables in controlling the dose—both total volume and rate.

When are the two approaches (bolus and infusion) interchangeable? And why select one over the other? The selection of infusion is usually limited to two reasons: (i) when a larger volume must be introduced than is practical in a bolus injection or (ii) tolerance is insufficient if the dose is given all at once (i.e., an infusions will “clear” a higher daily dose

than will a bolus injection). For safety studies, when a bolus can be used to clear a human infusion, dosing is a matter of judgment. If the planned clinical infusion will take less than a half an hour, practicality dictates that animal studies be accomplished by bolus. In other situations, pharmacokinetics (in particular, the half-life of the drug entity) should be considered in making the decision.

16.8.2 Test Systems for Parenteral Irritation

There are no regulatory guidelines or suggested test methods for evaluating agents for muscular or vascular irritation. Since such guidelines are lacking but the evaluation is necessary, those responsible for these evaluations have tried to develop and employ the most scientifically valid procedures.

Hagan (1959) first suggested a method for assessing IM irritation. His approach, however, did not include a grading system for evaluation of the irritation, and the method used the sacrospinalis muscles, which are somewhat difficult to dissect or repeatedly inject.

Shintani et al. (1967) developed and proposed the methodology that currently seems to be more utilized. It uses the lateral vastus muscle and includes a methodology for evaluation, scoring, and grading of irritation. Additionally, Shintani et al. investigated the effects of several factors such as the pH of the solution, the drug concentration, the volume of injection, the effect of repeated injections, and the time to maximum tissue response.

16.8.2.1 Acute Intramuscular Irritation in the Male Rabbit (USP, 1985)

1. Overview of study design

Each rabbit is injected as follows:

| Site (M. Vastus Lateralis) | Treatment (1.0 mL site ⁻¹) |
|----------------------------|--|
| Left | (Test article) |
| Right | (Vehicle) |

Day 1: Injection of all treatment groups—nine rabbits

Day 2: Sacrifice and evaluation: 24 h posttreatment group—three rabbits

Day 3: Sacrifice and evaluation: 48 h posttreatment group—three rabbits

Day 4: Sacrifice and evaluation: 72 h posttreatment group—three rabbits

2. Administration

2.1. Route: The test article is injected into the vastus lateralis of each rabbit.

2.2. Dose: The dose selected is chosen to evaluate the severity of irritation and represents a concentration

that might be used clinically. This volume has been widely used in irritation testing.

2.3. Frequency: Once only.

2.4. Duration: 1 day.

2.5. Volume: 1.0 mL site⁻¹.

3. Test system

3.1. Species, age, and weight range: Male New Zealand white rabbits weighing 2–5 kg are used. The New Zealand white rabbit has been widely used in muscle irritation research for many years and is a reasonable sized, even-tempered animals that are well adapted to the laboratory environment.

3.2. Selection: Animals to be used in the study are selected on the basis of acceptable findings from physical examinations and body weights.

3.3. Randomization: Animals are ranked by body weight and assigned a number between one and three. The order of number assigned (e.g., 1–3–2) is chosen from a table of random numbers. Animals assigned number 1 are in the 24 h post-treatment group; those assigned number 2 are in the 48 h posttreatment group; and those assigned number 3 are in the 72 h posttreatment group.

4. In-life observations

4.1. Daily observations: Once daily following dosing.

4.2. Physical examinations: Once within the 2 weeks before the first dosing day.

4.3. Body weight: Should be determined once before the start of the study.

4.4. Additional examinations may be done by the study director to elucidate any observed clinical signs.

5. Postmortem procedures

5.1. Irritation is evaluated as follows: Three rabbits are sacrificed by a lethal dose of barbiturate at approximately 24, 48, or 72 h after dosing. The left and right lateral vastus of each rabbit are excised. The lesions resulting from the injection are scored for muscle irritation on a numerical scale of 0–5 as follows (Shintani et al., 1967):

| Reaction Criteria | Score |
|--|-------|
| No discernable gross reaction | 0 |
| Slight hyperemia | 1 |
| Moderate hyperemia and discoloration | 2 |
| Distinct discoloration in comparison with the color of the surrounding area | 3 |
| Brown degeneration with small necrosis | 4 |
| Widespread necrosis with an appearance of “cooked meat” and occasionally an abscess involving the major portions of the muscle | 5 |

The average score for the nine rabbits is then calculated, and a category of irritancy is then assigned based on the following table:

| Average Score | Grade |
|----------------|----------|
| 0.0–0.4 | None |
| 0.5–1.4 | Slight |
| 1.5–2.4 | Mild |
| 2.5–3.4 | Moderate |
| 3.5–4.4 | Marked |
| 4.5 or greater | Severe |

16.8.2.2 Acute Intravenous Irritation in the Male Rabbit

The design here is similar to the IM assay, except that injections are made into the veins in specific muscle masses.

1. Overview of study design

Rabbits will be injected as follows:

| Group | No. of Animals | Treatment Site | Evaluation (h) |
|-------|----------------|--|----------------|
| 1 | 2 | M. vastus lateralis (left) and cervicodorsal subcutis (left) | 24 |
| | | M. vastus lateralis (right) and cervicodorsal subcutis (right) | 24 |
| 2 | 2 | M. vastus lateralis (left) and cervicodorsal subcutis (left) | 72 |
| | | M. vastus lateralis (right) and cervicodorsal subcutis (right) | 72 |
| 3 | 2 | Auricular vein (left) Auricular vein (right) | 24 and 72 |

Day 1: Injection of all groups (six rabbits)

Day 2: Evaluation of Group 3 (two rabbits). Sacrifice and evaluation of Group 1 (two rabbits)

Day 4: Evaluation of Group 3 (two rabbits). Sacrifice and evaluation of Group 2 (two rabbits)

2. Administration

2.1. Intramuscular: M. vastus lateralis

2.2. Subcutaneous: Cervicodorsal subcutis.

2.3. Intravenous: Auricular vein.

2.4. Dose: The doses and concentration selected are chosen to evaluate the severity of irritation. The dose volumes have been widely used in irritation testing.

2.5. Frequency: Once only.

2.6. Duration: 1 day.

2.7. Volume: M. vastus lateralis and cervicodorsal subcutis: 1.0 mL site⁻¹; auricular vein: 0.5 mL site⁻¹.

3. Test system

3.1. Species, age, and weight range: Male New Zealand white rabbits, weighing 2–5 kg, are used.

- 3.2. Selection: Animals to be used in the study are selected on the basis of acceptable findings from physical examinations.
- 3.3. Randomization: Animals are ranked by body weight and assigned a number between 1 and 3. The order of numbers assigned (e.g., 1–3–2) is chosen from a table of random numbers. Animals assigned number 1 are in Group 1; those assigned number 2 are in Group 2; and those assigned number 3 are in Group 3.

4. In-life observations

- 4.1. Daily observations: Once daily following dosing.
- 4.2. Physical examinations: Once within the 2 weeks before the first dosing day.
- 4.3. Body weight: Determined once before the start of the study.
- 4.4. Additional examinations may be done by the study director to elucidate any observed clinical signs.

5. Postmortem procedures

- 5.1. IM irritation is evaluated as follows: Rabbits are sacrificed by lethal dose of barbiturate approximately 24 and 72h after dosing. The left and right lateral vastus muscles of each rabbit are excised. The reaction resulting from injection is scored for muscle irritation using the scale shown on page 232.
- 5.2. SC (intracutaneous) irritation is evaluated as follows: Rabbits are sacrificed by a lethal dose of barbiturate approximately 24 and 72h after dosing. The SC injection sites are exposed by dissection, and the reaction is scored for irritation on a scale of 0–5 as follows (Shintani et al., 1967; USP, 1995a):

| Reaction Criteria | Score |
|---|-------|
| No discernible gross reaction | 0 |
| Slight hyperemia and discoloration | 1 |
| Moderate hyperemia and discoloration | 2 |
| Distinct discoloration in comparison with the color of the surrounding area | 3 |
| Small areas of necrosis | 4 |
| Widespread necrosis, possibly involving the underlying muscle | 5 |

| Average Score per Site | Irritancy Grade |
|------------------------|-----------------|
| 0.0–0.4 | None |
| 0.5–1.4 | Slight |
| 1.5–2.4 | Mild |
| 2.5–3.4 | Moderate |
| 3.5–4.4 | Marked |
| 4.5 or greater | Severe |

- 5.3. IV irritation is evaluated as follows: Rabbits are sacrificed by a lethal dose of barbiturate following the 72h irritation evaluation. The injection

site and surrounding tissue are grossly evaluated at approximately 24 and 72h after dosing on a scale of 0–3 as follows:

| Reaction Criteria | Score |
|--|-------|
| No discernible gross reaction | 0 |
| Slight erythema at injection site | 1 |
| Moderate erythema and swelling with some discoloration of the vein and surrounding tissue | 2 |
| Severe discoloration and swelling of the vein and surrounding tissue with partial or total occlusion of the vein | 3 |

| Average Score per Site | Irritancy Grade |
|------------------------|-----------------|
| 0.0–0.4 | None |
| 0.5–1.4 | Slight |
| 1.5–2.4 | Moderate |
| 2.5 or greater | Severe |

- 5.4. Additional examinations may be done by the study director to elucidate the nature of any observed tissue change.

16.9 PROBLEMS IN TESTING (AND THEIR RESOLUTIONS)

Some materials, by either their physicochemical or toxicological natures, generate difficulties in the performance and evaluation of dermal irritation tests. The most commonly encountered of these problems are presented in the following text:

- A. *Compound volatility* One is sometimes required or requested to evaluate the potential irritancy of a liquid that has a boiling point between room temperature and the body temperature of the test animal. As a result, the liquid portion of the material will evaporate off before the end of the testing period. There is no real way around the problem; one can only make clear in the report on the test that the traditional test requirements were not met, though an evaluation of potential irritant hazard was probably achieved (for the liquid phase would also have evaporated from a human that it was spilled on).
- B. *Pigmented material* Some materials are strongly colored or discolor the skin at the application site. This makes the traditional scoring process difficult or impossible. One can try to remove the pigmentation with a solvent; if successful, the erythema can then be evaluated. If use of a solvent fails or is unacceptable, one can (wearing thin latex gloves) feel the skin to determine if there is warmth, swelling, and/or rigidity—all secondary indicators of the irritation response.
- C. *Systemic toxicity* On rare occasions, the dermal irritation study begun only to have the animals die very rapidly after test material is applied.

16.9.1 Alternatives to *In Vivo* Parenteral Tests

IM and IV injections of parenteral formulations of pharmaceuticals can produce a range of discomfort including pain, irritation, and/or damage to muscular or vascular tissue. These are normally evaluated for prospective formulations before use in humans by histopathologic evaluation of damage in intact animal models, usually the rabbit. Attempts have been made to make this *in vivo* methodology both more objective and quantitative based on measuring the creatinine phosphokinase released in the tissue surrounding the injection site (Sidell et al., 1974). Currently, a protocol utilizing a cultured skeletal muscle cell line (L6) from the rat as a model has been evaluated in an interlaboratory validation program among 11 pharmaceutical laboratories. This methodology (Young et al., 1986) measures creatine kinase levels in media after exposure of the cells to the formulation of interest and predicts *in vivo* IM damage based on this end point. It is reported to give excellent rank-correlated results across a range of antibiotics (Williams et al., 1987). The current multilaboratory evaluation covers a broader structural range of compounds and has shown a good quantitative correlation with *in vivo* results for antibiotics and a fair correlation for a broader range of parenteral drug products. Likewise, Kato et al. (1992) have proposed a model that uses cultured primary skeletal muscle fibers from the rat. Damage is evaluated by the release of creatinine phosphokinase. An evaluation using six parenterally administered antibiotics (ranking their EC_{50} values) showed good relative correlation with *in vivo* results.

Another proposed *in vitro* assay for muscle irritancy for injectable formulations is the red blood cell hemolysis assay (Brown et al., 1989). Water-soluble formulations in a 1:2 ratio with freshly collected human blood are gently mixed for 5 min. The percentage red blood cell survival is then determined by measuring differential absorbance at 540 nm; this value is then compared to values for known irritants and nonirritants. Against a very small group of compounds (four), this assay reportedly accurately predicts muscle irritation.

16.10 PHOTOTOXICITY

The potential for sunlight (or selected other light frequencies) to transform a drug or product is both a useful tool for activating some drugs and a cause of significant adverse effects for others (such as the quinolone antibiotics (Horio et al., 1995; Lambert et al., 1996)).

16.10.1 Theory and Mechanisms

The portion of the solar spectrum containing the biologically most active region is from 290 to 700 nm.

The UV part of the spectrum includes wavelengths from 200 to 400 nm. Portions of the UV spectrum have distinctive

features from both the physical and biological points of view. The accepted designations for the biologically important parts of the UV spectrum are UVA, 400–315 nm; UVB, 315–280 nm; and UVC, 280–220 nm. Wavelengths less than 290 nm (UVC) do not occur at the earth's surface because they are absorbed predominantly by ozone in the stratosphere. The most thoroughly studied photobiological reactions that occur in skin are induced by UVB. The quinolones, for example, absorb light strongly in the 300–400 nm wavelength range. Although UVB wavelengths represent only approximately 1.5% of the solar energy received at the earth's surface, they elicit most of the known chemical phototoxic and photoallergic reactions. The visible portions of the spectrum, representing about 50% of the sun's energy received at sea level, include wavelengths from 400 to 700 nm. Visible light is necessary for such biological events as photosynthesis, the regulation circadian cycles, vision, and pigment darkening. Furthermore, visible light in conjunction with certain chromophores (e.g., dyes, drugs, and endogenous compounds which absorb light and therefore "give" color) and molecular oxygen induces photodynamic effects.

Understanding the toxic effects of light impinging on the skin requires knowledge of both the nature of sunlight and the skin's optical properties. Skin may be viewed as an optically heterogeneous medium, composed of three layers that have distinct refractive indices, chromophore distributions, and light-scattering properties. Light entering the outermost layer, the stratum corneum, is in part reflected—4–7% for wavelengths between 250 and 3000 nm (Anderson and Parrish, 1981)—due to the difference in refractive index between air and the stratum corneum. Absorption by urocanic acid (a deamination product of histidine), melanin, and proteins containing the aromatic amino acids tryptophan and tyrosine in the stratum corneum produces further attenuation of light, particularly at shorter UV wavelengths. Approximately 40% of UVB is transmitted through the stratum corneum to the viable epidermis. The light entering the epidermis is attenuated by scattering and, predominantly, absorption. Epidermal chromophores consist of proteins, urocanic acid, nucleic acids, and melanin. Passage through the epidermis results in appreciable attenuation of UVA and UVB radiation. The transmission properties of the dermis are largely due to scattering, with significant absorption of visible light by melanin, β -carotene, and the blood-borne pigments bilirubin, hemoglobin, and oxyhemoglobin. Lightly traversing these layers of the skin is extensively attenuated, most drastically for wavelengths less than 400 nm. Longer wavelengths are more penetrating. It has been noted that there is an "optical window"—that is, greater transmission—for light at wavelengths of 600–1300 nm, which may have important biological consequences.

Normal variations in the skin's optical properties frequently occur. The degree of pigmentation may produce variations in the attenuation of light, particularly between 300 and 400 nm, by as much as 1.5 times more in blacks than in Caucasians (Pathak, 1967). Alterations in the amount or distribution of other natural chromophores account for further variations in skin optical properties. Urocanic acid, deposited on the skin's surface during perspiration (Anderson and Parrish, 1981), and UV-absorbing lipids, excreted in sebum, may significantly reduce UV transmission through the skin. Epidermal thickness, which varies over regions of the body and increases after exposure to UVB radiation, may significantly modify UV transmission.

Certain disease states also produce alterations in the skin's optical properties. Alterations of the skin's surface, such as by psoriatic plaques, decrease transmitted light. The effect may be lessened by application of oils whose refractive index is similar to that of skin (Anderson and Parrish, 1981). Disorders such as hyperbilirubinemia, porphyrias, and blue skin nevi result in increased absorption of visible light due to accumulation or altered distribution of endogenous chromophoric compounds.

The penetration of light into and through dermal tissues has important consequences. This penetration is demonstrated in Figure 16.1. Skin, as the primary organ responsible for thermal regulation, is overperfused relative to its metabolic requirements (Anderson and Parrish, 1981).

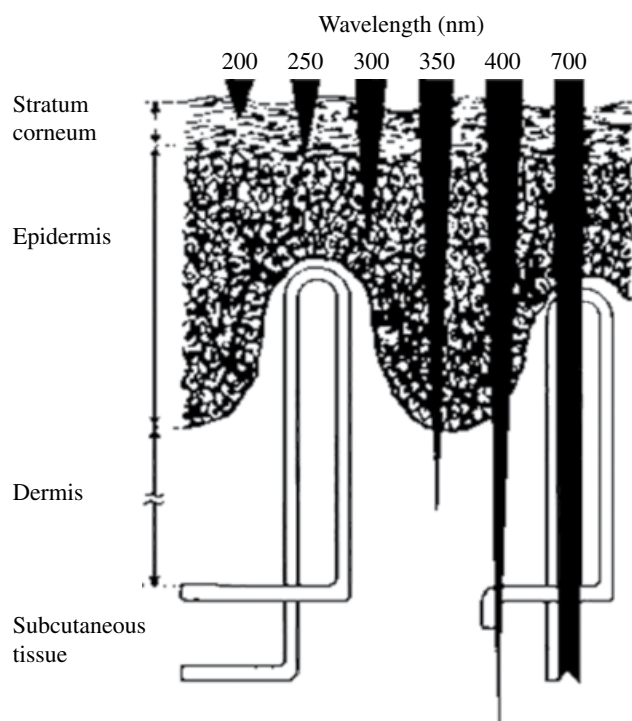


FIGURE 16.1 Schematic portraying penetration of light of varying wavelengths into skin.

It is estimated that the average cutaneous blood flow is 20–30 times than necessary to support the skin's metabolic needs. The papillary boundaries between epidermis and dermis allow capillary vessels to lie close to the skin's surface, permitting the blood and important components of the immune system to be exposed to light. The equivalent of the entire blood volume of an adult may pass through the skin, and potentially be irradiated, in 20 min. This corresponds to the time required to receive 1 or 2 MEDs (the MED is defined as the minimal dose of UV irradiation that produces definite, but minimally perceptible, redness 24 h after exposure). The accessibility of incidence radiation to blood has been exploited in such regimens and phototherapy of hyperbilirubinemia in neonates, where light is used as a therapeutic agent. However, in general, there is a potential for light-induced toxicity due to irradiation of blood-borne drugs and metabolites.

16.10.2 Factors Influencing Phototoxicity/Photosensitization

There are a number of factors which can influence an agent acting either as a phototoxin or a photoallergen. In addition to all those factors previously reviewed in Chapter 5, there are also the following:

1. The quantity and location of photoactive material present in or on the skin
2. The capacity of the photoactive material to penetrate into normal skin by percutaneous absorption as well as into the skin altered by trauma, such as maceration, irritation, and sunburn
3. The pH, enzyme presence, and solubility conditions at the site of exposure
4. The quantity of activating radiation to which the skin is exposed
5. The capacity of the spectral range to activate the materials on or within the skin
6. The ambient temperature and humidity
7. The thickness of the horny layer
8. The degree of melanin pigmentation of the skin
9. The inherent "photoactivity" of the chemical; does it weakly or strongly absorb light?

Basically, any material that has both the potential to absorb UV light (in the UVA or UVB regions) and the possibility of dermal exposure or distribution into the dermal region should be subject to some degree of suspicion as to potential phototoxicity. As shown in Table 16.5, a large number of agents have been identified as phototoxic or photoallergenic agents. Of these, tetrachlorosalicylanilide (TCSA) is the most commonly used as a positive control in animal studies.

TABLE 16.5 Known Phototoxic Agents

| In Humans | | In Animals | |
|--|---------------|---------------------------------|--------------|
| Compounds | Route | Compounds | Route |
| Aminobenzoic acid derivatives | Topical | Acridine | Topical |
| Amyldimethylamino benzoate, mixed <i>ortho</i> and <i>para</i> isomers | Topical | Amiodarone | Oral |
| Anthracene acridine | Topical | Anthracene | Topical |
| Bergapten (5-methoxypsoralen) | Topical | Bergapten (5-methoxypsoralen) | Topical |
| Cadmium sulfide | Tattoo | Bithionol | Topical |
| Chlorothiazides | Oral | Chlordiazepoxide | ip |
| Coal tar (multicomponent) | Topical | Chlorothiazide | ip |
| Dacarbazine | Infusion | Chlorpromazine | Topical |
| Disperse blue 35 (anthraquinone-based dye) | Topical | Demeclocycline | ip |
| Nalidixic acid | Oral | Griseofulvin | ip |
| Padimate A or Escalol 506 (amyl- <i>p</i> -dimethylamino benzoate) | Topical | Kynurenic acid | Oral |
| | | Nalidixic acid | Oral |
| Psoralens | Oral, topical | Prochlorperazine | ip |
| Quinolone (antibacterial) | Oral | Quinoline methanol | ip |
| | | Quinolone (antibacterial) | Oral |
| Tetracyclines | Oral | Tetracyclines | ip, topical |
| Xanthotoxin (8-methoxypsoralen) | Topical, oral | Xanthotoxin (8-methoxypsoralen) | Oral, ip, im |

16.10.3 Predictive Tests for Phototoxicity

Before we start on our description of the different methods, we will first cover some basic on light dosimetry. The intensity of the irradiation used in phototoxicity testing is determined with a light meter, which provides output as watts per square meter. The shelves on which the animals rest during the exposure periods are normally adjustable in order to control the dose of light to the exposure area. The irradiation from fluorescent lights will vary somewhat from day to day, depending on temperature, variations in line current, and so on. The dose the animals receive is generally represented as joules per square centimeter. A joule is equal to 1 W s^{-1} . Therefore, the dose of light is dependent on the time of exposure. For example, in their review, Lambert et al. (1996) discuss dosages of UVA light of 9 or 10 J cm^{-2} in the UVA spectral region. If the irradiation from the light is found to be 20 W m^{-2} at the exposure site, then the time of exposure required to obtain the target dose of light (in joules) is calculated as

$$\left(\text{Time of exposure} \right) = \frac{\text{W s}}{\text{J}} \frac{9 \text{ J}}{\text{cm}^2} \frac{\text{m}^2}{20 \text{ W}} \frac{10^4 \text{ cm}^2}{\text{m}^2} \frac{\text{min}}{60 \text{ s}} = 75 \text{ min}$$

If, with the same set of lights, 2 weeks later the irradiation is determined to be 19 W m^{-2} , then the exposure period would have to be 79 min.

The ICH promulgated a new guidance for phototoxicity testing in 2015 (ICH, 2015). It provides separate recommendations for drugs given systematically and drugs administered dermally/topically. In both cases it limits concerns to drugs having absorption between 290 and 700 nm with the 3T3 *in vitro* assay the primary screen for phototoxicity if there is absorption in this case.

For systemically administered drugs, a negative 3T3 precludes the need for further testing. If the 3T3 is positive, an *in vivo* phototoxicity study is called for.

For drugs administered by dermal routes, a positive 3TR evokes the need for *in vivo* phototoxicity and photosensitization tests.

16.10.4 3T3 In Vitro Test

The first is an *in vitro* alternative success story—the 3T3 NRU phototoxicity test, which uses BALB/c 3T3 (murine) cell line with cytotoxicity determination based on neutral red uptake to measure cell viability (OECD, 2004). While not a direct replacement alternative (as there is no *in vivo* equivalent test), it is an accepted screen for phototoxicity potential by FDA.

16.10.5 Rabbit Phototoxicity Test

The next test uses the rabbit. The traditional methodology for a predictive test for phototoxicity has been an intact

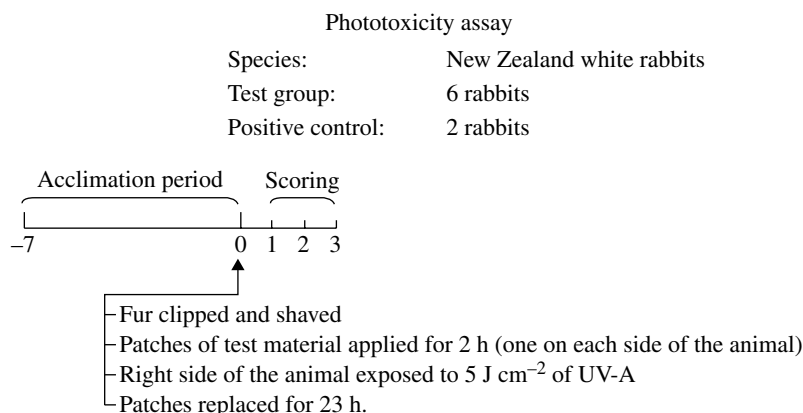


FIGURE 16.2 Line chart for design and conduct of phototoxicity assay using rabbits.

rabbit test (Marzulli and Maibach, 1970). This test is conducted as follows (and illustrated diagrammatically in Figure 16.2):

A. Animals and animal husbandry

1. Strain/species: Female New Zealand white rabbits.
2. Number: Six rabbits per test; two rabbits for positive control.
3. Age: Young adult.
4. Acclimation period: At least 7 days prior to study.
5. Food and water: Commercial laboratory feed and water are freely available.

B. Test article

1. A dose of 0.5 mL of liquid or 500 mg of a solid or semisolid will be applied to each test site.
2. Liquid substances will be used undiluted.
3. For solids, the test article will be moistened with water (500 mg test article/0.5 mL water or another suitable vehicle) to ensure good contact with the skin.
4. Positive control material will be a lotion containing 1% 8-methoxypsoralen.

C. Experimental procedures

1. Animals will be weighed on the first day of dosing.
2. On the day prior to dosing, the fur of the test animals will be clipped from the dorsal area of the trunk using a small animal clipper, and then shaved clean with a finer bladed clipper.
3. On the day of dosing, the animals will be placed in restraints.
4. One pair of patches (~2.5 × 2.5 cm) per test article will be applied to the skin of the back, with one patch on each side of the backbone.
5. A maximum of two pairs of patches may be applied to each animal, and the patches must be at least 2 in. apart.
6. The patches will be held in contact with the skin by means of an occlusive dressing for the 2-h exposure period.

7. After the 2-h exposure period, the occlusive dressing, as well as the patches on the right side of the animal, will be removed (aluminum foil).

8. The left side of the animal will be covered with opaque material.

9. The animal will then be exposed to approximately 5 J cm⁻² of UVA (320–400 nm).

10. After exposure to the UVA light, the patches on the right side of the animal, as well as the occlusive dressing, will be replaced.

11. The dressing will again be removed approximately 23 h after the initial application of the test article. Residual test article will be carefully removed, where applicable, using water (or another suitable vehicle).

12. Animals will be examined for signs of erythema and edema and the responses scored at 24, 48, and 72 h after the initial test article application according to the Draize reaction grading system previously presented in this volume.

13. Any unusual observation and mortality will be recorded.

D. Analysis of data

The data from the irradiated and nonirradiated sites are evaluated separately. The scores from erythema and eschar formation, and edema at 24, 48, and 72 h, are added for each animal (six values). The six values are then divided by three, yielding six individual scores. The mean of the six individual animal irritation scores represents the mean primary irritation score (maximum score=8, as in the PDI study). This method was developed after a human model had been developed.

16.10.6 Guinea Pig

Recently, a standardized protocol for using the guinea pig for phototoxicity testing has been proposed (Nilsson et al.,

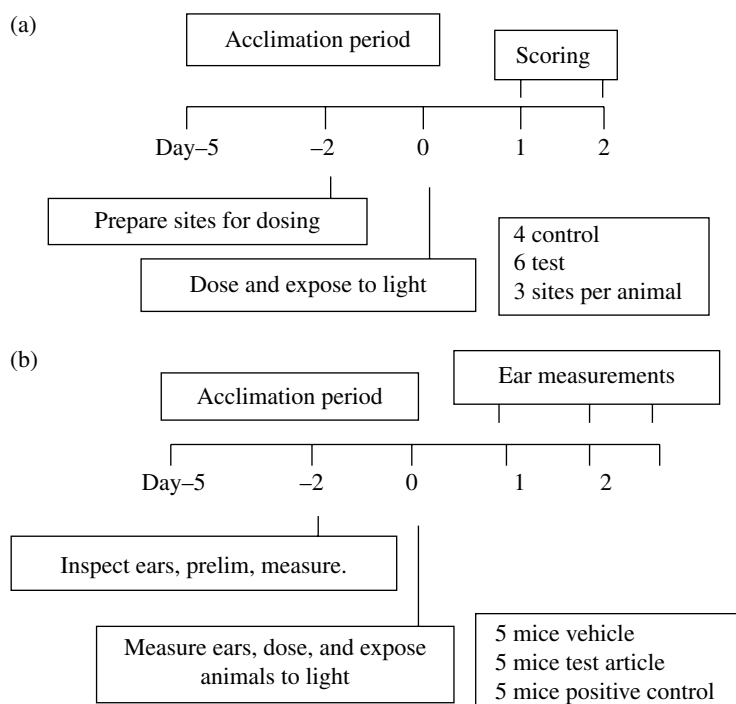


FIGURE 16.3 (a) Guinea pig and (b) mouse for phototoxicity testing.

1993), which has been the subject of an international validation exercise. This is detailed in Figure 16.3.

A. Animals and animal husbandry.

1. Strain/species: Male Hartley guinea pig.
2. Number: At least 10 (two groups).
Irradiation control: Four animals.
Test material treated: Six animals.
3. Age: Young adult, 300–500 g.
4. Acclimation period: At least 5 days.
5. Feed/water: *Ad libitum*.

B. Test material

1. Vehicle: Test assumes that material will be in solution. Use the most volatile, nonirritating organic solvent possible, for example, ethanol, acetone, dimethylacetamide, or some combination.
2. Treatment: There can be up to four sites per animal, each measuring 1.5×1.5 cm (2.25 cm²). In general, one side should be for a vehicle control and another for a positive control (8-methoxypsoralen (8-MOP), 0.005% in ethanol).
3. Dosage: A dose of 0.025–0.050 mL is applied using a micropipette to each site.

C. Experimental procedure

1. Animals will be weighed on the first day of dosing.
2. Preparation: Approximately 48 h prior to treatment, removed the hair from a 6×8 cm area on the back

with a fine clipper. On the day of dosing, animals are dosed as described previously. Tests are situated as to prevent mixing of test solutions after application. No patches or wraps are used.

3. Immediately after the dose application, the animals are placed in a restraint while keeping the test sites uncovered. Prior to irradiation the heads are covered to prevent ocular damage from the light exposure.
4. Thirty minutes after dosing, animals are exposed to a nonerythmogenic dose of light in the UVA band (should have peak intensity between 335 and 365 nm). The dose of light should be 9 or 10 J cm⁻² for UVA and 0.1–0.3 J cm⁻² for UVB.
5. Immediately after light exposure, the animals are wiped clean if necessary and returned to their home cages.
6. Animals are inspected and scored at 24- and 48-h postexposure according to the following:

- 0, no reaction
- 1, slight erythema
- 2, moderate erythema
- 3, severe erythema, with or without edema

The reader should note that this scoring scheme is the same one used for dermal sensitization scoring, whereas the scoring method for the rabbit model discussed previously is that used for dermal irritation studies.

7. Any unusual clinical signs noted during exposure should be noted. The following descriptive parameters can be calculated from the data:

$$\text{Phototoxic irritation index (PTII)} = \frac{\text{Number of positive sites} \times 100}{\text{Number of exposure sites}}$$

$$\text{Phototoxicity severity index (PSI)} = \frac{\text{Total of scores}}{\text{Total of observations}}$$

Lovell and Sanders (1992) had previously proposed a similar model of assessing topical phototoxicity potential in the guinea pig. Their model differed from that proposed by Nilsson et al. (1993) with regard to the following:

- Only one test site per animal was used.
- Test sites were smaller (about 1.6 cm²).
- Amounts applied were less (about 10 µL).
- Light intensity was set at 15 J cm⁻².
- Their paper made no reference to the use of a restrainer.
- Assessments were conducted at 4, 24, 48, and 72 h.

The scoring system was as follows:

- 0, normal
- 2, faint/trace erythema
- 4, slight erythema
- 6, definite erythema
- 8, well-developed erythema
- (Intermediate scores were indicated by odd numbers.)

They recommended the use of acidine (weak phototoxin) or anthracene (strong phototoxin) for positive controls.

16.10.7 Pyrogenicity

Pyrogenicity is the induction of a febrile (fever) response induced by the parenteral (usually IV or IM) administration of exogenous material. Pyrogenicity is usually associated with microbiological contamination of a final formulation, but it is now of increasing concern because of the growing interest in biosynthetically produced materials. Generally, ensuring sterility of product and process will guard against pyrogenicity for traditional pharmaceuticals. For biologically produced products, the FDA has promulgated the general guideline that no more than 5.0 units of endotoxin may be present per milligram of drug substance.

The US Pharmacopeia describes a pyrogen test using rabbits as a model (USP, 1995b). This test, which is the standard for limiting risks of a febrile reaction to an acceptable level, involves measuring the rise in body temperature

in a group of three rabbits for 3 h after injection of 10 mL of test solution.

16.10.7.1 Apparatus and Diluents Render the syringes, needles, and glassware free of pyrogens by heating at 250°F for not less than 30 min or by any other suitable method. Treat all diluents and solutions for washing and rinsing of devices or parenteral injection assemblies in a manner that will ensure that they are sterile and pyrogen free. Periodically perform control pyrogen tests on representative portions of the diluents and solutions that are used for washing or rinsing of the apparatus.

16.10.7.2 Temperature Recording Use an accurate temperature-sensing device, such as a clinical thermometer or thermistor or similar probe, that has been calibrated to ensure an accuracy of ±0.1° and has been tested to determine that a maximum reading is reached in less than 5 min. Insert the temperature-sensing probe into the rectum of the test rabbit to a depth of not less than 7.5 cm and, after a period of time not less than that previously determined as sufficient, record the rabbit's temperature.

16.10.7.3 Test Animals Use healthy, mature rabbits. House the rabbits individually in an area of uniform temperature (between 20 and 23°C) free from disturbances likely to excite them. The temperature should vary no more than ±3°C from the selected temperature. Before using a rabbit for the first time in a pyrogen test, condition it for not more than 7 days before use by a sham test that includes all of the steps as directed under procedure, except injection. Do not use a rabbit for pyrogen testing more frequently than once every 48 h, nor prior to 2 weeks following a maximum rise in its temperature of 0.6° or more while being subjected to the pyrogen test, or following its having been given a test specimen that was adjusted to be pyrogenic.

16.10.7.4 Procedure Perform the test in a separate area designated solely for pyrogen testing and under environmental conditions similar to those under which the animals are housed. Withhold all food from the test rabbits during the period of the test. Access to water is allowed at all times but may be restricted during the test. If probes measuring rectal temperature remain inserted throughout the testing period, restrain the rabbits with loose-fitting Elizabethan collars that allow the rabbits to assume a natural resting posture. Not more than 30 min prior to the injection of the test dose, determine the "control temperature" of each rabbit; this is the base for the determination of any temperature increase resulting from the injection of a test solution. In any one group of test rabbits, use only those rabbits whose control temperatures do not vary by more than 1°C from each other, and do not use any rabbit having a temperature exceeding 39.8°C.

Unless otherwise specified in the individual protocol, inject 10 mL of the test solution per kilogram of body weight into an ear vein in each of the three rabbits, completing each injection within 10 min after the start of administration. The test solution is either the product, constituted if necessary as directed in the labeling, or the material under test. For pyrogen testing of devices or injection assemblies, use washing or rinsing of the surfaces that come in contact with the parenterally administered material or with the injection site or internal tissues of the patient. Ensure that all test solutions are protected from contamination. Perform the injection after warming the test solution to a temperature of $37 \pm 2^\circ\text{C}$. Record the temperature at 1, 2, and 3 h subsequent to the injection.

16.10.7.5 Test Interpretation and Continuation Consider any temperature decreases as zero rises. If no rabbit shows an individual rise in temperature of 0.6° or more above its respective control temperature, and if the sum of the three individual maximum temperature rises does not exceed 1.4° , the product meets the requirements for the absence of pyrogens. If any rabbit shows an individual temperature rise of 0.6° or more, or if the sum of the three individual maximum temperature rises exceeds 1.4° , continue the test using five other rabbits. If not more than three of the eight rabbits show individual rises in temperature of 0.6° or more, and if the sum of the eight individual maximum temperature rises does not exceed 3.7° , the material under examination meets the requirements for the absence of pyrogens.

In Vitro Pyrogenicity *In vitro* pyrogenicity testing (or bacterial endotoxin testing) is one of the great success stories for *in vitro* testing. Some 15 years ago, the limulus amebocyte lysate (LAL) test was developed, validated, and accepted as an *in vitro* alternative (Cooper, 1975; Weary and Baker, 1977) to the rabbit test. An *in vitro* test for estimating the concentration of bacterial endotoxins that may be present in or on a sample of the article(s) to which the test is applied uses LAL that has been obtained from aqueous extracts of the circulating amebocytes of the horseshoe crab, *Limulus polyphemus*, and that has been prepared and characterized for use as an LAL reagent for gel-clot formation. The test's limitation is that it detects only the pyrogens of gram-negative bacteria. This is generally not significant, since most environmental contaminants that gain entrance to sterile products are gram negative (Devleeshouwer et al., 1985).

Where the test is conducted as a limit test, the specimen is determined to be positive or negative to the test judged against the endotoxin concentration specified in the individual monograph. Where the test is conducted as an assay of the concentration of endotoxin, with calculation of confidence limits of the result obtained, the specimen is judged to comply with the requirements if the result does not exceed (i) the concentration limit specified in the

individual monograph and (ii) the specified confidence limits for the assay. In either case the determination of the reaction end point is made with parallel dilutions of redefined endotoxin units.

Since LAL reagents have also been formulated to be used for turbidimetric (including kinetic) assays or colorimetric readings, such tests may be used if shown to comply with the requirements for alternative methods. These tests require the establishment of a standard regression curve, and the endotoxin content of the test material is determined by interpolation from the curve. The procedure includes incubation for a preselected time of reacting endotoxin and control solutions with LAL reagent and reading the spectrophotometric light absorbance at suitable wavelengths. In the case of the turbidimetric procedure, the reading is made immediately at the end of the incubation period. In the kinetic assays, the absorbance is measured throughout the reaction period, and rate values are determined from those readings. In the colorimetric procedure, the reaction is arrested at the end of the preselected time by the addition of an appropriate amount of acetic acid solution prior to the readings. A possible advantage in the mathematical treatment of results, if the test is otherwise validated and the assay suitably designed, could be the confidence interval and limits of potency from the internal evidence of each assay itself.

16.11 HEMOCOMPATIBILITY

The standard test (and its major modifications) currently used for this purpose is technically an *in vitro* one, but it requires a sample of fresh blood from a dog or other large donor animal. The test was originally developed by the National Cancer Institute for use in evaluating cancer chemotherapeutic agents (Prieur et al., 1973) and is rather crude, though definitive.

The variation described here is one commonly utilized. It uses human blood from volunteers, eliminating the need to keep a donor colony of dogs. The test procedure is described in the following text:

1. **Test system** Human blood. Collect 30 mL heparinized blood for whole blood and plasma (three tubes) and 30 mL clotted blood for serum (two tubes) from each of six donors.
2. **Precipitation potential**
 - 2.1. For each donor, set up and label eight tubes: 1 through 8.
 - 2.2. Add 1 mL serum to tubes 1 through 4.
 - 2.3. Add 1 mL plasma to tubes 5 through 8.
 - 2.4. Add 1 mL formulation to tubes 1 through 5.
 - 2.5. Add 1 mL vehicle to tubes 2 and 6.

- 2.6. Add 1 mL physiological saline to tubes 3 and 7 (negative control).
- 2.7. Add 1 mL 2% nitric acid to tubes 4 and 8 (positive control).
- 2.8. Observe tubes 1 through 8 for qualitative reactions (e.g., precipitation or clotting) before and after mixing.
- 2.9. If a reaction is observed in the formulation tubes (tubes 1 and/or 5), dilute the formulation in an equal amount of physiological saline ($\frac{1}{2}$ dilution) and test 1 mL of the dilution with an equal amount of plasma and/or serum. If a reaction still occurs, make serial dilutions of the formulation in saline (i.e., $\frac{1}{4}$, $\frac{1}{8}$, etc.).
- 2.10. If a reaction occurs in the vehicle tubes (tubes 2 and/or 6), repeat in a manner similar to that in step 2.9.
3. *Hemolytic potential*
 - 3.1. For each donor, set up and label eight tubes: 1 through 8.
 - 3.2. Add 1 mL whole blood to each tube.
 - 3.3. Add 1 mL formulation to tube 1.
 - 3.4. Add 1 mL vehicle to tube 2.
 - 3.5. Add 1 mL of $\frac{1}{2}$ dilution of formulation in saline to tube 3.
 - 3.6. Add 1 mL of $\frac{1}{2}$ dilution of vehicle in saline to tube 4.
 - 3.7. Add 1 mL of $\frac{1}{4}$ dilution of formulation in saline to tube 6.
 - 3.8. Add 1 mL of $\frac{1}{4}$ dilution of vehicle in saline to tube 6.
 - 3.9. Add 1 mL of physiological saline to tube 7 (negative control).
 - 3.10. Add 1 mL of distilled water to tube 8 (positive control).
 - 3.11. Mix by gently inverting each tube three times.
 - 3.12. Incubate tubes for 45 min at 37°C.
 - 3.13. Centrifuge 5 min at 1000 g.
 - 3.14. Separate the supernate from the sediment.
 - 3.15. Determine hemoglobin concentrations to the nearest 0.1 g dL⁻¹ on the supernate (plasma).
 - 3.16. If hemoglobin concentrations of the above dilutions are 0.2 g dL⁻¹ (or more) greater than the saline control, repeat the procedure, adding 1 mL of further serial dilutions (1/8, 1/16, etc.) of formulation or vehicle to 1 mL of blood until the hemoglobin level is within 0.2 g dL⁻¹ of the saline control.

There are two proposed, true *in vitro* alternatives to this procedure (Mason et al., 1974; Kambic et al., 1976), but neither has been either widely evaluated or accepted.

REFERENCES

- Anderson RR, Parrish JA. (1981) The optics of skin. *J. Invest. Dermatol.* 77:13–19.
- Auletta C. (1994) Vaginal and rectal administration. *J. Am. Coll. Toxicol.* 13:48–63.
- Avis KE. (1985) Parenteral preparations. In: *Remington's Pharmaceutical Sciences* (Gennaro AR, Ed.). Mack Publishing Company, Easton, PA, pp. 1518–1541.
- Ballard BE. (1968) Biopharmaceutical considerations in subcutaneous and intramuscular drug administration. *J. Pharm. Sci.* 57:357–378.
- Brown S, Templeton D, Prater DA, Potter CC. (1989) Use of an *in vitro* hemolysis test to predict tissue irritancy in an intramuscular formation. *J. Parenter. Sci. Technol.* 43:117–120.
- Chan PW, Hayes AW. (2014) Acute toxicology and eye irritancy. In: *Principles and Methods of Toxicology* (Hayes AW and Kruger CL, Eds.), 6th Ed. CRC Press, Boca Raton, FL, pp. 1117–1171.
- Cooper JF. (1975) Principles and applications of the limulus test for pyrogen in parenteral drugs. *Bull. Parenter. Drug Assoc.* 3:122–130.
- CPSC. (1976) *Illustrated Guide for Grading Eye Irritation by Hazardous Substances*. CPSC, Washington, DC.
- Devleeshouwer MJ, Cornil MF, Dony J. (1985) Studies on the sensitivity and specificity of the limulus amebocyte lysate test and rabbit pyrogen assays. *Appl. Environ. Microbiol.* 50:1509–1511.
- Draize JH, Woodard G, Calvery HO. (1944) Method for the study of irritation and toxicity of substances applied topically to the skin and mucous membranes. *J. Reprod. Fertil.* 20:85–93.
- Eckstein P, Jackson M, Millman N, Sobero A. (1969) Comparison of vaginal tolerance test of spermical preparations in rabbits and monkeys. *J. Reprod. Fertil.* 20:85–93.
- Gad SC. (2000) *In Vitro Toxicology*, 2nd Ed. Taylor & Francis, Philadelphia, PA.
- Gad SC. (2009) Alternatives to *in vivo* studies in toxicology. In: *General and Applied Toxicology* (Ballantyne B, Marrs TC and Syversen T, Eds.), 3rd Ed. John Wiley & Sons, Ltd, Chichester, pp. 779–808.
- Gad SC, Chengelis CP. (1998) *Acute Toxicity: Principles and Methods*, 2nd Ed. Academic Press, San Diego, CA.
- Gad SC, Walsh RD, Dunn BJ. (1986) Correlation of ocular and dermal irritancy of industrial chemicals. *Ocul. Dermal Toxicol.* 5(3):195–213.
- Garramone JP. (1986) *Vascular Access Port Model SLA. Users Manual*. Norfolk Medical Products, Skokie, IL.
- Gautheron P, Dukic M, Alix D, Sina J. (2004) Bovine corneal opacity and permeability test: an *in vitro* assay of corneal irritancy. *Fundam. Appl. Toxicol.* 18:442–449.
- Gray JE. (1978) Pathological evaluation of injection injury. In: *Sustained and Controlled Release Drug Delivery Systems* (Robinson J, Ed.). Marcel Dekker, New York, pp. 351–405.
- Guillot JP, Gonnet JF, Clement C, Caillard L, Trahaut R. (1982) Evaluation of the cutaneous-irritation potential of 56 compounds. *Food Chem. Toxicol.* 20:563–572.

- Hagan EC. (1959) *Appraisal of the Safety of Chemicals in Foods, Drugs and Cosmetics*. Association of Food and Drug Officials of the United States, Austin, TX, p. 19.
- Horio T, Kohachi K, Ogawa A, Inoue T, Ishirara M. (1995) Evaluation of photosensitizing ability of quinolones in guinea pigs. *Drugs*. 9(Suppl. 2):283–285.
- ICH. (2015) *S10 Photosafety Evaluation of Pharmaceuticals*. International Conference on Harmonisation, Geneva, Switzerland.
- Kambic HE, Kiraly RJ, Yukihiro N. (1976) A simple *in vitro* screening test for blood compatibility of materials. *J. Biomed. Mater. Res. Symp.* 7:561–570.
- Kato I, Harihara A, Mizushima Y. (1992) An *in vitro* method for assessing muscle irritation of antibiotics using rat primary cultured skeletal muscle fibers. *Toxicol. Appl. Pharmacol.* 117:194–199.
- Lambert L, Warner W, Kornhauser A. (1996) Animal models for phototoxicity testing. *Toxicol. Methods*. 2:99–114.
- Lovell WW, Sanders DJ. (1992) Phototoxicity testing in guinea pigs. *Food Chem. Toxicol.* 30:155–160.
- Marzulli FM, Maibach HK. (1970) Perfume phototoxicity. *J. Soc. Cosmet. Chem.* 21:685–715.
- Mason RG, Shermer RW, Zucker W h, Elston RC, Blackwelder WC. (1974) An *in vitro* test system for estimation of blood compatibility of biomaterials. *J. Biomed. Mater. Res.* 8:341–356.
- Mathias CGT. (1983) Clinical and experimental aspects of cutaneous irritation. In: *Dermatotoxicology* (Margull FM and Maibach HT, Eds.). Hemisphere Publishing, New York, pp. 167–183.
- Nilsson R, Maurer T, Redmond N. (1993) A standard protocol for phototoxicity testing: results from an interlaboratory study. *Contact Dermatitis*. 28:285–290.
- OECD. (2004) *Test No. 432 In Vitro 3T3 NRU Phototoxicity Test, OECD Guidelines for the Testing of Chemicals, Section 4*. Organisation for Economic Co-operation and Development, Paris.
- Pathak MA. (1967) Photobiology of melanogenesis: biophysical aspects. In: *Advances in Biology of Skin, Vol. 8, The Pigmentary System* (Montagna W and Ulf F, Eds.). Pergamon, New York, pp. 400–419.
- Prieur DJ, Young DM, Davis RD, Cooney DA, Homan ER, Dixon RL, Guarino AM. (1973) Procedures for preclinical toxicologic evaluation of cancer chemotherapeutic agents: protocols of the laboratory of toxicology. *Cancer Chemother. Rep.* 3. 4:1–30.
- Shintani S, Yamazaki M, Nakamura M, Nakayama I. (1967) A new method to determine the irritation of drugs after intramuscular injection in rabbits. *Toxicol. Appl. Pharmacol.* 11:293–301.
- Sidell FR, Calver DL, Kaminskis A. (1974) Serum creatine phosphokinase activity after intramuscular injection. *JAMA*. 228:1884–1887.
- Theeuwes F, Yum SI. (1976) Principles of the design and operation of generic osmotic pumps for the delivery of semisolid or liquid drug formulations. *Ann. Biomed. Eng.* 4:343–353.
- USP. (1985) *Intramuscular Irritation Test. United States Pharmacopeia*. USP Convention, Rockville, MD, pp. 1180–1183.
- USP. (1995a) *Intracutaneous Test, United States Pharmacopoeia*. USP Convention, Rockville, MD, pp. 1201–1202.
- USP. (1995b) *Bacterial Endotoxins Test. United States Pharmacopeia*. USP Convention, Rockville, MD, pp. 1696–1697.
- Weary M, Baker B. (1977) Utilization of the limulus amoebocyte lysate test for pyrogen testing of large-volume parenterals, administration sets and medical devices. *Bull. Parenter. Drug Assoc.* 31:127–133.
- Weil CS, Scala RA. (1971) Study of intra- and interlaboratory variability in the results of rabbit eye and skin irritation tests. *Toxicol. Appl. Pharmacol.* 19:276–360.
- Wilhelm K-P, Zhai H, Maibach HI. (2012) *Dermatotoxicology*, 8th Ed. Informa, Philadelphia, PA.
- Williams PD, Masters BG, Evans LD, Laska DA, Hattendorf GH. (1987) An *in vitro* model for assessing muscle irritation due to parenteral antibiotics. *Fundam. Appl. Toxicol.* 9:10–17.
- Young MF, Tobretta LD, Sophia JV. (1986) Correlative *in vitro* and *in vivo* study of skeletal muscle irritancy. *Toxicologist*. 6(1):1225.

PHARMACOKINETICS AND TOXICOKINETICS IN DRUG SAFETY EVALUATION

17.1 INTRODUCTION

Among the cardinal principles of both toxicology and pharmacology is that the means by which an agent comes in contact with or enters the body (i.e., the route of exposure or administration) does much to determine the nature and magnitude of its effects (Goldstein et al., 1974; Pratt and Taylor, 1990). Accordingly, an understanding of routes of administration and their implications for absorption is essential. The therapeutic index (calculated based on plasma and therefore absorbed levels) is the ratio between the levels that cause adverse effects and the levels that cause desired therapeutic effects in humans. This is the clinical analog of the margin of safety in toxicology studies, which used to be based on administered doses but now is based on the plasma levels at different animal doses (Yacobi et al., 1989).

Safety assessment studies usually involve a control group of animals (untreated and/or dosed with formulation only) and at least three treated groups receiving low-, intermediate-, and high-dose levels of the drug entity of interest via a route approximately that is used in man (as closely as possible). Frequently there will also be “recovery” groups to determine if any observed effects are reversible (and if so, to what extent). In most instances, the high-dose level is expected to elicit some toxic effects in the animals, expressed as anything from decreased food consumption and/or below-normal body weight gain to mortality or actual body weight loss and severe clinical signs, and has been selected after consideration of earlier data, perhaps from dose range-finding studies, or at least to dose as high as possible by the intended route. The other

two dose levels are anticipated to not cause toxic effects or only cause moderate effects. Generally, but not always (e.g., nonsteroidal anti-inflammatory drugs in rodents), the “low-dose” level is a several-fold multiple of the expected human therapeutic or exposure level (generally more so in rodents than in nonrodents). However, without knowing the true relationship of these dose levels to each other with respect to the absorption, distribution, and elimination of the new molecular entity as reflected by its pharmacokinetics (PKs), it is difficult to see how meaningful extrapolations concerning safety margins can be made from the toxicity data obtained.

Pharmacokinetic studies can provide information on several aspects of the earlier mentioned factors, the knowledge of which greatly facilitates assessment of safety of the therapeutic entity. Six such aspects can be mentioned:

1. Relationship between the dose levels used and the relative extent of absorption of the test compound
2. Relationship between the protein binding of the test compound and the dose levels used
3. Relationship between pharmacological or toxicological effects and the kinetics of the test compound
4. Effect of repeated doses on the kinetics of the test compound
5. Relationship between the age of the animal and the kinetics of the test compound
6. Relationship between the dose regimens of the test compound used in the toxicity studies and those employed clinically in man

ICH guidelines (ICH, 2004a, b) dictate a clearly defined set of objectives for toxicokinetic (TK) studies:

- *Primary* To describe the systemic exposure achieved in animals and its relationship to dose level and the time course of toxicity studies
- *Secondary*
 1. To relate the exposure achieved in toxicity studies to toxicological findings and contribute to the assessment of the relevance of these findings to clinical safety
 2. To support the choice of species and treatment regimen in nonclinical toxicity studies
 3. To provide information which, in conjunction with the toxicity findings, contributes to the design of subsequent nonclinical toxicity studies

These data may be obtained from all animals on a toxicity study, or from representative subgroups, or from satellite groups, or from separate studies.

If toxicology can be described as being the study of the effects of a chemical on an organism, metabolism can be described as the opposite—the effects of the organism on the chemical; metabolism refers to a process by which a drug (xenobiotic) is chemically modified by an organism (Roberts and Renwick, 2014). It is part of the overall process of disposition of xenobiotic (ADME)—the process of how a chemical gains access to the inner working of an organism (absorption), how it moves around inside an organism (distribution), how it is changed by the organism (metabolism), and how it is eventually eliminated from the organism (elimination). The EPA definition of biotransformation or metabolism is “...the sum of processes by which a xenobiotic (foreign chemical) is handled by a living organism.” The mathematical formulas used to describe and quantify these processes are collectively known as PKs. The EPA definition of PKs is “... quantitation and determination of the time course and dose dependency of the absorption, distribution, biotransformation, and excretion of chemicals.” The acronym ADME (absorption, distribution, metabolism, and excretion) has been used to describe the multifaceted biological process. The term metabolism has also come into common jargon to describe the entire process. This science has long played a central role in pharmaceutical development but has played less a role in the development of other types of products. The purpose of this chapter is to introduce the basic concepts of PKs (ADME) and describe the practices in the conduct of studies, as described in the regulations under ICH and OECD which require such data for non-pharmaceutical products, and to provide some real-world examples.

17.2 REGULATIONS

FDA believes that data from studies on the absorption, distribution, metabolism, and excretion of a chemical provide insight into mechanisms of toxicity and are essential in the design and evaluation of results from other toxicity studies. Such data should be provided for all drugs and significant impurities. Recommendations for obtaining data on the metabolism and PKs of these substances are presented in the ICH guidelines and the FDA Redbook II (2000). In general, it is required that this information be obtained as part of initial and subsequent repeat dose studies with a drug. EMEA has promulgated separate guidelines (EMEA, 2006) for evaluating the PKs of protein therapeutics.

17.3 PRINCIPLES

17.3.1 Preliminary Work

Before one ventures into GLP general (systemic) toxicity studies, some initial characterization of comparative species metabolism and understanding of initial routes of metabolism are essential. The studies used for these early characterizations are performed using *in vitro* methods. These can be performed using either enzyme (microsomal) preparations or isolated cells (usually hepatocytes) from the species of interest (initially, humans and a panel of potential animal models—mouse, rat, dog, minipig, and nonhuman primate (NHP)). Valuable assessments to enable subsequent decisions are metabolic stability (how rapidly the drug is metabolized by the microsomes or hepatocytes of different species); comparative metabolic “fingerprinting” by different species (either microsomes or hepatocytes) (see Dow, 2006); identification of potential metabolites, in which cytochrome P450 (CYP) isozymes are involved in the drug’s metabolism; and identification if the drug inhibits or induces a higher rate of metabolic activity by a panel of CYPs (therefore allowing a preliminary understanding of potential drug/drug interactions (DDIs)). Remembering that it is only the “free” molecules of a drug which have therapeutic (or toxicological) potential, one can also evaluate plasma protein binding *in vitro* (Smith and Kerns, 2010).

The design and analysis of PK studies require a broad understanding of the underlying concepts and principles inherent in the ADME process and in our current (and evolving) technology for studying such. Each of these four principal areas is overviewed from a practical basis as it relates to toxicology. First, however, one should consider the fundamental terminology used in PK studies (Table 17.1).

TABLE 17.1 Fundamental Terms Used in PK Studies

| Absolute bioavailability | The bioavailability of a dosage form relative to an intravenous administration |
|-----------------------------|--|
| Absorption | The process by which a xenobiotic and its metabolites are transferred from the site of absorption to the blood circulation |
| Accumulation | The progressive increase of chemical and/or metabolites in the body. Accumulation is influenced by the dosing interval and half-life of the chemical. The process can be characterized by an “accumulation factor,” which is the ratio of the plasma concentration at steady state to that following the first dose in a multiple dosing regimen |
| Analyte | The drug entity assayed in biological samples |
| Area under the curve (AUC) | The concentration of chemical and/or metabolites in the blood (or plasma/serum) integrated over time. This is typically considered the best indicator of exposure |
| Bioavailability | The rate and extent to which a xenobiotic entity enters the systemic circulation intact, following oral or dermal administration. It is sometimes expanded to include therapeutically active metabolites. Also known as the comparative bioavailability |
| Biotransformation | The process by which a xenobiotic is structurally and/or chemically changed in the body by either enzymatic or nonenzymatic reactions. The product of the reaction is a different composition of matter or different configuration than the original compound |
| Clearance | The volume of biological fluid which is totally cleared of xenobiotic in a unit time |
| C_{\max} | The maximum mean concentration of the chemical in the plasma. Also known as the peak plasma concentration |
| Concomitant TKs | Toxicokinetic measurements performed in the toxicity study, either in all animals or in representative subgroups or in satellite groups |
| Disposition | All processes and factors which are involved from the time a chemical enters the body to the time when it is eliminated from the body, either intact or in metabolite form |
| Distribution | The process by which an absorbed xenobiotic and/or its metabolites partition between blood and various tissues/organs in the body |
| Dosage form | The formulation (diet, lotion, capsule, solution, etc.) administered to animals or man |
| Dose proportionality | The relationship between doses of a chemical and measured parameters, usually including tests for linearity |
| Enterohepatic circulation | The process by which xenobiotics are emptied via the bile into the small intestine and then reabsorbed into the hepatic circulation |
| Enzyme induction | The increase in enzyme content (activity and/or amount) due to xenobiotic challenge, which may result in more rapid metabolism of a chemical |
| Enzyme inhibition | The decrease in enzymatic activity due to the effect of xenobiotic challenge |
| Excretion | The process by which the administered compound and/or its biotransformation products are eliminated from the body |
| Exposure | Exposure is represented by PK parameters demonstrating the local and systemic burden on the test species with the test compound and/or its metabolites. The area under the matrix level concentration-versus-time curve (AUC) and/or the measurements of matrix concentrations at the expected peak concentration time C_{\max} , or at some other selected time $C_{(\text{time})}$, are the most commonly used parameters. Other parameters might be more appropriate in particular cases |
| First-order kinetics | Kinetic processes, the rate of which is proportional to the concentration |
| First-pass effect | The phenomenon whereby xenobiotics may be extracted or metabolized following enteral absorption before reaching the systemic circulation |
| Flux | Term (which takes area into consideration) used to describe the movement of a chemical across a barrier. Most typically used to describe the absorption of a chemical across the skin as micrograms per square centimeter per hour |
| Half-life | The time elapsed for a given chemical entity concentration or amount to be reduced by a factor of two |
| Hepatic clearance | The rate of total body clearance accounted for by the liver |
| Kel | The elimination constant for a chemical in plasma. Typically calculated using the formula $K_{el} = -\ln[10] \times b$ where b is the slope of the linear regression line of the log of the mean plasma concentrations vs. time from t_{\max} to 24 h |
| Lag time | The interval between compound administration and when the compound concentration is measurable in blood |
| Metabolite characterization | The determination of physiochemical characteristics of the biotransformation product(s) |
| Metabolite identification | The structural elucidation of the biotransformation product(s) |
| Metabolite profile | The chromatographic pattern and/or aqueous/nonaqueous partitioning of the biotransformation products of the administered compound |

(Continued)

TABLE 17.1 (Continued)

| Absolute bioavailability | The bioavailability of a dosage form relative to an intravenous administration |
|---|--|
| Monitor | To take a small number of matrix samples (e.g., 1–3) during a dosing interval to estimate $C_{(time)}$ and/or C_{max} |
| Nonlinear kinetics (saturation kinetics) | Kinetic processes, the rate of which is not directly proportional to the concentration |
| Presystemic elimination | The loss of that portion of the dose that is not bioavailable. This would include, among others, loss through intestinal and gut-wall metabolism, lack of absorption, and first-pass hepatic metabolism |
| Profile | To take (e.g., 4–8) matrix samples during a dosing interval to make and estimate C_{max} and/or $C_{(time)}$ and area under the matrix level concentration-versus-time curve (AUC) |
| Protein binding | The complexation of a xenobiotic and/or its metabolite(s) with plasma or tissue proteins |
| Relative bioavailability | The bioavailability relative to a reference or standard formulation or agent |
| Renal clearance | The rate of total body clearance accounted for by the kidney. Its magnitude is determined by the net effects of glomerular filtration, tubular secretion and reabsorption, renal blood flow, and protein binding |
| Satellite | Groups of animals included in the design and conduct of a toxicity study, treated and housed under conditions identical to those of the main study animals, but used primarily for TKs |
| Steady state | An equilibrium state where the rate of chemical input is equal to the rate of elimination during a given dose interval |
| Support | In the context of a toxicity study—to ratify or confirm the design of a toxicity study with respect to PK and metabolic principles. This process may include two separate steps: <ol style="list-style-type: none"> 1. Confirmation using TK principles that the animals on a study were exposed to appropriate systemic levels of the administered compound and/or its metabolite(s) 2. Confirmation that the metabolic profile in the species used was acceptable (the data to support this will normally be derived from metabolism studies in animals and in humans) |
| T_{max} | The sampling time point at which C_{max} occurs |
| Total clearance | The volume of biological fluid totally cleared of xenobiotic per unit time, usually including hepatic clearance and renal clearance |
| TKs | The study of the kinetics of absorption, distribution, metabolism, and excretion of toxic or potentially toxic chemicals |
| Validate | In the context of an analytical method—to establish the accuracy, precision, reproducibility, response function, and specificity of the analytical method with reference to the biological matrix to be examined and the analyte to be quantified |
| Volume of distribution (V_d) | A hypothetical volume of body fluid into which the chemical distributes. It is not a “real” volume, but is a proportionality constant relating the amount of chemical in the body to the measured concentration in blood or plasma |

17.3.2 Absorption

Absorption describes the process by which a chemical crosses a biological membrane to gain access to the inner workings of an organism. For mammals, this process results in the entry of the chemical into the bloodstream or systemic circulation. In this case, the process is also called systemic absorption. Pharmaceutical products, procedures, and devices, such as hypodermic needles or catheters, can be used to bypass biological barriers. Other products gain access to the systemic circulation via the oral, dermal, buccal, or inhalatory route of administration.

For a material to be toxic (local tissue effects are largely not true toxicities by this definition), the first requirement is that it be absorbed into the organism (for which purpose being in the cavity of the gastrointestinal (GI) tract does not qualify). Most pharmaceuticals are intended to gain such access.

There are characteristics which influence absorption by the different routes, and these need to be understood by any person trying to evaluate and/or predict the toxicities of

different moieties. Some key characteristics and considerations are summarized in the succeeding text by route.

A. Oral and Rectal Routes (Gastrointestinal Tract)

1. Lipid-soluble compounds (nonionized) are more readily absorbed than water-soluble compounds (ionized):
 - a. Weak organic bases are in the nonionized, lipid-soluble form in the intestine and tend to be absorbed there.
 - b. Weak organic acids are in the nonionized, lipid-soluble form in the stomach and one would suspect they would be absorbed there, but the intestine is more important because of time and area of exposure.
2. Specialized transport systems exist for some moieties: sugars, amino acids, pyrimidines, calcium, and sodium.
3. Almost everything is absorbed—at least to a small extent (if it has a molecular weight below 10000).

4. Digestive fluids may modify the structure of a drug.
 5. Dilution increases toxicity because of more rapid absorption from the intestine, unless stomach contents bind the moiety.
 6. Physical properties are important—for example, dissolution of metallic mercury is essential to allow absorption.
 7. Age—neonates have a poor intestinal barrier.
 8. Effect of fasting on absorption depends on the properties of the chemical of interest.
- B. Inhalation (Lungs)
1. Aerosol deposition:
 - a. Nasopharyngeal—5 μm or larger in man, less in common laboratory animals.
 - b. Tracheobronchial—1 to 5 μm
 - c. Alveolar—1 μm
 2. If a solid, mucociliary transport may serve to clear from lungs to GI tract.
 3. Lungs are anatomically good for absorption:
 - a. Large surface area (50–100 m^2)
 - b. High blood flow
 - c. Close to blood (10 μm between gas media and blood)
 4. Absorption of gases is dependent on the solubility of the gas in blood:
 - a. Chloroform, for example, has high solubility and is all absorbed; respiration rate is the limiting factor.
 - b. Ethylene has low solubility and only a small percentage is absorbed; blood flow limited absorption.
 - c. Parenteral routes.
 - d. Dermal routes.

As a generalization, there is a pattern of relative absorption rates which extends between the different routes that are commonly employed. This order of absorption (by rate from fastest to slowest and, in a less rigorous manner, in degree to absorption from most to least) is $\text{IV} > \text{inhalation} > \text{IM} > \text{IP} > \text{SC} > \text{oral} > \text{ID} > \text{other dermal}$.

Absorption (total amount and rate), distribution, metabolism, and species similarity in response are the reasons for selecting particular routes in toxicology. In acute studies, however, these things are rarely known to us. So the cardinal rule for selecting routes of use in acute testing is to use those routes which mirror the intended route for human exposure. If this route of human exposure is uncertain or if there is the potential for either a number of routes or the human absorption rate and pattern being greater, then the common practice becomes that of the most conservative approach.

This approach stresses maximizing potential absorption in the animal species (within the limits of practicality) and selecting from among those routes commonly used in the laboratory that which gets the most material into the animal's system as quickly and completely as possible to evaluate the potential toxicity.

In general, drugs cross biological barriers by one of three mechanisms: active transport, facilitated transport, and passive transport. In active transport, the chemical is specifically recognized by the organism, which then expends energy to take the chemical up, even against a concentration gradient. In facilitative transport, the organism produces a carrier molecule which reacts with the target molecule to form a complex which more easily traverses the membrane, but no energy is expended to take up the complex. Such complexes do not flow against a concentration barrier. The simplest mechanism is passive transport or diffusion. Here, a drug flows down a concentration gradient (from high concentration to a lower concentration) and must passively (no energy expended by organism) cross a biological membrane. Passive transfer or diffusion is the most common (if not the only) mechanism involved in the absorption of the vast majority of approved drugs. It should be remembered that for purposes here, "concentration gradient" must be considered in relationship to partition coefficient. That is, a gradient will reflect the relative solubilities of drug in polar (water) and nonpolar (lipid) matrices or tissues. The other mechanisms involved in absorption will not be further discussed here.

Drugs in solution have a natural tendency (more rigorously defined by the laws of thermodynamics) to move down a concentration gradient. That is to say, the individual molecules of solute tend to move from a region of high concentration toward regions of lower concentration. Also, the movement of a chemical across a permeable barrier, such as a biological membrane, is a process called diffusion, as illustrated by Figure 17.1. For most products, these biological barriers are either the wall of the GI tract, the lining of the pulmonary system, or the skin.

Absorption from the GI tract is controlled by a variety of factors. These include the acid/base characteristics of the chemical (described as the pK_a), the solubility, the nature of the delivery (e.g., diet vs. gavage), the nature of any vehicle (suspensions vs. solutions or aqueous vs. nonaqueous), and the GI tract of the species under study. Gad (2007, 2015) provide much greater detail on this subject.

Ionized or charged organic moieties do not readily pass through the lipophilic cell membranes of the epithelial cells that line the GI tract. Thus, more acidic molecules tend to be more readily absorbed from the stomach, while more alkaline materials tend to be absorbed from the small intestine. This is because at the acidic pH of the stomach, acidic chemicals tend to be nonionized. More alkaline chemicals tend to be more ionized in the stomach and less

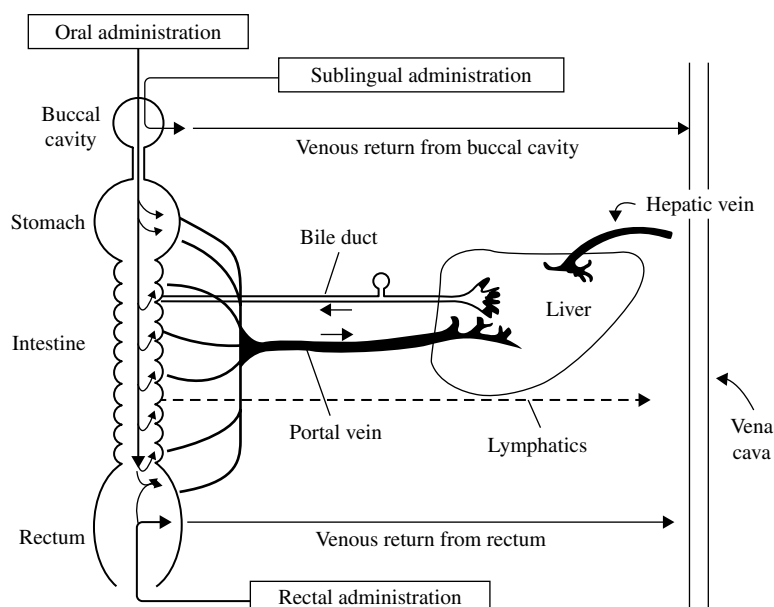


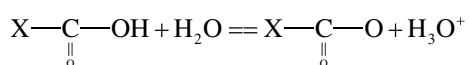
FIGURE 17.1 Passage of chemical moieties from GI tract into bloodstream.

TABLE 17.2 Receptors Slowing Gastric Emptying

| Receptor (Site) | Stimulus | Example | Sensitivity |
|---|------------------------------------|--|--------------|
| Osmoreceptor (duodenum) | Osmotic effect (except tryptophan) | Glucose Amino acids Electrolytes | Least |
| Acid receptor (proximal duodenum and jejunum) | Acids with $pK_a < 5$ | Citric acid HCl | Intermediate |
| Fat receptor (jejunum) | Fatty acids | Sodium myristate | Most |
| Tryptophan receptor (duodenum and jejunum) | Tryptophan | Tryptophan | — |

Source: Adapted from Hunt and Knox (1968) and Minami and MacCallum (1984).

ionized in the gut. The equilibrium reaction for acidic dissociation can be represented by this equation:



Like all chemical equations, this one has an equilibrium constant. The discussion of basic chemistry is outside the purview of this book. Readers who may need a refresher are referred to Tse and Jaffe (1991). For every chemical, a pK_a can be calculated based on equilibrium constant, which represents the proportion of ionized and unionized material in solution. The lower the pK_a of a molecule, the more likely it is to be nonionized in the stomach.

One can manipulate the movement of a drug through the GI tract (and particularly the stomach) by formulation and/or by feeding, concurrent with drug administration. Some of the factors which slow gastric emptying (and therefore some to increase drug absorption) are presented in Table 17.2.

17.3.2.1 Absorption from the Pulmonary System Of the three routes discussed here, absorption from the pulmonary system is perhaps the most rapid. Systemic absorption of inhaled materials is highly dependent on the physical properties of the inhaled materials which dictate how easily the materials reach the alveoli of the deep lung. Proteins may be readily absorbed when instilled in the deep lungs—as an example, note the inhalable insulins. Gases and vapors easily penetrate into the deep lung. For mists and dusts, absorption will be highly dependent on particle size. In general, the larger the particles, the less they will penetrate the pulmonary system. The term impaction describes the deposition of particles in the respiratory tract. Particles of less than $0.2\mu\text{m}$ are preferentially deposited in the pulmonary portion of the respiratory system, whereas particles over $2\mu\text{m}$ do not reach the alveolar epithelium in great number. Particles from 1 to $4\mu\text{m}$ tend to be distributed over the length of the system, whereas particles over $4\mu\text{m}$ tend to be deposited in the nasal region. Aerosolized particles of greater than $20\mu\text{m}$ do not

commonly occur in nature. Tidal volume will also influence impaction. In general, the larger the tidal volume, and thus the more forceful the inhalation process, the more deeply particles of all sizes tend to be driven into the lung.

Once deposited, materials must be in solution before they can be absorbed. Hence, materials in an aerosolized solution will be more readily absorbed than materials that are delivered as solid particles (e.g., dusts). Solid materials must be able to go into solution *in situ* in order to be absorbed. Particle size influences dissolution rate. Large particles dissolve more slowly (for any given material) than small particles due to the differences in surface area. Once in solution, the same laws of passive diffusion apply to materials in the lung as they apply to material in the GI tract. The large surface area and the rich blood flow at the alveoli make for ideal conditions for rapid absorption into the systemic circulation. Absorption across the mucosa lining the upper airways is less rapid. Materials that do not dissolve are ingested by pulmonary macrophages and either broken down there or moved out of the lungs by the upward movement of the bronchociliary tree.

For gases and vapors, the amount absorbed is highly dependent on the partial pressure and solubility of the gas in blood. Let's take the simple case of a gas that is not metabolized and is excreted by exhalation (e.g., an anesthetic gas or a Halon-type fire-extinguishing agent). At any given concentration (or partial pressure) in the atmosphere, the concentration in the blood will reach a steady state. Accordingly, prolonged exposure does not lead to continual buildup.

At equilibrium, the concentration in the blood is depicted by the formula (also known as the Ostwald coefficient) $X_b/X_a = S$, where X_b is the concentration in the blood and X_a is the concentration in the inspired air. Thus, if one knows the S for a given chemical and the target concentration for a given exposure, one can predict what the resulting concentration may be at equilibrium. Additionally, the lower the S value (i.e., the lower the solubility in blood), the more rapidly the chemical will achieve equilibrium.

17.3.2.2 Absorption across the Skin An aqueous carrier may be used for a variety of dermal products. In fact, carriers can be designed to limit the transportation or the penetration of the active ingredient (such as an insect repellent) if the desired effect is to keep the active ingredient on the surface of the skin. Once again, however, only those materials that are dissolved will be available for penetration across the skin to gain access to the systemic circulation. For almost all drugs in or about to enter clinical trials, dermal penetration is a passive process. The relative thickness of the skin makes absorption (into the systemic circulation) slower than the absorption across the GI or pulmonary barriers. This is compounded by the fact that the stratum corneum function is to be impervious to the environment. One of the skin's major

functions is protection from infection. Once a drug penetrates into the dermis, it may partition into the subcutaneous fat. Essentially, absorption across the skin is a two-step process with the first being penetration and deposition into the skin and the second being release from the skin into the systemic circulation. The pattern of blood levels obtained via dermal penetration is generally one with a delayed absorption and slow buildup to more of a plateau than a peak. Blood levels of chemicals absorbed via the dermal route are generally low (Garner and Mathews, 1998).

Given the overwhelming influence of the physical properties of skin in determining bioavailabilities via the dermal route, assessment of dermal penetration is one area in metabolism and toxicology where the use of *in vitro* methods can be effectively used to predict *in vivo* results and to screen chemicals. Apparatus and equipment that one can use to maintain sections of skin (obtained from euthanized animals or from human cadavers or surgical discard) for such experiments exist (Holland et al., 1984; Bronaugh, 1998). These apparatus are set up to maintain the metabolic integrity of the skin sample between two reservoirs: the one on the stratum corneum side, called the application reservoir, and the one on the subcutaneous side, called the receptor reservoir. One simply places radiolabeled test material in the application reservoir and collects samples at various time points from the receptor fluid.

The rate of penetration can be presented by the traditional kinetic formulas to obtain a penetration rate constant. Given that exposed surface area also plays a role in the amount of material absorbed, the concept of flux is also important.

Major considerations in determining the quantity of material that is absorbed into the skin, and eventually released into the systemic circulation, are primarily dependent on three factors: the surface area exposed, the volume and concentration of the material applied, and the nature of the vehicle:

Surface area all things being equal, it is clear that the greater the surface exposed, the higher the achieved internal dose.

Volume the volume of material will obviously play a role in total dose, but it is not as straightforward as the relationship to surface area. Theoretically, the maximum absorption is obtained when the material is spread as thin and uniform as possible; piling material on so that it is literally rolling off the animal serves no practical purpose. In fact, it is not a sound practice when dealing with an *in vivo* animal experiment as it makes it more likely for the material to be available for oral ingestion.

Concentration the higher the concentration in a formulation, the higher the flux achieved of drug molecules across the skin. Of course, the nature of the vehicle that the drug is being carried in may also have a profound influence on absorption.

17.3.2.3 Parameters Controlling Absorption The absorption of a chemical into the skin is a function of the nature of the molecule, the behavior of the vehicle, and the status of the skin. Three major variables account for differences in the rate of absorption or flux of different topical chemicals or of the same molecule in different vehicles: the concentration of the molecule in the vehicle, the partition coefficient of chemical between the stratum corneum and the vehicle, and the diffusion coefficient of molecule in the stratum corneum.

The rate of diffusion is proportional to the concentration of molecule in the vehicle. The relationship is linear only at low molecular concentrations and only applies to soluble molecule in the vehicle. The latter factor may explain the variable therapeutic effects of different formulations of the same drug molecule. The partition coefficient is a measure of the molecule's ability to escape from the vehicle and is defined as the equilibrium solubility of molecule in the surface of the stratum corneum relative to its solubility in the vehicle. Increased lipid solubility favors penetration of molecule through the skin by increasing the solubility in the relatively lipophilic stratum corneum. The diffusion coefficient indicates the extent to which the matrix of the barrier restricts the mobility of the molecule. Increases in molecular size of the molecule will increase frictional resistance and decrease the diffusion coefficient (Bronaugh, 1998); molecules over 1000 Da usually will not be absorbed easily into normal adult skin.

Finally, intact stratum corneum is an excellent barrier, but in disease states that compromise the skin barrier, the resistance to absorption is rapidly lost and absorption can be facilitated. Such compromised skin can be humanly simulated by using either a dermatome or tape stripping the skin site in question.

17.3.3 Distribution

Once the chemical gains access to the body, it is carried by the bloodstream and distributed to the different organs. The preferential organ of deposition is determined by a variety of factors: the two most important are blood flow to the organ and the affinity of the chemical for that organ. Affinity is governed by two general characteristics. Firstly, the product may be designed to have a specific affinity for a specific molecular entity in a target cell. For example, an anticholinesterase insecticide will tend to accumulate in the cells that have the highest concentration of cholinesterase. Secondly, the product may have a nonspecific or general chemical attraction for a specific cell type. The more highly lipophilic a chemical is, the more likely it is to distribute and remain in adipose tissue. Blood flow will also have a major impact on distribution, as chemicals will be distributed more readily to those organs that are more highly perfused. A highly lipophilic chemical may first be deposited in the

brain due to the fact that it is richly perfused and then be distributed to body fat with time.

Once a material is absorbed, distribution of a compound in most early toxicology studies is usually of limited interest. This is unfortunate, as it is the preferential distribution of the drug to the therapeutic target that is desired. Some factors which can serve to alter distribution are listed in Section 17.3.3.

For most drugs, the rate of disposition or loss from the biological system is independent of rate and input, once the agent is absorbed. Disposition is defined as what happens to the active molecule after it reaches a site in the blood circulation where concentration measurements can be made (the systemic circulations, generally). Although disposition processes may be independent of input, the inverse is not necessarily true because disposition can markedly affect the extent of availability. Agents absorbed from the stomach and the intestine must first pass through the liver before reaching the general circulation (Figure 17.1). Thus, if a compound is metabolized in the liver or excreted in bile, some of the active molecule absorbed from the GI tract will be inactivated by hepatic processes before it can reach the systemic circulation and be distributed to its sites of action. If the metabolizing or biliary excreting capacity of the liver is great, the effect on the extent of availability will be substantial. Thus, if the hepatic blood clearance for the chemical is large, relative to hepatic blood flow, the extent of availability for this chemical will be low when it is given by a route that yields first-pass metabolic effects.

Likewise, metabolism is generally of only limited concern in most acute studies. There are some special cases, however, in which metabolic considerations must be factored in seeking to understand differences between routes and the effects which may be seen.

The first special case is parenteral routes, where the systemic circulation presents a peak level of the moiety of interest to the body at one time, tempered only by the results of a single pass through the liver.

The second special case arises from inhalation exposures. Because of the arrangements of the circulatory system, inhaled compounds enter the full range of systemic circulation without any "first-pass" metabolism by the liver. Keberle et al. (1971) and O'Reilly (1972) have published reviews of absorption, distribution, and metabolism that are relevant.

17.3.3.1 Protein Binding The degree to which a drug binds to plasma proteins will highly influence its distribution. Albumin, the most prominent of the many proteins found in mammalian plasma, carries both positive and negative charges with which a polar compound can associate by electrostatic attraction. As with all such reactions, it can be described by the following equations. The more avidly bound the material, the less will be distributed to surrounding fluids as part of a solution and only that portion that is free in solution will be available for diffusion into the tissues.

17.3.3.2 Water Solubility The solubility of a chemical has a direct bearing on its distribution. Recall that only molecules that are in solution will be available for absorption.

As mentioned earlier, only that portion that is free in solution will be available for diffusion into the tissues. Hence, the more material that is in solution, the more that will be available for diffusion.

17.3.3.3 Volume of Distribution If one takes the dose administered (milligrams) and divides it by the plasma concentration of the test material (milligram per milliliter), the result is a volume number:

$$\frac{\text{Dose}}{\text{Concentration}} = \text{Volume}$$

One can take this process a step further and extrapolate back from a plasma time curve to the y axis. This is theoretically the plasma concentration (C_0) that would occur if, upon being administered, the material is instantly distributed throughout the body. The volume number obtained with the preceding equation becomes

$$\frac{\text{Dose}}{C_0} = V_D$$

where V_D represents the apparent volume of distribution, a proportionality constant that reflects the relation of the concentration of a xenobiotic in plasma to the total amount

of the entity in the body. Materials that are avidly bound to plasma proteins will have a high volume of distribution, while materials that are avidly taken by the tissues (e.g., deposit fat) will have a low volume of distribution. V_D is a parameter that is simple to calculate yet gives one important information about the distribution of the chemical under investigation.

The available volumes and masses for distribution vary from species to species, as summarized in Tables 17.3 and 17.4.

17.3.4 Metabolism/Biotransformation

Metabolism describes the process by which chemicals are changed by the body. In fact, very few foreign chemicals that come to enter the body are excreted unchanged. Most are chemically modified. In general, metabolism results in chemicals that are more polar and water-soluble, and more easily excreted (La Du et al., 1972). Examples of more common metabolic conversions are shown in Table 17.5. In general, the vast majority of lipophilic chemicals are first oxidized via the CYP-dependent mixed function oxidase system of the liver. This is the process classically called phase I metabolism. CYP exists as a family of isozymes (the CYP gene superfamily) with varying but overlapping substrate affinities and varying responses to different inducing agents. For a review of the molecular biology of the CYP gene superfamily, the reader is referred to Meyer (1994). Induction is the process whereby exposure to a chemical leads to increased activity of the mitochondrial

TABLE 17.3 Volume and Half-Life of Body Water in Selected Species

| Species | Sex | Exchangeable Body Water (% of Body Weight) | Half-Life (Days) |
|-------------------|-----|---|------------------|
| Mouse | F | 58.5 | 1.13 |
| Rat | M | 59.6 | 2.53 |
| Rabbit | F | 58.4 | 3.87 |
| Dog | M | 66.0 | 5.14 |
| Cynomolgus monkey | M | 61.6 | 7.80 |
| Rhesus monkey | M | 61.6 | 7.80 |
| Man | M,F | 55.3 | 9.46 |

TABLE 17.4 Typical Organ Weights in Adult Laboratory Animals

| Organ | Percent of Body Weight | | | | |
|----------|------------------------|-------|------|--------|--------|
| | Rat | Mouse | Dog | Rabbit | Monkey |
| Liver | 3.5 | 6 | 3.5 | 3 | 2.5 |
| Kidney | 0.8 | 1.6 | 0.5 | 0.8 | 0.5 |
| Heart | 0.4 | 0.4 | 0.8 | 0.3 | 0.4 |
| Spleen | 0.3 | 0.5 | 0.3 | 0.04 | 0.1 |
| Brain | 0.5 | 0.6 | 0.8 | 0.4 | 3 |
| Adrenals | 0.02 | 0.01 | 0.01 | 0.02 | 0.03 |
| Lung | 0.6 | 0.6 | 1 | 0.6 | 0.7 |

TABLE 17.5 Summary of Prominent Phase I Biotransformation Reactions

| Reaction | Enzyme | Location | Example/Comments |
|------------|---------------------------------|----------------------|--|
| Hydrolysis | Carboxylesterase | Ubiquitous | Vinyl acetate to acetate and acetaldehyde |
| | Peptidase | Blood, lysosomes | Amino-, carboxy-, and endopeptidase which cleave peptides at specific amino acid linkages |
| Reductions | Epoxide hydratase | Microsomes, cytosol | Conversion of styrene-7,8-epoxide to styrene-7,8-glycol |
| | Azo and nitro reduction | Gut microflora | Sequential conversion of nitrobenzene to aniline |
| | Carbonyl reductase | Cytosol | Conversion of haloperidol to reduced haloperidol (a secondary alcohol) |
| | Disulfide reduction | Cytosol | Glutathione-dependent reduction of disulfiram to diethyldithiocarbamate |
| Oxidation | Sulfoxide reduction | Cytosol | Thioredoxin-dependent reduction of sulindac to sulindac sulfide |
| | Quinone reduction | Cytosol, microsomes | DT-diaphorase reduction of menadione to hydroquinone |
| | Reductive dehalogenation | Microsomes | Conversion of pentabromoethane to tetrabromoethane (releasing free bromide ion) |
| | Alcohol dehydrogenase | Cytosol | Conversion of ethanol to acetaldehyde (DAD/DADH-dependent reversible reaction) |
| | Aldehyde dehydrogenase | Mitochondria/cytosol | Conversion of acetaldehyde to acetate |
| | Aldehyde oxidase | Liver cytosol | FAD-dependent metalloenzyme, oxidation of benzaldehyde to benzoic acid |
| | Xanthine oxidase | Cytosol | Oxidation of purine derivative, conversion of allopurinol to alloxanthine |
| | Monoamine oxidase | Mitochondria | FAD-dependent oxidative deamination of monoamines, for example, primaquine |
| | Diamine oxidase | Cytosol | Pyridoxal-dependent, copper-containing enzyme. Conversion of allylamine to acrolein |
| | Prostaglandin oxidase | Microsomes | Cooxidation reaction can activate chemicals in tissues low in cytochrome P450, for example, nephrotoxicity of acetaminophen, oxidation of phenylbutazone |
| | Flavin-containing monooxygenase | Microsomes | FAD-dependent oxidation of nucleophilic nitrogen, sulfur, and phosphorus heteroatoms, for example, conversion of nicotine to nicotine 1'-N-oxide, cimetidine to cimetidine S-oxide |
| | Cytochrome P450 | Microsomes | |

mixed-function oxidase (MMFO) system due to an increase in the amount of the CYP isozyme. The isoenzymes induced by a variety of chemicals are given in Table 17.6, and example compounds which inhibit specific CYPs are presented in Table 17.6. In a practical sense, a drug can induce its own metabolism. Hence, repeated dosing with a chemical may lead to lower blood levels at the end, for example, of a 13-week study than at the beginning. There could also be alterations in the spectrum of metabolites produced, such that an agent could become more, or less, toxic with repeated dosing depending on the nature of the metabolites. It is not unusual during a subchronic or chronic toxicity test for tolerance to occur. There may be signs of toxicity early in the study, but even with continued daily dosing, the signs abate. This phenomenon, particularly in rodents, is frequently due to microsomal induction, whereby the chemical has induced its own metabolism, and more rapid clearance of the parent chemical occurs. It should be noted that the P450 system is not the only drug-metabolizing system. As Table 17.7 summarizes, there are at least five major metabolic systems in mammals.

TABLE 17.6 Examples of Xenobiotics Metabolized by Human P450

| | |
|--------|--|
| CYP1A1 | Benzo[a]pyrene and other polycyclic aromatic hydrocarbons |
| CYP1A2 | Acetaminophen 2-Acetylaminofluorene 4-Aminobiphenyl 2-Aminofluorene 2-Naphthylamine Amino acid pyrolysis products (DiMeQx, MeIQ, MeIQx, Glu P-1, Glu P-2, IQ, PhIP, Trp P-1, Trp P-2) |
| CYP2A6 | 6-Aminochrysene Cyclophosphamide Ifosfamide N-Nitrosodimethylamine |
| CYP2B6 | 6-Aminochrysene Cyclophosphamide Ifosfamide |

(Continued)

TABLE 17.6 (Continued)

CYP2C8, 9, 18 (Note: 2C9 is absent in 15–30% of Asians)

Tolbutamide

Taxol

CYP2C19

Diazepam

Diphenylhydantoin

Hexobarbital

Propranolol

CYP2D6 (Note: absent in 7% of Caucasians)

Bufuralol

CYP2E1

Acetaminophen

Acrylonitrile

Benzene

Carbon tetrachloride

Chloroform

Chlorzoxazone

Dichloromethane

1,2-Dichloropropane

Ethylene dibromide

Ethylene dichloride

Ethyl carbamate

N-Nitrosodimethylamine

Styrene

Trichloroethylene

Vinyl chloride

CYP3A4 (50% of all marketed drugs are marketed here)

Acetaminophen

Aflatoxin B₁ and G₁

6-Aminochrysene

Benzo[*a*]pyrene 7,8-dihydrodiol

Cyclophosphamide

Ifosfamide

Midazolam

Nifedipine

Testosterone

1-Nitropyrene

Sterigmatocystin

Senecionine

tris(2,3-Dibromopropyl) phosphate

CYP4A9/11

None known

After the drug has been metabolically oxidized, it can be further metabolized. In fact, it is possible for the metabolites to also be substrates of the MMFO or the other metabolic systems shown in Table 17.7 and thus be further metabolized themselves.

The route of metabolic activation of the classic carcinogen benzo[*a*]pyrene is due to such a mechanism. The biology of these reactive intermediates has been extensively studied. Glutathione is among the most common organic intracellular chemicals in all mammalian species, being present at a concentration of up to 10 mM, and glutathione *S*-transferase is very active. Glutathione is a tripeptide (glutamine–cysteine–glycine). The sulfhydryl group of cysteine is the business end of the molecule where the reaction with the nucleophilic reactive intermediate takes place. After that, the glutathione conjugate is further metabolized to a cysteinyl-acetyl moiety. These moieties are called mercapturic acids and are generally found in the urine. The relative predominance of mercapturic acid over other metabolites may be considered a rough indication of how “reactive” the intermediates may have been. Teleologically, it is tempting to speculate that it is a very well-designed protective mechanism. So long as intracellular glutathione concentrations remain above a critical level, the destructive actions of active metabolites can be held in check. Thus, a small dose of a chemical (bromobenzene is a good example) may cause no liver damage, while a large dose may. This is also a good example of one of the aspects of TKs versus PKs where a high dose of a chemical can become toxic due to saturation of a detoxification pathway.

The glutathione *S*-transferase pathway is sometimes in biochemical competition with the epoxide hydratase pathway in that both deactivate intermediates of the MMFO system. Epoxide hydratase is a microsomal enzyme that acts specifically to deactivate epoxide intermediates by the addition of water across the C–O bond to form a diol. As a very broad generality, the glutathione *S*-transferase pathway tends to be more prominent in rodents, while the epoxide hydratase pathway tends to be more dominant in nonrodents.

TABLE 17.7 A Comparison of the Key *In Vitro* Drug-Metabolizing Experimental Systems (Liver Microsomes), Liver Postmitochondrial Supernatant (S9), Liver Cytosol, and Hepatocytes and Their Contents of Major Drug-Metabolizing Enzymes

| <i>In Vitro</i> System | P450 | MAO | UGT | ST | GST |
|------------------------|------|-----|--------------|--------------|--------------|
| Microsomes | + | – | ^a | – | ^b |
| S9 | + | – | ^a | ^a | + |
| Cytosol | – | – | ^a | ^a | ^c |
| Hepatocytes | + | + | + | + | + |

GST, glutathione *S*-transferase; MAO, monoamine oxidase; P450, cytochrome P450 isoforms; ST, sulfotransferase; UGT, UDP-glucuronosyltransferase.

^a Activity of this drug-metabolizing enzyme requires the addition of specific cofactors, for instance, UDP-glucuronic acid (UDPGA) for UGT activity and 3′-phosphoadenosine 5′-phosphosulfate (PAPS) for ST activity.

^b Membrane-bound GST but not the soluble GST are found in the microsomes.

^c Soluble GST but not membrane-bound GST are found in the cytosol.

The hydroxyl- or diol-containing metabolites of the MMFO can be further metabolized by so-called phase II (synthetic) metabolism whereby they are conjugated to/from glucuronides and/or sulfates (so-called ethereal sulfates). Amines can also be substrates. The net effect of phase II reactions is to create a more polar molecule that is more water-soluble and therefore readily excretable. While there are species differences, glucuronides are actively transported and excreted in the bile into the GI tract. Sulfates are excreted more predominantly in the urine. Both glucuronides and sulfates, however, can be found in both the urine and the feces. Like the MMFO pathway, glutathione *S*-transferase, UDP-glucuronyl transferase, and epoxide hydratase are inducible; that is, treatment with exogenous chemicals will increase the amount of enzyme proteins present.

Outside of the MMFO-mediated (phase I) reactions, there are a few other major reactions that are worthy of note. The two major ones involve ester hydrolysis and alcohol and aldehyde dehydrogenases. All mammalian species have an extensive ability to hydrolyze the ester bond. The products of the reactions then can go on to be further metabolized. In the pharmaceutical industry, this property has been utilized to synthesize prodrugs, that is, chemicals that have desirable pharmaceutical properties (generally increased water solubility) and are not converted to their active moiety until hydrolyzed in the body.

The activity of alcohol dehydrogenase is one with which we should all be familiar. It oxidizes alcohols to aldehydes. The aldehydes produced by this reaction can go on to be further metabolized to a carboxylic acid, if they are not sterically hindered. Side chain constituents of aromatic compounds can also be substrates for this reaction sequence, producing side chain carboxylates. The oxidation of alcohols to aldehydes can also be a form of metabolic activation as aldehydes can have potent physiological actions. Fortunately, aldehyde dehydrogenase has a very high activity when compared to alcohol dehydrogenase, so that the aldehydes do not accumulate. Inhibition of aldehyde dehydrogenase by disulfiram (Antabuse) leads to the accumulation of acetaldehyde, causing nausea, dizziness, and flushing. Like disulfiram, some pesticides contain dithiocarbamates and have the potential of causing this type of reaction.

Hopefully, this brief description of the major metabolic pathways has given one some appreciation of the richness of the processes. The different sites of oxidation, the possibility of additional oxidative metabolism of metabolites, and the differences in phase II reactions all lead to a multiplicity of possible metabolites. Over 100 different metabolites of the human pharmaceutical chlorpromazine have been isolated and identified. When analyzed by high-performance liquid chromatography (HPLC), for example, the parent chemical and the different (detectable) metabolites will form a pattern of different peaks. This is referred to as the metabolic fingerprint or profile of a chemical. Different species may have

different profiles. Ideally, in doing a risk assessment, one would like to know the similarity between the metabolic pattern in animals used in the toxicology studies and that produced by human beings. This is only infrequently available for most nonpharmaceutical products; pesticides, for example, are rarely given intentionally to human subjects for the purposes of such a metabolic study. The technology now exists, however, to address this potential problem. Cell lines with human CYP have been developed that can provide some indication of the similarities between human metabolism of a chemical and that of experimental animals. At least they may be able to assist in identifying the major oxidative metabolite(s). For nonpharmaceutical products, it may be an unusual circumstance that would require one to identify potential human metabolites as part of a marketing application; however, it may be useful for one to know that the technology exists to do so.

The processes of metabolic conversion are frequently involved in the mechanisms of toxicity and carcinogenicity.

17.3.4.1 Metabolic Activation As mentioned, most nonnutritive chemicals pass through the GI tract by passive absorption and then enter the mesenteric circulation. The venous circulation from the mesentery flows through the portal vein into the liver. The metabolic action of the liver literally sits between the GI tract and the general systemic circulation. Thus, even chemicals that may be highly absorbed from the GI tract could appear only sparingly in systemic circulation if they are highly metabolized by the liver. The combination of absorption from the GI tract and intestinal and hepatic metabolism leads to what is called the first-pass effect. An extension of this is the fact that the gut flora contain glucuronidases, which can cleave glucuronides of chemicals and/or metabolites that are then available to be reabsorbed. This process is called enterohepatic circulation.

17.3.4.2 Induction of P450 Metabolism and Isoenzymes When organisms are exposed to certain xenobiotics, their ability to metabolize a variety of chemicals is increased. This phenomenon can produce either a transitory reduction in the toxicity of a drug or an increase (if the metabolite is the more toxic species). However, this may not be the case with compounds that require metabolic activation. The exact toxicological outcome of such increased metabolism is dependent on the specific xenobiotic and its specific metabolic pathway. Since the outcome of a xenobiotic exposure can depend on the balance between those reactions that represent detoxification and those that represent activation, increases in metabolic capacity may at times produce unpredictable results.

The ability of different drugs to differentially inhibit and/or induce individual CYP isoenzymes has become critical in assessing the potential safety of drug molecules. Table 17.8

TABLE 17.8 Examples of Xenobiotics Activated by Human Cytochrome P450 Isoenzymes

| | |
|----------------------------------|---|
| CYP1A1 | CYP2D6 |
| Benzo[a]pyrene and other | Buforolol |
| Polycyclic aromatic hydrocarbons | Codeine |
| CYP1A2 | Timolol |
| Acetaminophen | Metoprolol |
| 2-Acetylaminofluorene | CYP2E1 |
| 4-Aminobiphenyl | Acetaminophen |
| 2-Aminofluorene | Acrylonitrile |
| 2-Naphthylamine | Benzene |
| CYP2A6 | Carbon tetrachloride |
| N-Nitrosodiethylamine | Chloroform |
| Butadiene | Dichloromethane |
| Coumarin | 1,2-Dichloropropane |
| CYP2B6 | Ethylene dibromide |
| 6-Aminochrysene | Ethylene dichloride |
| Cyclophosphamide | Ethyl carbamate |
| Ifosfamide | N-Nitrosodimethylamine |
| CYP2C8 | Styrene |
| Taxol | Trichloroethylene |
| CYP2C9 | Vinyl chloride |
| Diclofenac | CYP3A4 |
| Phenytoin | Acetaminophen |
| Piroxicam | Aflatoxin B ₁ and G ₁ |
| Tolbutamide | 6-Aminochrysene |
| CYP2C19 | Benzo[a]pyrene 7,8-dihydrodiol |
| Diazepam | Cyclophosphamide |
| Diphenylhydantoin | Ifosfamide |
| Hexobarbital | 1-Nitropyrene |
| Propranolol | Sterigmatocystin |
| | Senecionine |
| | tris(2,3-Dibromopropyl) phosphate |
| | CYP4A9/11 |
| | None known |

Source: Adapted from Parkinson (1996).

presents an overview of some of what we have come to know about differential metabolism by P450 isoenzymes. Draper et al. (1998) had published on the use of human liver microsomes for determining the levels of activity or inhibition a drug has on the formation of 6 β -hydroxytestosterone as a model for CYP3A activity (1) and chlorzoxazone for CYP2E1 activity (2). If, for example, a chemical under study competitively inhibits the metabolism of these model substrates in these systems, then it is a substrate for that human isozyme. Using these more recently available *in vitro* systems, it is much easier to perform cross-species comparisons with regard to biotransformation. It is now easier to determine how similar the routes of metabolism are in the experimental animals with comparison to those in man without having to administer the chemical to human subjects. Human and animal microsome preparations may be used as models to identify patterns of metabolites *in vitro*, allowing for better selection of model species for safety studies, and competition for or inhibition of activation of

specific isoenzymes can be evaluated to identify potential problems of DDI in patients (Levy et al., 2000).

17.3.4.3 Species Differences Species differences in metabolism are among the principal reasons that there are species differences in toxicity. Differences between species in CYP isozyme content and amounts are one of the most common reasons for differences in metabolism. For example, Monostory et al. (1997) published a paper comparing the metabolism of panomifene (a tamoxifen analog) in four different species. These data serve to address that the rates of metabolism in the nonhuman species were most rapid in the dog and slowest in the mouse. Thus, one should not *a priori* make any assumptions about which species will have the more rapid metabolism. Of the seven metabolites, only one was produced in all four species. Both the rat and the dog produced the two metabolites (M5 and M6) produced by human microsomes. So how does one decide which species best represents man? One needs to consider the chemical

structure of the metabolites and the rates at which they are produced. In this particular case, M5 and M6 were relatively minor metabolites in the dog, which produced three other metabolites in larger proportion. The rat produced the same metabolites at a higher proportion, with fewer other metabolites than the dog. Thus, in this particular instance, the rat, rather than the dog, was a better model. Likewise, Table 17.9 offers a comparison of excretion patterns between three species for a simple inorganic compound. Table 17.10 presents a summary of interspecies differences between species in drug metabolism and PKs, and Table 17.11 presents differences in total liver concentrations of CYP enzymes in each of six mammalian species.

A more thorough review on species differences in PKs has been presented by Smith (1991), Gad and Chengelis (1998), and Gad (2015).

17.3.4.4 Sex-Related Differences in Rodents Not only are there differences in absorption, distribution, and metabolism between species, but also there may be differences between sexes within a species (Mugfor and Kidderis, 1998).

TABLE 17.9 Differences in the Disposition of 2,4-Dichlorophenoxyacetic Acid

| Species | Sex | Urine | Feces |
|---------|-----|-------|-------|
| Rat | M | 31.2 | 2.7 |
| | F | 16.5 | 1.1 |
| Mouse | M | 12.7 | 2.8 |
| | F | 26.8 | 6.7 |
| Hamster | M | 4.9 | 2.5 |
| | F | 33.9 | 14.5 |

Note: All animals dosed orally with radiolabeled 2,4-D, 200mg kg⁻¹. Results are expressed as percent of ¹⁴C dose recovered. Urine was collected for 8h and feces for 24h.

TABLE 17.10 “General Rules” on Interspecies Differences in DMPK

| Species | DMPK Characteristics |
|-----------|---|
| Human | Polymorphisms (e.g., CYP 2C9, CYP 2C19, CYP 2D6, NAT1, NAT2) |
| Dog | Low acetylation, high capacity for deacetylation Different absorption due to higher pH in gastrointestinal tract than in humans (consider use of synthetic gastric fluid to mimic human situation) |
| Rat | Often gender differences which are not observed in other species Abundant tetrahydrofolate (protects, e.g., against methanol ocular damage) |
| Rabbit | Low sulfation |
| (Mini)Pig | Low sulfation Gastrointestinal conditions similar to humans |
| Cat | Low glucuronidation High sulfation |

TABLE 17.11 Total Liver Content of CYP Enzymes in Humans, Monkeys, Pigs, and Dogs

| Species | Total Liver Content of CYP (nmol/mg Protein) | Reference |
|--------------------|--|---|
| Humans | 0.29±0.06 (n=12) | Stevens et al. (1993), |
| | 0.307±0.16 (n=18) | Shimada et al. (1997) |
| Rhesus monkeys | 0.95±0.08 (n=6) | Stevens et al. (1993) |
| Cynomolgus monkeys | 1.03±0.11 (n=5) | Shimada et al. (1997) |
| Minipigs | 0.81±0.15 (n=9) | Nebbia et al. (2003) |
| Pigs | 0.22±0.12 (n=3) | Shimada et al. (1997), |
| | 0.46±0.07 (n=12) | Myers et al. (2001), Skaanild and Friis (1999) |
| Dogs | 0.39±0.04 (n=6) | Shimada et al. (1997) |

Source: Data from Dalgaard (2015, pp. 80–92).

Griffin et al. (1997), for example, had demonstrated sex-related differences in the metabolism of 2,4-dichlorophenoxyacetic acid (2,4-D) (Table 17.9). They noted that while there were differences between sexes, they tended to be quantitative (rates), not qualitative (metabolites). Differences between species were greater than sex-related differences. With regard to sex-related differences, it is noteworthy that males do not always have the higher rates, as Griffin et al. have shown; in hamsters, the females metabolize 2,4-D more rapidly than males. In general, male rats tend to have higher activity than female rats, especially with regard to CYP-dependent activity.

In the case of 2,4-D, the only urinary metabolite is 2,4-D-glucuronide, but the half-life of 2,4-D was 138 min in males and 382 min in females.

17.3.5 Excretion

Excretion encompasses the process by which chemicals or their metabolites are transported out of the body. There are three possible major routes of excretion and a handful of minor ones. The major routes of excretion for chemicals, and in particular their metabolites, are:

17.3.5.1 Urine The kidneys filter the entire cardiac output multiple times each day and thus provide a large opportunity for the removal of chemicals from the bloodstream. How much of a xenobiotic is actually excreted is dependent on three factors or processes:

1. The glomerular membrane has pores of 70–80 Å, and under the positive hydrostatic conditions in the glomerulus, all molecules smaller than about 20000 Da are filtered. Proteins and protein-bound compounds thus remain in the plasma, and about 20% of the nonbound entity is carried with 20% of the plasma water into the glomerular filtrate.

2. Because the glomerular filtrate contains many important body constituents (e.g., glucose), there are specific active uptake processes for them. Also, lipid-soluble chemicals diffuse back from the tubule into the blood, especially as the urine becomes more concentrated because of water reabsorption. The pH of the urine is generally lower than that of the plasma, and therefore pH partitioning tends to increase the reabsorption of weak acids. The pH of the urine can be altered appreciably by treatment with ammonium chloride (decreases pH) or sodium carbonate (increases pH); the buffered plasma shows little change.
3. Xenobiotics may be secreted actively into the renal tubule against a concentration gradient by anion and cation carrier processes. These processes are saturable and of relatively low specificity; many basic or acidic compounds and their metabolites (especially conjugation products) are removed by them. Because the dissociation rate for the chemical–albumin complex is rapid, it is possible for highly protein-bound compounds to be almost completely cleared at a single passage through the kidney.

17.3.5.2 Feces The most important mechanism allowing circulating foreign compounds to enter the gut is in the bile. The biological aspects of this mechanism have been reviewed, and certain pertinent points have emerged. The bile may be regarded as a complementary pathway to the urine, with small molecules being eliminated by the kidney and large molecules in the bile. Thus, the bile becomes the principal excretory route for many drug conjugates. Species differences exist in the molecular weight requirement for significant biliary excretion, which has been estimated as 325 ± 50 in the rat, 440 ± 50 in the guinea pig, and 475 ± 50 in the rabbit. In the rat, small molecules (<350 Da) are not eliminated in the bile or large molecules (>450 Da) in the urine, even if the principal excretory mechanism is blocked by ligation of the renal pedicles or bile duct, respectively (Wang and Reuning, 1994). Compounds of intermediate molecular weight (350–450Da) are excreted by both routes, and ligation of one pathway results in increased use of the other.

Foreign compounds may also enter the gut by direct diffusion or secretion across the gut wall, elimination in the saliva, pH partitioning of bases into the low pH of the stomach, and elimination in the pancreatic juice.

17.3.5.3 Expired Air Volatile compounds or metabolites can be extensively excreted by passage across pulmonary membranes into the airspace of the lungs and then expulsion from the lungs in expired air.

Minor routes for excretion can include tears, saliva, sweat, exfoliated keratinocytes, hair, and nasal discharge. These are of concern or significance only in rare cases. Accordingly, quantitation of excretion typically requires

collection of urine and feces (and occasionally expired air) over a period of time.

17.4 PHARMACOKINETICS

The interplay of absorption, distribution, metabolism, and excretion processes results in changes in the concentration of the test chemical in different organs with time (Shargel and Yu, 1999; Roberts and Renwick, 2014). With regard to the practical concerns of monitoring human exposure, the organ of interest is the blood. Blood is generally considered the central compartment. Determining the concentration of the chemical in plasma gives one an assessment of exposure. Mathematical formulas are used to quantitatively describe this exposure (Bauer, 2001).

17.5 LABORATORY METHODS

The actual means by which PK information is collected is through the conduct of one or more of specific studies, employing a wide range of available analytical techniques. Administered therapeutic molecules can be identified and quantified in relevant samples collected in accordance with carefully designed and executed protocols. While it is most common to use either rats or mice as the rodent species in all cases, for small molecules, nonrodent species (dog, monkey, or minipig) are most commonly selected based on the compatibility of their metabolism of the molecule of interest to that of humans. There has been a lot of work on the capabilities of these species in recent years (Dalgaard, 2015)—see Table 17.8.

17.5.1 Analytical Methods

There are three broad categories of analytical techniques now available—instrumental (cold chemical), radiolabeled, and immunological. Each of these has advantages and disadvantages. Only an overview of these techniques will be given here—detailed explanations are beyond the scope of this text. These methodologies are all directed at being able to identify and/or quantify a chemical (and/or its metabolites) in various biological matrices.

17.5.1.1 Instrumental Methods These bioanalytical methods are also sometimes called cold chemistry methods (Crooks et al., 2014). These generally start from a place of isolating the compound or compounds of interest, for which the workhorse methodology is HPLC. A wide variety of specialized columns are used to achieve desired separation. At the end of the column, where separation of molecular entities has been achieved, the outflow of the column can be directed to any of a wide variety of detection instruments,

including various forms of detectors intrinsic to the HPLC. In general, all of the cold chemistry methodologies have less sensitivity (higher detection limits) than radiochemical or immunological methods.

Mass spectrometry (MS); nuclear magnetic resonance (NMR) spectroscopy; electron spin resonance (ESR) spectroscopy; ultraviolet, infrared, and visible spectrophotometry; and mass spectroscopy are all well-established detection methodologies.

17.5.1.2 Radiochemical Methods The massive expansion of our understanding of TKs since the late 1970s is to a large degree a reflection of the wide use of radioactive isotopes as tracers of chemical and biological processes. Appropriately radiolabeled test compounds are commonly used in TK studies, providing a simple means of following the administered dose in the body. This is particularly important when specific analytical methods are unavailable or too insensitive. The use of total radioactivity measurements allows an estimation of the total exposure to drug-related material and facilitates the achievement of material balance and the routes of excretion.

The most commonly used radionuclides in drug metabolism and disposition studies are carbon-14 (^{14}C) and tritium (^3H), both of which are referred to as beta emitters. Since these beta-emitting isotopes have relatively long half-lives,

their radioactive decay during an experiment is insignificant. Additionally, they provide sufficient emission energy for measurement and are relatively safe to use, as indicated by the data in Table 17.12. Although individual beta particles can have any energy up to the maximum, E_{max} , the basic quantity in determining the energy imparted to tissues by beta emitters is the average energy, E_{β} . The range is the maximum thickness the beta particles can penetrate. Beta particles present virtually no hazard when they originate outside the body (Shapiro, 1981).

During the synthesis of radiolabeled compounds, the label is usually introduced as part of the molecular skeleton in a metabolically stable and, with tritium, nonexchangeable position. The in vivo stability of ^{14}C labels is often reflected by the extent of $[^{14}\text{C}]$ carbon dioxide formation. The biological stability of ^3H labels can be estimated by the extent of tritiated water formation. The tritiated water concentration (dpm mL^{-1}) in urine samples collected during a designated time interval after dosing, assumedly after equilibrium is reached between urine and the body water pool, is determined. This value is extrapolated from the midpoint of the collection interval to zero time, based on the known half-life of tritiated water in the given species. The percentage of the radioactive dose that is transformed to tritiated water ($\% \text{ } ^3\text{H}_2\text{O}$) can be calculated using the following equation:

$$\text{Percent } ^3\text{H}_2\text{O} = \frac{^3\text{H}_2\text{O concentration at zero time} \times \text{exchangeable body water volume}}{\text{Radioactivity dose}} \times 100\%$$

Values for the exchangeable body water content as well as the half-life of tritiated water in some mammalian species that can be applied to the preceding equation were shown earlier in Table 17.13. If the molecule is likely to or is known to fragment into two major portions, it may be desirable to monitor both fragments by differential labeling (^3H and ^{14}C).

The chemical and radiochemical purity of the labeled compound must be ascertained prior to use. In practice a

value of 95% or greater is usually acceptable. The desired specific activity of the administered radioactive compound depends on the dose to be used as well as the species studied. Doses of ^{14}C on the order of $5 \mu\text{Ci kg}^{-1}$ for the dog and $20 \mu\text{Ci kg}^{-1}$ for the rat have been found adequate in most studies, while doses of ^3H are usually two to three times higher owing to lower counting efficiency of this isotope.

Liquid scintillation counting is the most popular technique for the detection and measurement of radioactivity. In order to count a liquid specimen such as plasma, urine, or digested blood or tissues directly in a liquid scintillation spectrometer, an aliquot of the specimen is first mixed with a liquid scintillant. Aliquots of blood, feces, or tissue homogenates are air-dried on ash-free filter papers and combusted in a sample oxidizer provided with an appropriate absorption medium and a liquid scintillant prior to counting. The liquid scintillant plays the role of an energy transducer, converting energy from nuclear decay into light. The light generates electrical signal pulses which are analyzed according to their timing and amplitude and are subsequently recorded as a count rate, for example, counts per minute (cpm). Based on the counting

TABLE 17.12 Properties of Primary Radioisotopes Employed in PKs

| Property | ^3H | ^{51}Cr | ^{14}C | ^{125}I |
|-------------------------------------|--------------|------------------|-----------------|------------------|
| Half-life (year) | 12.3 | 27.8 (days) | 5730 | 13 (days) |
| Maximum beta energy (MeV) | 0.0186 | 0.752 | 0.156 | 2.150 |
| Average beta energy (MeV) | 0.006 | 0.049 | | |
| Range in air (mm) | 6 | 300 | | |
| Range in unit density material (mm) | 0.0052 | 0.29 | | |

TABLE 17.13 Selected Factors that May Affect Chemical Distribution to Various Tissues

| |
|--|
| Factors relating to the chemical and its administration |
| Degree of binding of chemical to plasma proteins (i.e., agent affinity for proteins) and tissues |
| Chelation to calcium, which is deposited in growing bones and teeth (e.g., tetracyclines in young children) |
| Whether the chemical distributes evenly throughout the body (one-compartment model) or differentially between different compartments (two- or more compartment model) |
| Ability of chemical to cross the blood–brain barrier |
| Diffusion of chemical into the tissues or organs and degree of binding to receptors that are and are not responsible for the drug's beneficial effects |
| Quantity of chemical given |
| Route of administration/exposure |
| Partition coefficients (nonpolar chemicals are distributed more readily to fat tissues than polar chemicals) |
| Interactions with other chemicals that may occupy receptors and prevent the drug from attaching to the receptor, inhibit active transport, or otherwise interfere with a drug's activity |
| Molecular weight of the chemical |
| Factors relating to the test subject |
| Body size |
| Fat content (e.g., obesity affects the distribution of drugs that are highly soluble in fats) |
| Permeability of membranes |
| Active transport for chemicals carried across cell membranes by active processes |
| Amount of proteins in blood, especially albumin |
| Pathology or altered homeostasis that affects any of the other factors (e.g., cardiac failure and renal failure) |
| The presence of competitive binding substances (e.g., specific receptor sites in tissues bind drugs) |
| pH of blood and body tissues |
| pH of urine ^a |
| Blood flow to various tissues or organs (e.g., well-perfused organs usually tend to accumulate more chemicals than less well-perfused organs) |

^a The pH of urine is usually more important than the pH of blood.

efficiency of the radionuclide used, the count rate is then converted to the rate of disintegration, for example, disintegrations per minute (dpm), which is a representation of the amount of radioactivity present in the sample.

17.5.1.3 Immunoassay Methods Radioimmunoassay (RIA) allows measurement of biologically active materials which are not detectable by traditional cold chemistry techniques. RIAs can be used to measure molecules that cannot be radiolabeled to detectable levels *in vivo*. They also are used for molecules unable to fix complement when bound to antibodies, or they can be used to identify cross-reacting antigens that compete and bind with the antibody. These are discussed further in Section 17.11.

17.6 SAMPLING METHODS AND INTERVALS

17.6.1 Blood

Since blood (plasma and serum) is the most easily accessible body compartment, the blood concentration profile is most commonly used to describe the time course of drug disposition in the animal. With the development of sensitive analytical methods that require small volumes (100–200 μ L) of blood, ADME data from individual rats can be obtained by serial sample collection. Numerous cannulation techniques have been utilized to facilitate repeated blood collection, but the animal preparation procedures are elaborate and tedious and are incompatible with prolonged sampling periods in studies involving a large number of animals. In contrast, noncannulation methods such as collection from the tail vein, orbital sinus, or jugular vein are most practical. Significant volumes of blood can be obtained from the intact rat by cardiac puncture, although this method can cause shock to the animal system and subsequent death; therefore, cardiac puncture is usually only a terminal collection method.

Blood collection from the tail vein is a simple and rapid, nonsurgical method which does not require anesthesia. A relatively large number of serial samples can be obtained within a short period of time. However, this method is limited to relatively small sample volumes (~250 μ L per sample). Although larger volumes can be obtained by placing the rat in a warming chamber, this procedure could significantly influence the disposition of the test compound and therefore is not recommended for routine studies. Blood collected from the cut tail has been shown to provide valid concentration data for numerous compounds.

The rat is placed in a suitable restrainer with the tail hanging freely. The tail is immersed in a beaker of warm water (37–40°C) for 1–2 min to increase the blood flow and thereby visibility of the vein. Using surgical scissors or a scalpel, the tail is completely transected approximately 5 mm above the tip. The tail is then gently “milked” by sliding the fingers down the tail from its base. It should be noted that excessive “milking” could cause damage to the blood capillaries or increase the white cell count in the blood. A heparinized micropipette of desired capacity (25–250 μ L) is held at a 30–45° downward angle in contact with the cut end of the tail. This allows blood to fill the micropipette by capillary action. Application of gentle pressure with a gauze pad for approximately 15 s is sufficient to stop bleeding. A sufficient number of serial blood samples may be obtained to adequately describe the blood level profile of a compound.

If plasma is required, the blood may be centrifuged after sealing one end of the filled micropipette and placing it in a padded centrifuge tube. The volume of plasma is determined by measuring the length of plasma as a fraction of the length of the micropipette, multiplied by the total capacity of the pipette. The tube is then broken at the plasma/red blood cell

(RBC) interface and the sample is expelled using a small bulb. If serum is needed, the blood should be collected without using anticoagulants in the sampling tube.

Serial blood samples can also be collected from the orbital sinus, permitting rapid collection of larger (1–3 mL) samples. However, since 2008, there have been initiatives to collect smaller samples from fewer animals in an attempt to reduce animal usage.

Three new approaches for blood sampling for PK/TK analysis have, however, come into consideration for use in rodents to facilitate use of fewer animals.

The first of these is microsampling, by means of a microdialysis or capillary sampling needle. There is minimal trauma to the rodents, and in taking minimal amounts of blood each time, it does not compromise the rat or mouse (or hamster). Accordingly, “main study” animals can be used and satellite groups of animals are not required—significantly reducing the number of animals required for a study (particularly for longer duration studies where sampling at multiple time points is required).

A variation on this is the use of dried blood spot sampling (DBSS). With liquid plasma samples, a significant amount of each sample collected is lost in adherence to the walls of a series of glass or plastic collection and handling vessels; DBSS collects directly onto a chemically inert medium from a capillary sampling tube or the Cutex® collecting device. As long as the molecules to be analyzed are chemically stable under these conditions, again smaller samples (and fewer animals) are required.

The third approach is sparse sampling, where only a small number (usually 3) of animals are sampled at each time point, with sampling being rotated between time points (such as the traditional 0.5, 1, 2, 4, 8, and 12 h), so that levels of blood collected from any one animal are sufficient to avoid notable traumatization. Again, either no satellite group of animals or only a smaller such group is required.

17.6.2 Excreta

Excretion samples commonly collected from the rat include urine, feces, bile, and expired air. By using properly designed cages and techniques, the samples can be completely collected so that the mass balance is readily determined. These samples also serve to elucidate the biotransformation characteristics of the compound.

These samples can be easily collected through the use of suitable metabolism cages. Since rodents are coprophagic, the cage must be designed to prevent the animal from ingesting the feces as it is passed. Other main features of the cage should include the ability to effectively separate urine from feces with minimal cross-contamination, a feed and water system that prevents spillage and subsequent contamination of collected samples, and collection containers that can be easily removed without disturbing the animal. Also, the cage

should be designed so that it can be easily disassembled for cleaning or autoclaving.

Following dose administration, rats are placed in individual cages. The urine and feces that collect in containers are removed at predetermined intervals. The volume or weight of urine and the weight of feces are measured. After the final collection, the cage is rinsed, normally with ethanol or water, to assure complete recovery of excreta and radioactivity. If the rats are also used for serial blood sampling, it is important that bleeding be performed inside the cage to avoid possible loss of urine or feces.

17.6.3 Bile

The bile is the pathway through which an absorbed compound is excreted in the feces. In order to collect this sample, surgical manipulation of the animal is necessary.

17.6.4 Expired Air

For ^{14}C -labeled chemicals, the tracer carbon may be incorporated in vivo into carbon dioxide, a possible metabolic product. Therefore, when the position of the radiolabel indicates the potential for biological instability, a pilot study to collect expired air and monitor its radioactivity content should be conducted prior to initiating a full-scale study. Expired air studies should also be performed in situations where the radiolabel has been postulated to be stable but analyses of urine and feces from the TK study fail to yield complete recovery (mass balance) of the dose.

Following drug administration, the rat is placed in a special metabolism cage. Using a vacuum pump, a constant flow of room air ($\sim 500 \text{ mL min}^{-1}$) is drawn through a drying column containing anhydrous calcium sulfate impregnated with a moisture indicator (cobalt chloride) and is passed into a second column containing Ascarite II®, where it is rendered carbon dioxide-free. The air is then drawn in through the top of the metabolism cage. Exhaled breath exiting the metabolism cage is passed through a carbon dioxide adsorption tower, where the expired $^{14}\text{CO}_2$ is trapped in a solution such as a mixture of 2-ethoxyethanol and 2-aminoethanol (2:1). The trapping solution is collected, replaced with fresh solution, and assayed at designated times postdose so that the total amount of radioactivity expired as labeled carbon dioxide can be determined.

17.6.5 Milk

The study of passage of a xenobiotic into milk serves to assess the potential risk to breastfed infants in the absence of human data. The passage into milk can be estimated as the milk–plasma ratio of drug concentrations at each sampling time or that of the area under the curve (AUC) values. Approximately 30 rats in their first lactation are used. The litter size is adjusted

TABLE 17.14 Approximate Volumes of Pertinent Biological Fluids in Adult Laboratory Animals

| Fluid | Rat | Mouse | Dog | Rabbit | Monkey |
|--|-----|-------|-----|--------|--------|
| Blood (mL kg ⁻¹) | 75 | 75 | 70 | 60 | 75 |
| Plasma (mL kg ⁻¹) | 40 | 45 | 40 | 30 | 45 |
| Urine (mL kg ⁻¹ day ⁻¹) | 60 | 50 | 30 | 60 | 75 |
| Bile (mL kg ⁻¹ day ⁻¹) | 90 | 100 | 12 | 120 | 25 |

to about 10 within 1–2 days following parturition. The test compound is administered to the mothers 8–10 days after parturition. The rats are then divided into groups for milk and blood collection at designated times postdose. All sucklings are removed from the mother rats several hours before milking. Oxytocin, 1 IU per rat, is given intramuscularly 10–15 min before each collection of milk to stimulate milk ejection. The usual yield of milk is about 1 mL from each rat. Blood is obtained immediately after milking. In order to minimize the number of animals used, the sucklings can be returned to the mother rat which can then be milked again 8–12 h later.

In all the fluid sampling techniques mentioned earlier, the limitations of availability should be kept in mind. Table 17.14 presents a summary of such availability for the principal model species.

For topical exposures, determining absorption (into the skin and into the systemic circulation) requires a different set of techniques. For determining how much material is left, skin washing is required. There are two components to skin washing in the recovery of chemicals. The first component is the physical rubbing and removal from the skin surface. The second component is the surfactant action of soap and water. However, the addition of soap effects the partitioning. Some compounds may require multiple successive washing with soap and water applications for removal from skin.

Skin tape stripping can be used to determine the concentration of drug in the stratum corneum at the end of a short application period (30 min) and by linear extrapolation predicts the percutaneous absorption of that chemical for longer application periods. The chemical is applied to skin of animals or humans, and after a 30-min skin contact application time, the stratum corneum is blotted and then removed by successive tape applications. The tape strippings are assayed for chemical content. There is a linear relationship between this stratum corneum reservoir content and percutaneous absorption. The major advantages of this method are (i) the elimination of urinary and fecal excretion to determine absorption and (ii) the applicability to nonradiolabeled determination of percutaneous absorption, because the skin strippings contain adequate chemical concentrations for nonlabeled assay methodology.

Finally, a complete determination of the distribution and potential departing of a chemical and its metabolites requires some form of measurement or sampling of tissues/organs. Autoradiography provides a nonquantitative means of doing such, but quantitation requires actual collection and sampling

TABLE 17.15 Blood Samples Required so that Certain TK Parameters Can Be Obtained and Calculated

| Parameter | Blood Sample Required | Information Obtained |
|-------------------------|--|------------------------|
| C_{\min} (C_{24}) | 24-h | Accumulation |
| CT | T -h | Proof of absorption |
| C_{\max} (C peak) | Several ^a | Rate of absorption |
| T_{\max} (T peak) | Several ^a | Rate of absorption |
| AUC | Several ^a | Extent of absorption |
| $t_{1/2}$ | Several ^a | Various |
| Accumulation ratio | Several after first and repeated doses | Extent of accumulation |

^a Several samples to define concentration-versus-time profile.

of tissues. Table 17.4 provides guidance as to the relative percentage of total body mass that the organs constitute in the common model species.

17.6.5.1 Sampling Interval To be able to perform valid TK analysis, it is necessary not only to properly collect samples of appropriate biological fluids but also to collect a sufficient number of samples at the current intervals. Both of these variables are determined by the nature of the answers sought. Useful parameters in TK studies are C_{\max} , which is the peak plasma test compound concentration; T_{\max} , which is the time at which the peak plasma test compound concentration occurs; C_{\min} , which is the plasma test compound concentration immediately before the next dose is administered; AUC, which is the area under the plasma test compound concentration-versus-time curve during a dosage interval; and $t_{1/2}$, which is the half-life for the decline of test compound concentrations in plasma. The samples required to obtain these parameters are shown in Table 17.15. C_{\min} requires one blood sample immediately before a dose is given and provides information on accumulation. If there is no accumulation in plasma, the test compound may not be detected in this sample.

Several C_{\min} samples are required at intervals during the toxicity study to check whether accumulation is occurring. CT is a blood sample taken at a chosen time after dosing and provides proof of absorption as required by the GLP regulations, but little else. C_{\max} requires several blood samples to be taken for its accurate definition, as does T_{\max} ; these two parameters provide information on rate of absorption. AUC also requires several blood samples to be taken so that it can

be calculated: it provides information on extent of absorption. $t_{1/2}$, the half-life, requires several samples to be taken during the terminal elimination phase of the test compound concentration-versus-time curve: this parameter provides information on various aspects such as any change in the kinetics of the test compound during repeated doses or at different dose levels. Depending on the other parameters obtained, the accumulation ratio can be calculated from C_{\min} , C_{\max} , and/or AUC when these are available after the first dose and after several doses to steady state.

Operational and metabolic considerations generally make urine sampling and assay of limited value for TK purposes.

17.7 STUDY TYPES

Metabolic and PK data from a rodent species and a nonrodent species (usually the dog) used for repeat dose safety assessments (14 days, 28 days, 90 days, or 6 months) are recommended. If a dose dependency is observed in metabolic and PK or toxicity studies with one species, the same range of doses should be used in metabolic and PK studies with other species. If human metabolism and PK data also are available, this information should be used to help select test species for the full range of toxicity tests and may help to justify using data from a particular species as a human surrogate in safety assessment and risk assessment.

Metabolism and PK studies have greater relevance when conducted in both sexes of young adult animals of the same species and strain used for other toxicity tests with the test substance. The number of animals used in metabolism and PK studies should be sufficient to reliably estimate population variability. This usually means a separate (but parallel) set of groups of animals in rodent studies. A single set of intravenous (IV) and oral dosing results from adult animals, when combined with some *in vitro* kinetic results, may provide an adequate data set for the design and interpretation of short-term, subchronic, and chronic toxicity studies.

Studies in multiple species may clarify what appear to be contradictory findings in toxicity studies (i.e., equal milligram per kilogram body weight doses having less effect in one species than in another). If disposition and metabolite profiles are found to be similar, then differences in responses among species could more reliably be attributed to factors other than differences in metabolism. Studies of the PKs and metabolism of a substance in neonatal and adolescent animals provide information about any changes in metabolism associated with tissue differentiation and development. Animals with fetuses of known gestational age should be used for determining the disposition of the test substance in the fetus. Dosage is by (to the maximum extent possible) the intended clinical route.

An acute IV study can provide accurate rates of metabolism—without interference from intestinal flora—plus

rates of renal and biliary elimination, if urine and bile are collected. This route also avoids the variability in delivered dose associated with oral absorption and ensures that the maximum amount of radiolabel is excreted in the urine or bile for purposes of detection. Once IV data and parameters are available, they can be used with plasma concentrations from limited oral studies to compute intestinal absorption via the ratio of areas under the (plasma and/or urine) curves or via simulations of absorption with GI absorption models.

In single-dose PK studies of oral absorption, the primary concerns are with the rates and extent of absorption and peak plasma or target tissue concentrations of the test substance. If the test vehicle affects gastric emptying, it may be necessary to use both fasted and nonfasted animals for PK studies.

Blood (RBCs, plasma, and serum), urine, and feces are the most commonly collected samples. In addition, a few representative organ and tissue samples should be taken, such as liver, kidney, fat, and suspected target organs. Sampling times should depend on the substance being tested and the route of administration. In general, an equal number of blood samples should be taken in each phase of the concentration-versus-time curve. IV studies usually require much shorter, and more frequent, sampling than is required for oral dosing. Time spacing of samples will depend on the rates of uptake and elimination. In a typical IV study, blood and tissue samples are taken in a “powers of 2” series, that is, samples at 2, 4, 8, 16, and 30 (32) min, 1, 2, 4, 8, and 16 h. Similar coverage could be obtained with only 7 time points by using a “powers of 3” series: 3, 9, and 30 (27) min, 1, 3, 9, and 24 (27) h. Oral dosing studies usually extend to at least 72 h. Such a sampling scheme over 3 days would provide better data coverage for a more thorough evaluation of absorption, elimination, enterohepatic recirculation, and excretion processes.

The number of animals used in metabolism and PK studies should be large enough to reliably estimate population variability. In the case of rats and mice, tissue and/or blood sample size is usually the limiting factor: analysis of the substance may require 1 mL or more of blood, but this is not usually practical or safe to obtain multiple blood samples of this size from one animal and may confound or complicate PK or safety evaluations. As a consequence, a larger number of animals are required (3–4 per time point, 7–9 time points) when small rodents are used. Such an approach has the advantage of allowing limited sampling of critical tissues (e.g., liver, fat) at each time point, an option which is usually unavailable with large animals. The use of humans and large animals generally permits collection of multiple (serial) blood samples. For outcrossing populations like humans and large animals, individual differences in the rates of biotransformation are likely to be greater than those of inbred rodent populations; under these circumstances, more samples/sexes/groups may be needed to reliably estimate variability.

Individual metabolism cages are recommended for collecting urine and feces in oral dosing studies. Excreta should be collected for at least five elimination half-lives of the test substance. When urine concentrations will be used to determine elimination rates, sampling times should be less than one elimination half-life (taken directly from the bladder in IV studies); otherwise, samples should be taken at equal time intervals.

The results of the preliminary biotransformation/kinetic study, together with the current regulatory metabolism studies and the 28- and 90-day studies, should allow the selection of a relatively small number of appropriate tissues and/or fluids for monitoring purposes. Satellite groups of animals will provide the material for analysis. Methods must be developed to analyze the nonradioactive test chemical. Obviously it is important to monitor blood. It is accessible and convenient, and in certain circumstances, sequential sampling from the same animal may be important. The most useful aspect of blood is that the results can be compared with those obtained in man (see succeeding text). It is important, however, not to be constrained by this aspect. The most relevant tissues and body fluids should also be analyzed. These are target organs (if known) and indicator organs, tissues, or fluids, that is, those in which the concentration of pesticide or metabolite is a measure of that in the whole animal. In cases where distribution varies with dose (if shown in the preliminary study), a larger number of organs/tissues would be chosen for monitoring.

Whether the parent drug or metabolite (or both) is chosen for analysis depends on the preliminary study. In principle, analysis for the parent compound should always be carried out; however, there are situations (e.g., rapid metabolism) when this is quite futile and a major retained metabolite should be used. Covalently bound metabolites are addressed in the succeeding text.

Four occasions may be adequate for monitoring:

1. One month (equilibrium between intake of chemical and elimination of metabolites should be established; the time relates to the 28-day preliminary study)
2. Three months (confirmation of results at one month; relates to the 90-day study)
3. One year (coincides with the interim kill)
4. Two years (effects of age; coincides with termination of study)

Consideration should be given to the analysis of moribund animals.

17.7.1 Whole-Body Autoradiography

Autoradiography is the production of an image in a photographic emulsion by the emission from a radioactive element. The term autoradiography is preferred to radioautography.

Prefixes are added to words to further classify the concept. Therefore, the process is "auto-" radiography for a "self-" radiograph and not a "radio-" autograph or one's transmitted signature (Waddell, 1972).

Whole-body autoradiography (WBA) has been used with increasing frequency as a means of identifying tissues which concentrate test substances. This technique allows a small number of animals (5–10) to be used for screening purposes with a minimal investment in manual labor. FDA encourages the use of WBA with IV dosing as a means of screening and selecting tissues of greatest relevance for later oral dosing studies. Animals used for WBA should generally be sacrificed during primary consideration in selecting specific tissues.

The most comprehensive technique currently available for the initial survey of the distribution of a drug is that of WBA. The species of animals used include mice, rats, hamsters, monkeys, pigs, dogs, and ferrets. The most widely used animal has been the mouse, which has the advantages of requiring few isotopes and being easier to section.

The animals are anesthetized and then frozen by immersion at various times after administration of the labeled compound in hexane or acetone cooled with dry ice. Since the freezing in the interior of the animals occurs slowly, large ice crystals form within these tissues; hence, subcellular localization of compounds is not possible.

The selection of times for freezing an animal after injection of a drug must be based on the information available on the rate of elimination of the compound from the animal by metabolism and excretion. In general, a geometric increase in time intervals is most useful. In order to have time intervals for comparison, we routinely have employed freezing times which are approximately multiples of three, namely, 2 min, 6.5 min, 20 min, 1 h, 3 h, 9 h, and 24 h. In certain cases, rapid elimination of the drug by the kidneys must be circumvented by ligation of the renal pedicles to avoid apparent localization from failure of the agent to reach equilibrium.

The animal is frozen into a block of carboxymethyl cellulose ice on the microtome stage. Although the Jung type K microtome has been used, the Leitz model 1300 sledge microtome is more suitable, for its smaller size allows it to be mounted in an ordinary commercial freezer instead of a walk-in freeze room. The microtome stage must be designed for mounting in the vice of the front end of the stage.

Sections that are from 5 μ m to approximately 80 μ m thick are taken onto #800 Scotch tape (Minnesota Mining & Manufacturing Company). Before removal from the freezer, the sections must be allowed to dry thoroughly so that no ice remains, which melts and allows movement of the isotope. After drying, if covered to prevent condensation of moisture on the sections, the sections may be transferred from the freezer to room temperature.

X-ray films which produce the most satisfactory autoradiograms are Kodak industrial type AA and Gevaert

Structurix D-7. Both are fine-grained films which have been demonstrated not to produce chemical artifacts. Approximately six times faster, Kodak No-Screen and Kodirex may be used for rapid screening and timing of autoradiograms. However, they occasionally produce artifacts and should not be relied on for interpretation. Some investigators have used photographic emulsions such as Ilford G-5 (10 μ m thick, pre-applied to glass plates). The increased cost and likelihood of breakage, however, hardly justify the small improvement in resolution for whole-body sections.

Exposure of the photographic emulsion to the radioactivity of the tissue section should be at freezer temperatures to prevent autolysis of the tissue. After exposure of the x-ray film, sections with isotopes which have a long half-life may be placed against fresh x-ray film for additional sets of autoradiograms with either a longer or shorter exposure time. This procedure is useful for revealing relative concentrations of radioactivity for areas that have either very high or very low concentrations after the first exposure. When no further autoradiograms are needed, the section can be stained with histological dyes to verify localizations of radioactivity.

Compounds that fluoresce under ultraviolet light can be visualized in the tissue sections and their locations recorded with color film. Whole-body tissue sections can be used for histochemical localizations for comparison with the autoradiograms. Furthermore, the areas can be removed and extracted, and the extract chromatographed to identify the chemical nature of the radioactivity revealed by the autoradiogram.

Although the whole-body technique will allow localization of an increased concentration of an isotope in a tissue or occasionally a cell type, other techniques must be used for single cells and subcellular localization. A nuclear tract plate is prepared by dipping the plate in a 12% solution of glycerin in absolute ethyl alcohol and allowing it to drain for 10 min in a vertical position before approximating the section to tape. After the emulsion is exposed, soaking in xylene removes the tape but leaves the section attached to the nuclear tract plate. The Ilford G-5 nuclear tract plates with 10 μ m emulsions are most satisfactory. The increased resolution gained by the finer-grained Ilford K and L emulsions is warranted only for tissues that are well preserved and relatively free of ice crystal artifacts. Kodak NTB emulsions seem to produce more pressure artifacts than the Ilford plates.

Comparison of various techniques of autoradiography for diffusible compounds clearly demonstrates that no solutions can be used in processing the tissue. These investigators have dried thin sections of liver and uterus at temperatures below -60°C. These freeze-dried sections were dry mounted on microscope slides which had been precoated with either Kodak NTB-3 or NTB-10 emulsion. Other techniques which thawed the frozen section, embedded the tissue in paraffin, or dipped the section in liquid emulsion were demonstrated to translocate diffusible compounds. Many other similar attempts have been and are currently being

made to localize diffusible compounds by autoradiography at the electron microscope level.

17.7.2 Mass Balance Studies

Another method for evaluating the distribution of drug throughout the body while also assessing clearance rates from the body is the mass balance study. In this study format, the drug (in a radiolabeled form) is administered by the desired route. Excreta (urine and feces, as well as occasionally expired air) is collected over a suitable interval of time after dosing (usually 24–48 h in rodents and up to 72 h in large animals) by placing test animals in special collection chambers or cages.

At the end of this interval, blood is collected, the animal is euthanized, key tissues are collected and weighed, and then the amount of label in each is determined. The rest of the body is separately weighed and processed, and residual compound (measured as radioactivity) is determined. It is usually the case that accounting for 85%+ of administered drug is considered acceptable, but 90–95% is desirable. Rodents should yield recovery values of 95%+, while recovery in dogs and monkeys may be more often 85%+.

17.7.2.1 In Vitro Studies *In vitro* measurements employing enzymes, subcellular organelles, isolated cells, and perfused organs may be used to augment the dose–response information available from less extensive metabolic and PK studies. Because *in vitro* systems are less complex than whole animals, elucidation of a test compound's metabolic pathways and the pathways' kinetic characteristics may be facilitated. Such systems can be used to measure binding, adduct and conjugate formation, transport across cell membranes, enzyme activity, enzyme substrate specificity, and other singular objectives. Biochemical measurements that can be made using *in vitro* systems include intrinsic clearances of enzymes in an organ or tissue, kinetic constants for an enzyme, binding constants, and the affinity of the test compound and its metabolites for the target macromolecules. The activity of a hepatic drug-metabolizing enzyme *in vivo* may be approximated by kinetic constants that are calculated from *in vitro* studies; when a first-order approximation is used, the ratio of V_{\max} to K_m is equal to the intrinsic clearance of the drug. *In vitro* measurements made using readily accessible tissues and body fluids from animals and man may also be useful in elucidating mechanisms of toxicity.

17.8 ANALYSIS OF DATA

Data from all metabolism and PK studies should be analyzed with the same PK model and results should be expressed in the same units. Concentration units are acceptable if the

organ or sample size is reported, but percent of dose/organ is usually a more meaningful unit. In general, all samples should be analyzed for metabolites that cumulatively represent more than 1% of the dose.

A variety of rate constants and other parameters can be obtained from IV and oral dosing data sets, provided that good coverage of the distribution, elimination, and absorption (oral dose) phase is achieved. Typical parameters calculated to characterize the disposition of a test substance are half-lives of elimination and absorption; area under the concentration-versus-time curve (AUC) for blood (or plasma); total body (renal and metabolic) clearance (CL); volume of distribution (V_d); bioavailability (F); and mean residence time (MRT) and mean absorption time (MAT). Some of these parameters, such as half-lives and elimination rates, are easily computed from one another; the half-life is more easily visualized than the rate constant.

Computation of oral absorption (k_a) and elimination (E) rates is often complicated by the "flip-flop" of the absorption and elimination phases when they differ by less than a factor of 3. Because of these analysis problems, computation of absorption and elimination rates should not be attempted on the basis of oral dosing results alone, and therefore, it is preferable to have both oral and IV PK data in hand when determining oral absorption and elimination rates.

Blood-tissue uptake rates (k_{ij}) can often be approximated from data at early ($t < 10$ min) time points in IV studies, provided that the blood has been washed from the organ (e.g., liver) or the contribution from blood to the tissue residue is subtracted (fat). High accuracy is not usually required since these parameters can be optimized to fit the data when they are used in more complex models. Tissue-blood recycling rates (k_{ji}) and residence times can be computed from partition coefficients if estimates of uptake rates are available.

Tissue-blood partition coefficients (R_{ji}) should be determined when steady state has been achieved. Estimates based on samples obtained during the elimination phase following a single dose of the test substance may lead to underestimates of this ratio in both eliminating and non-eliminating tissues unless its half-life is very long. Correction of these values for elimination has been described by several authors.

It may be important to determine the degree of plasma protein and RBC binding of the test substance; calculation of blood clearance rates using plasma or serum concentrations of the substance that have not been adjusted for the degree of binding may under- or overestimate the true rate of clearance of the test substance from the blood. This is usually done through experiments *in vitro*.

Two classical methods used in the analysis of PK data are the fitting of sums of exponential functions (two- and three-compartment mammillary models) to plasma and/or tissue data and, less frequently, the fitting of arbitrary polynomial functions to the data (noncompartmental analysis).

Noncompartmental analysis is limited in that it is not descriptive or predictive; concentrations must be interpolated from data. The appeal of noncompartmental analysis is that the shape of the blood concentration-versus-time curve is not assumed to be represented by an exponential function and, therefore, estimates of metabolic and PK parameters are not biased by this assumption. In order to minimize errors in parameter estimates that are introduced by interpolation, a large number of data points that adequately define the concentration-versus-time curve are needed (Gabrielsson and Weiner, 1997).

Analysis of data using simple mammillary compartmental models allows the estimation of all of the basic parameters mentioned earlier, if data for individual tissues are analyzed with one- or two-compartment models and if combined with results from two- to three-compartment analyses of blood data. "Curve stripping" analysis can be applied to such simple models through the use of common spreadsheet programs (i.e., Lotus 1-2-3 or Excel), as long as a linear regression function is provided in the program. Optimization of the coefficients and exponents estimated may require the use of more sophisticated software: a number of scientific data analysis packages such as RS/1 and SigmaPlot have the necessary capabilities. Specialized programs such as NONLIN, CONSAM, or SIMUSOLV will be needed when more complex models must be analyzed. Coefficients and exponents from mammillary models can be used to calculate other parameters; however, they should not be taken too literally, since mammillary models assume that all inputs are to a central pool (blood), which equilibrates without limitation into other compartments. This approach does not include details such as blood flow limitations, anatomical volumes, or other physiological limits in the animal.

Physiologically based pharmacokinetic (PBPK) models were developed to overcome the limitations of simple mammalian models. Physiologically based models describe the disposition of test substances via compartmental models which incorporate anatomical, biochemical, and physiological features of specific tissues in the whole animal (Connally and Anderson, 1991; Vinegar and Jepson, 1996). The types of information added include organ-specific blood flows, volumes, growth models, and metabolism rates. Metabolic parameters often are obtained from *in vitro* studies (i.e., enzyme reaction rates in cultured hepatocytes, plasma protein binding, etc.), while other parameters are becoming available as standard parameters in the literature. Parameters from mammillary models can be used to compute the value of parameters used in PBPK models, using tissue-specific blood flows, anatomical volumes, and other information (literature values). Estimation of parameters for a simple mammillary model is often the first data reduction step in creating a physiological model.

Because PBPK models are based on physiological and anatomical measurements and all mammals are inherently

similar, they provide a rational basis for relating data obtained from animals to humans. Estimates of predicted disposition patterns for test substances in humans may be obtained by adjusting biochemical parameters in models validated for animals; adjustments are based on experimental results of animal and human *in vitro* tests and substitution of appropriate human tissue sizes and blood flows. Development of these models requires special software capable of simultaneously solving multiple (often very complex) differential equations, some of which were mentioned earlier. Several detailed descriptions of data analysis have been reported.

17.8.1 Use of Data from Metabolism and Pharmacokinetic Studies

Information from metabolism and PK studies can be used in the design and analysis of data from other toxicity studies. Some examples are described in the succeeding text.

17.8.1.1 Design of Toxicity Studies The concentration-versus-time curve and peak and steady-state concentrations of the test substance in blood or plasma provide information on the distribution and persistence of the substance in the animal which may suggest essential elements in the design of the toxicity studies. For example, when metabolic and PK studies indicate that the test compound accumulates in the bone marrow, long-term toxicity tests should include evaluation of the test compound's effect on hematopoietic function and morphology. If a test compound is found to accumulate in milk, an investigator may need to plan to perform reproductive toxicity studies with *in utero* exposure and a nursing phase (cross-fostering study). In addition, information from metabolic and PK studies can be used to predict the amount of test compound that enters biological compartments (tissues, organs, etc.) that may not suffer a toxic insult but may serve as depots for indirect or secondary exposure.

17.9 PHYSIOLOGICALLY BASED PHARMACOKINETIC (PBPK) MODELING

Pharmacokinetic parameters are descriptive in nature. They quantitatively describe the manner in which a test material is absorbed and excreted, such that a specific blood or tissue level is achieved or maintained. In the past, experiments had to be done by every route of administration to gather the data appropriate for describing the PK behavior of a chemical administered by different routes. The development of more sophisticated and readily accessible computers has led to the development of a different approach, that of PK modeling. In this computerized model, different compartments are represented as shown in boxes and the movement of the material in and out of the compartments is defined by the rate constants. These can be determined either *in vivo* or *in vitro*.

Other physiological parameters are brought into play as well, such as octanol/water partition coefficient, blood flow through an organ, respiration rate (for the inhalation route of exposure), rate of microsomal metabolism, and so on.

Pharmacokinetic modeling is the process of developing mathematical explanations of absorption, distribution, metabolism, and excretion of chemicals in organisms. Two commonly used types of compartmental PK models are (i) data based and (ii) physiologically based. The data-based PK models correspond to mathematical descriptions of the temporal change in the blood/tissue level of a xenobiotic in the animal species of interest. This procedure considers the organism as a single homogeneous compartment or as a multicompartamental system with elimination occurring in specific compartments of the model. The number, behavior, and volume of these hypothetical compartments are estimated by the type of equation chosen to describe the data, and not necessarily by the physiological characteristics of the model species in which the blood/tissue concentration data were acquired. It can provide a powerful tool for both the discovery of new drugs and the optimization of their development (Lavé et al., 2007).

Whereas these data-based PK models can be used for interpolation, they should not be used for extrapolation outside the range of doses, dose routes, and species used in the study on which they were based. In order to use the data-based models to describe the PK behavior of a chemical administered at various doses by different routes, extensive animal experimentation would be required to generate similar blood time course data under respective conditions. Even within the same species of animal, the time-dependent nature of critical biological determinants of the disposition (e.g., tissue glutathione depletion and resynthesis) cannot easily be included or evaluated with the data-based PK modeling approach. Further, due to the lack of actual anatomical, physiological, and biochemical realism, these data-based compartmental models cannot easily be used in interspecies extrapolation, particularly to predict PK behavior of chemicals in humans. These various extrapolations, which are essential for the conduct of dose-response assessment of chemicals, can be performed more confidently with a PBPK modeling approach. This chapter presents the principles and methods of PBPK modeling as applied to the study of toxicologically important chemicals.

PBPK modeling is the development of mathematical descriptions of the uptake and disposition of chemicals based on quantitative interrelationships among the critical biological determinants of these processes. These determinants include partition coefficients, rates of biochemical reactions, and physiological characteristics of the animal species. The biological and mechanistic basis of the PBPK models enables them to be used, with limited animal experimentation, for extrapolation of the kinetic behavior of chemicals from high dose to low dose, from one exposure route to another, and from test animal species to people.

The development of PBPK models is performed in four interconnected steps: model representation, model parameterization, model stimulation, and model validation. Model representation involves the development of conceptual, functional, and computational descriptions of the relevant compartments of the animal as well as the exposure and metabolic pathways of the chemical. Model parameterization involves obtaining independent measures of the mechanistic determinants, such as physiological, physicochemical, and biochemical parameters, which are included in one or more of the PBPK model equations. Model simulation involves the prediction of the uptake and disposition of a chemical for defined exposure scenarios using a numerical integration algorithm, simulation software, and computer. Finally, the model validation step involves the comparison of the *a priori* predictions of the PBPK model with experimental data to refute, validate, or refine the model description and the characterization of the sensitivity of tissue dose to changes in model parameter values. PBPK models after appropriate testing and validation can be used to conduct extrapolations of the PK behavior of chemicals from one exposure route/scenario to another, from high dose to low dose, and from one species to another.

The PBPK model development for a chemical is preceded by the definition of the problem, which in toxicology may often be related to the apparent complex nature of toxicity. Examples of such apparent complex toxic responses include nonlinearity in dose–response, sex/species differences in tissue response, differential response of tissues to chemical exposure, qualitatively and/or quantitatively different responses for the same cumulative dose administered by different routes/scenarios, and so on. In these instances, PBPK modeling studies can be utilized to evaluate the PK basis of the apparent complex nature of toxicity induced by the chemical. One of the values of PBPK modeling, in fact, is that accurate description of target tissue dose often resolves behavior that appears complex at the administered dose level.

The principal application of PBPK models is in the prediction of the target tissue dose of the toxic parent chemical or its reactive metabolite. Use of the target tissue dose of the

toxic moiety of a chemical in risk assessment calculations provides a better basis of relating to the observed toxic effects than the external or exposure concentration of the parent chemical. Because PBPK models facilitate the prediction of target tissue dose for various exposure scenarios, routes, doses, and species, they can help reduce the uncertainty associated with the conventional extrapolation approaches. Direct application of modeling includes:

- High-dose–low-dose extrapolation
- Route–route extrapolation
- Exposure scenario extrapolation
- Interspecies extrapolation

17.10 POINTS TO CONSIDER

Probably the most important thing to remember is that what we historically (and currently) measure and model in PKs (levels of free molecules in plasma) are a more readily available surrogate for what we are really interested in—levels of drug molecules at the sites of action. These are almost always within cells—and intracellular concentrations are much more difficult to measure (Dollery, 2013).

Stereoisomerism will influence metabolism and toxicity. For example, Lu et al. (1998) reported a comparison between (*S*)-(–)-ifosfamide and (*R*)-(+)-ifosfamide. They demonstrated that there were significant differences between the two stereoisomers with regard to PK behavior and major metabolite formation, as shown in Table 17.16.

When considering safe and efficacious use of drugs in children (pediatrics), it is important to realize that neonates and juveniles are not just scaled-down versions of adults—both their physiologies and their metabolisms/PKs are frequently different. Subsequently the PKs, pharmacodynamics, and toxicodynamics for neonates and juveniles require their own evaluations (Kearns, 2015).

In addition, treatment of animals with phenobarbital not only increased overall rates of metabolism and clearance

TABLE 17.16 Examples of Stereoselective Differences in Metabolism (*R*) versus (*S*) Ifosfamide

| Parameter | Phenobarbital | <i>R</i> | <i>S</i> | <i>R/S</i> |
|---------------------------|---------------|----------|----------|------------|
| Term half-life (min) | – | 34.3 | 41.8 | 0.820 |
| | + | 19.8 | 19.41 | 1.02 |
| AUC (μM·min) | – | 4853 | 6259 | 0.820 |
| | + | 1479 | 1356 | 1.03 |
| 2-Dechloro metabolite AUC | – | 799 | 2794 | 0.287 |
| | + | 229 | 1205 | 0.186 |
| 3-Dechloro metabolite AUC | – | 1380 | 996 | 1.41 |
| | + | 192 | 1175 | 0.159 |

Source: Adapted from Lu et al. (1998, pp. 476–482).

Note: Animals were pretreated with phenobarbital (80 mg/kg) for 4 days.

but also shifted the metabolite patterns. One of the more common methods used for determining an exposure to (or the amount of) a metabolite produced is to determine an AUC for the metabolite. Further, one of the more common methods for representing a racemically preferred metabolite is to calculate the ratio of the *R* to the *S*. For example, the 3-dechloro metabolite of ifosfamide was produced in higher amounts from the *R* enantiomer, while the 2-dechloro metabolite was the major metabolite produced from the *S* enantiomer in naive animals. Treatment with phenobarbital shifted the metabolism so that the 3-dechloro metabolite was no longer the major metabolite for the *S* enantiomer.

17.11 BIOLOGICALLY DERIVED MATERIALS

The progress and products of biotechnology have brought some new challenges to the assessment of PKs and TKs. While the reasons for why this data is needed (demonstrating exposure, displaying dose dependency, correlating any findings of toxicity to exposure, and determining steady state for systemic agent levels) are certainly as compelling as with traditional drugs, there are a whole set of special problems involved (Dennis et al., 2002; Baumann, 2006).

These special concerns for biologically derived products are:

Assay Sensitivity/Specificity

- Needs to be at 1 ng mL⁻¹ or lower.
- Cross-reactivity to native protein may confound results.
- If test article is the same as native protein, how do you tell the difference?
- Western blot can be used to demonstrate specificity.
- Antibody interference may occur with assay.

Low Systemic Levels

- Rapid metabolism: Metabolites may be endogenous proteins or amino acids.
- Extensive metabolism: Metabolites may be incorporated into cell structures rapidly.
- Rapid distribution.
- Rapid hepatic clearance.
- Route of administration may bypass first-pass metabolism:
 - SC
 - ICV, IT
 - Buccal

Endogenous Protein

- May cross-react and lead to false-positive blood levels
- Can radiolabel to tell the difference between administered molecule and endogenous molecule:
 - However, the label may lead to different distribution
 - What is the specific activity if diluted with unlabeled endogenous material?

Sample Volume

- May need to be large to increase sensitivity
- May need to be small because of competing assays:
 - Immune factors (antibodies, globulins)
 - Hormones
 - Disease state modifiers
 - In humans, concomitant medications

Distribution

- Rapidly cleared from blood.
- Frequently distributed mid-lymphatics.
- Target and off-target receptor binding found in a portion of drug molecule population in the body rapidly predominate.
- Pharmacodynamics very different from PKs.
- Delivery rarely by oral route.
- Available test material supply will be very limited in early development.

The upshot of these points is that it may not be practical to follow established guidelines for ADME evaluation. Binding proteins, immunoreactive metabolites, and antibodies could interfere with the immunoassays used to measure the activity of biotechnologically derived pharmaceuticals. The link between immunoreactivity and pharmacological activity may be difficult to establish, making the data difficult to interpret. In radiolabeled distribution studies, if the label alters the physicochemical and biological properties of the test material, its PK behavior may change. These analytical difficulties make accurate characterization of the distribution, metabolism, and excretion of a protein more difficult. However, as immunoassay-based methods of measuring levels of large molecules have improved, this situation has improved, and as more information on species-specific PKs of large molecules (such as monoclonal antibodies (mAb)) has become available, cross-species scaling of the PKs of such molecules has also become possible (Zhao et al., 2015).

AUC and C_{\max} are commonly measured to identify safety ratios for new chemical entities. Since the analytical methods used for biotechnologically derived pharmaceuticals may lack

specificity, a clinical marker of biological activity or efficacy may sometimes be more appropriate than exposure data.

It is therefore essential that before pivotal (repeat dose) preclinical studies are initiated, bioanalytical assay development must be completed. This has to cover potential test species and normal (and diseased humans). The assays must be validated in the sampling matrix of the toxicity test species, and one should also develop suitable assays for antibodies to the test article.

17.11.1 Immunoassay Methods

Competitive inhibition of radiolabeled hormone antibody binding by unlabeled hormone (either as a standard or an unknown mixture) is the principle of most RIAs. A standard curve for measuring antigen (hormone) binding to antibody is constructed by placing known amounts of radiolabeled antigen and the antibody into a set of test tubes. Varying amounts of unlabeled antigen are added to the test tubes. Antigen–antibody complexes are separated from the antigen and the amount of radioactivity from each sample is measured to detect how much unlabeled antigen is bound to the antibody. Smaller amounts of radiolabeled antigen–antibody complexes are present in the fractions containing higher amounts of unlabeled antigen. A standard curve must be constructed to correlate the percentage of radiolabeled antigen bound with the concentration of unlabeled antigen present.

Two methods are commonly employed in RIAs to separate antigen–antibody complexes. The first, the double-antibody technique, precipitates antigen–antibody complexes out of solution by utilizing a second antibody, which binds to the first antibody. The second most commonly used method is the dextran-coated activated charcoal technique. Addition of dextran-coated activated charcoal to the sample followed immediately by centrifugation absorbs free antigen and leaves antigen–antibody complexes in the supernatant fraction. This technique works best when the molecular weight of the antigen is 30kDa or less. Also, sufficient carrier protein must be present to prevent adsorption of unbound antibody.

Once a standard curve has been constructed, the RIA can determine the concentration of hormone in a sample (usually plasma or urine). The values of hormone levels are usually accurate using the RIA, but certain factors (e.g., pH or ionic strength) can affect antigen binding to the antibody. Thus, similar conditions must be used for the standard and the sample.

Problems of RIAs include lack of specificity. This problem is usually due to nonspecific cross-reactivity of the antibody. RIA represents an analytical approach of great sensitivity. Unlike assays that often require large amounts of tissue (or blood), the greater sensitivity of the RIAs or mAb techniques can be achieved using small samples of biological fluids. Some of these RIA methodologies are more useful

than others and to some extent depend on the degree of hormonal cross-reactions or, in the case of mAb methods, their degree of sensitivity.

Enzyme-linked immunosorbent assay (ELISA) is comparable to the immunoradiometric assay (IRMA) except that an enzyme tag is attached to the antibody instead of a radioactive label. ELISAs have the advantage of no radioactive materials and produce an end product that can be assessed with a spectrophotometer. The molecule of interest is bound to the enzyme-labeled antibody, and the excess antibody is removed for IRMAs. After excess antibody has been removed or the second antibody containing the enzyme has been added (two-site assay), the substrate and cofactors necessary are added in order to visualize and record enzyme activity. The level of molecule of interest present is directly related to the level of enzymatic activity. The sensitivity of the ELISAs can be enhanced by increasing the incubation time for producing substrate.

IRMAs are like RIAs in that a radiolabeled substance is used in an antigen–antibody reaction, except that the radioactive label is attached to the antibody instead of the hormone. Furthermore, excess of antibody, rather than limited quantity, is present in the assay. All the unknown antigen becomes bound in an IRMA rather than just a portion, as in a RIA; IRMAs are more sensitive. In the one-site assay, the excess antibody that is not bound to the sample is removed by addition of a precipitating binder. In a two-site assay, a molecule with at least two antibody binding sites is adsorbed onto a solid phase, to which one of the antibodies is attached. After binding to this antibody is completed, a second antibody labeled with ^{125}I is added to the assay. This antibody reacts with the second antibody binding site to form a “sandwich,” composed of antibody–hormone–labeled antibody. The amount of hormone present is proportional to the amount of radioactivity measured in the assay.

With enzyme-multiplied immunoassay technique (EMIT) assays, enzyme tags are used instead of radiolabels. The antibody binding alters the enzyme characteristics, allowing for measurement of target molecules without separating the bound and free components (i.e., homogeneous assay). The enzyme is attached to the molecule being tested. This enzyme-labeled antigen is incubated with the sample and with antibody to the molecule. Binding of the antibody to the enzyme-linked molecule either physically blocks the active site of the enzyme or changes the protein conformation so that the enzyme is no longer active. After antibody binding occurs, the enzyme substrate and cofactor are added, and enzyme activity is measured. If the sample contains subject molecules, it will compete with enzyme-linked molecules for antibody binding, enzyme will not be blocked by antibody, and more enzyme activity will be measurable.

Most protein drug entities can now be assessed using mAb techniques. It is possible to produce antisera containing

TABLE 17.17 Advantages and Disadvantages of Monoclonal Antibodies Compared with Polyclonal Antisera

| Advantages | Disadvantages |
|---|--------------------------------|
| Sensitivity | Overly specific |
| Quantities available | Decreased affinity |
| Immunologically defined | Diminished complement fixation |
| Detection of neoantigens on cell membrane | Labor intensive; high cost |

a variety of polyclonal antibodies that recognize and bind many parts of the molecule. Polyclonal antisera can create some nonspecificity problems such as cross-reactivity and variation in binding affinity. Therefore, it is oftentimes desirable to produce a group of antibodies that selectively bind to a specific region of the molecule (i.e., antigenic determinant). In the past, investigators produced antisera to antigenic determinants of the molecule by cleaving the molecule and immunizing an animal with the fragment of the hormone containing the antigenic determinant of interest. This approach solved some problems with cross-reactivity of antisera with other similar antigenic determinants, but problems were still associated with the heterogeneous collection of antibodies found in polyclonal antisera.

The production of mAbs offers investigators a homogeneous collection of antibodies that could bind selectively to a specific antigenic determinant with the same affinity. In addition to protein isolation and diagnostic techniques, mAbs have contributed greatly to RIAs.

While mAbs offer a highly sensitive, specific method for detecting antigen, sometimes increasing mAb specificity compromises affinity of the antibody for the antigen. In addition, there is usually decreased complement fixation, and costs are usually high for preparing and maintaining hybridomas that produce mAbs (Table 17.17).

The mAb techniques provide a means of producing a specific antibody for binding antigen. This technique is useful for studying protein structure relations (or alterations) and has been used for devising specific RIAs.

17.11.1.1 Metabolism and Elimination Biologics are usually not excreted unchanged in urine. They are degraded to small peptides and individual amino acids with pathways equally and generally understood for endogenous compounds. Their metabolites (amino acids) are reutilized in the endogenous amino acid pool for the *de novo* biosynthesis of structural or functional body proteins.

The metabolism of biologics is highly dependent on structure (including sugars), charge (density and distribution), size, and hydrophilicity/lipophilicity. Sites of metabolism of biologics are the liver, the kidneys, and the blood and the extravascular sites of administration. In the liver, hepatocytes are mainly responsible for the catabolism of

biologics using carrier-mediated membrane transport as well as endocytosis/pinocytosis for transport process. The kidneys play a major role in the catabolism of many small polypeptides. After being filtered by the glomeruli, some proteins are reabsorbed by the proximal tubule by endocytosis, while small amino acid chains are hydrolyzed at the brush border. Controversy exists surrounding glomerular filtration selectivity regarding size, molecular conformation, and charge of the protein (Tang et al., 2004). The often observed incomplete bioavailability of biologics after extravascular injection can be attributed to local metabolism. Catabolism at extravascular sites has been observed for, for example, insulin, calcitonin, and interferon- β (for review, see Mohler et al., 1992).

Several biologics, especially antibodies, show high inter-individual variation of PK parameters, which is mainly clearance driven. It is now clear that mAbs which target cellular antigens have far more complex, nonlinear PKs such that the half-life of these drugs can be both dose and time dependent (Lobo et al., 2004). When antigen concentration is high, half-life is short because the mAb is rapidly cleared from the blood through antigen–mAb interaction. As the antigen is depleted, clearance decreases and half-life is consequently prolonged. As the mAb accumulates, a new steady state is reached. Eventually, the target is totally depleted, at which time the clearance of mAb will be at its slowest. At this point, half-life will be at its longest, approaching the half-life of endogenous IgG (~21 days). More probable than total target depletion is saturation of the target–mAb binding with similar consequences.

Due to catabolism of proteins to (mostly) endogenous amino acids, classical biotransformation studies are performed for small molecules are not needed. Additionally, limitations of current analytical methods to detect and distinguish metabolites, and the putative lack of pharmacological or toxicological activity of the metabolites, remain obstacles. Similarly, mass balance studies usually used to determine the excretion pathways of small molecules (and their metabolites) are not used for biologics.

The majority of therapeutic biologics, especially after chronic administration, elicit an immune response in test animal species and often also in humans. This is an inherent property when administering nonhuman sequence proteins of sufficient size. Initial success in reducing immunogenicity has been achieved by replacing biologics obtained from nonhuman sources with human sequences. Antibody formation can also occur in immunocompetent recipients after treatment with products derived from human sera and tissues and also with recombinant human proteins that are identical or nearly identical in sequence to native human proteins. The mechanism for generation of antibodies of recombinant human proteins is not well understood. In most cases, the underlying mechanism is the breaking of immune tolerance that typically exists to self-antigens

(Baumann, 2006). Other reasons for immunogenicity relate to manufacturing, formulation, and storage (e.g., aggregates). These are especially addressed when modifications of these processes are performed which might influence the physicochemical properties of the product. An immune response to a product does not mean it cannot be developed. However, the development and use of a product may be complicated and, in rare cases, impossible. In certain cases, some patients develop antibodies which neutralize the biological activity of the therapeutic product and become unresponsive to treatment. Alterations in the PK

profile due to immunomediated clearance mechanisms may affect the PK profiles and the interpretation of the preclinical toxicity data. Last but not least, safety issues like immunomediated toxicity may be raised. Detection and characterization of the immune response in patients are expected by the authorities (USFDA, 2002).

The development and usefulness of appropriate animal models for testing immunogenicity are still unclear. Conventional animal models have poorly predicted immunogenicity problems in humans. One limitation of traditional animal models is that tolerance, a key aspect of the immune

TABLE 17.18 Selected Human Transporters Compared to Monkeys, Pigs, and Dogs

| Selected Human Transporters | Accession Numbers (GenBank) | % Homology (Zhang et al., 2000) of Amino Acids | | |
|---|---|--|--------------------------------|---|
| Name (Synonym) (Super Family) (<i>Gene</i>) | Human, Monkey, Pig, Dog | Number of Amino Acids: Humans/Animals | | |
| Pgp (MDR1) (ABC efflux transporter) (<i>ABCB1</i>) | NP_000918.2 NP_001028059.1 XP_003130253.2 NP_001028059.1 | 96 1280/1283 | 89 1280/1286 | 90 1280/1280 |
| BCRP (ABC efflux transporter) (<i>ABCG2</i>) | NP_004818.2 NP_001028091.1 NP_999175.1 NP_001041486 | 97 655/654 | 84 655/656 | 83 655/655 |
| OATP1B1 (SLC influx transporter) (<i>SLCO1B1</i>) | NP_006437.3 XP_001097704.1 — — | 92 691/691 | No homologue | No homologue |
| OATP1B3 (SLC influx transporter) (<i>SLCO1B3</i>) | NP_062818.1 NP_001028113.1 XP_003355596.1 NP_001159519.1 | 93 702/702 | 72 702/684 | 72 702/692 |
| OAT1 (SLC influx transporter) (<i>SLC22A6</i>) | NP_004781.2 XP_0011115901 NP_001001261.1 XP_533258.1 | 95(97) ^a 563/550 | 90(89) ^b 563/547 | 90 563/550 |
| OAT3 (SLC influx transporter) (<i>SLC22A8</i>) ^c | NP_004245.2 BAD99108 NP_999620.1 XP_533257.3 | 96(96) ^a 542/542 | 82(81) ^c 542/543 | 78 542/613 Insignificantly expressed in the liver ^d |
| OCT1 (SLC influx transporter) (<i>SLC22A1</i>) | NP_003048.1 XP_002747211.1 NP_999154.1 XP_850971.2 | 91 554/506 | 75 554/554 | 74 554/521 |
| OCT2 (SLC influx transporter) (<i>SLC22A2</i>) | NP_003049.2 XP_002804005.1 NP_999067.1 XP_533466.2 | 94 555/708? | 86 555/554 | 84 555/533 |

Source: Extracted from Dalgaard (2015).

Note: Values were derived from NCBI (2012) using BLAST (<http://blast.ncbi.nlm.nih.gov/Blast.cgi>).

^a Tahara et al. (2005).

^b Hagos et al. (2002).

^c Hagos et al. (2005).

^d Bleasby et al. (2006).

response, is highly species specific. However, characterization of the immune response in research and preclinical development is necessary to get a valid interpretation of the preclinical efficacy and safety data.

Not only the compound structure itself but also the administration route may affect immunogenicity. Extravascular injection is known to stimulate antibody formation more than IV application. This is most likely due to increased immunogenicity of protein aggregates and precipitates formed at the injection site.

To lower the systemic clearance and increase elimination half-lives, several strategies have been developed including polyethylene glycol (PEG) attachment (PEGylation), glycosylation, or fusion to proteins with decreasing clearance and prolonged serum half-lives. PEGylation improves the PK behaviors by increasing the effective size of the protein, with most significant effects for proteins smaller than 70 kDa. However, conjugation of the protein may also alter receptor affinity and biodistribution, changing the concentration–response profile for the protein independent of effects on PKs. PEGylation can also reduce immunogenicity and aggregation.

17.12 POINTS TO CONSIDER

In the last decade, our understanding of the function and significance of transporters in both systemic and organ-specific absorption of drug molecules (and their retention in potential therapeutic or toxicologically significant organs) has become clear. Table 17.18 presents a summary comparison of these functions in humans and the three major nonrodent species.

REFERENCES

- Bauer LA. (2001) *Applied Clinical Pharmacokinetics*. McGraw-Hill, New York.
- Baumann A. (2006) Early development of therapeutic biologics–pharmacokinetics. *Curr. Drug Metab.* 7:15–21.
- Bleasby K, Castle JC, Roberts CJ, Cheng C, Bailey WJ, Sina JF, Kulkarni AV, Hafey MJ, Evers R, Johnson JM, Ulrich RG, Slatter JG. (2006) Expression profiles of 50 xenobiotic transporter genes in humans and pre-clinical species: a resource for investigations into drug disposition. *Xenobiotica*. 36(10–11):963–988.
- Bronaugh RL. (1998) Methods for in vitro percutaneous absorption. In: *Dermatotoxicology Methods* (Marzulli FN and Maibach HI, Eds.). Taylor & Francis, Philadelphia, PA.
- Connally R, Anderson M. (1991) Biologically based pharmacokinetic models: tools for toxicological research risk assessment. *Annu. Rev. Pharmacol. Toxicol.* 31:503–523.
- Crooks PA, Hendrickson HP, Worthen DR, Byrd GD, deBethizy JD, Caldwell WG. (2014) Modern instrumental methods for studying mechanisms of toxicity. In: *Principles and Methods of Toxicology* (Hayes AW and Kruger CL, Eds.), 6th Ed. CRC Press, Boca Raton, FL, pp. 1965–2029.
- Dalgaard L. (2015) Comparison of minipig, dog, monkey and human drug metabolism. *J. Pharmacol. Toxicol. Methods*. 74:80–92.
- Dennis MS, Zhang M, Meng YG, Kadkhodayan M, Kirchokofer D, Combs D, Damicot LA. (2002) Albumin binding as a general strategy for improving the pharmacokinetics of proteins. *J. Biol. Chem.* 277:35035–35043.
- Dollery CT. (2013) Intracellular drug concentrations. *Clin Pharmacol. Ther.* 93:263–266.
- Dow P. (2006) In vitro species comparisons and metabolic identification. In: *Current Protocols in Pharmacology* (Enna SJ and Williams M, Eds.). John Wiley & Sons, Inc., Hoboken, NJ, pp. 7.10.1–7.10.14.
- Draper A, Madan A, Smith K, Parkinson A. (1998) Development of a non-high pressure liquid chromatography assay to determine testosterone hydroxylase (CYP3A) activity in human liver microsomes. *Drug Metab. Dispos.* 26:299–304.
- EMA. (January 2006) *Guideline on the Clinical Investigation of the Pharmacokinetics of Therapeutic Proteins*. European Medicines Agency, London.
- FDA. (2007) Chapter IV. Guidelines for toxicity studies. In: *Toxicological Principles for the Safety of Food Ingredients, Redbook II*. Center for Food Safety and Applied Nutrition, FDA, Washington, DC, pp. 138–148.
- Gabrielsson J, Weiner D. (1997) *Pharmacokinetic/Pharmacodynamic Data Analysis: Concepts and Applications*. Apotekarsocieteten, Stockholm.
- Gad SC. (Ed.) (2007) *Toxicology of the Gastrointestinal Tract*. CRC Press, Boca Raton, FL.
- Gad SC. (Ed.) (2015) *Animal Models in Toxicology*, 3rd Ed. Marcel Dekker, Inc., New York.
- Gad SC, Chengelis CP. (1998) *Acute Toxicology Testing*, 2nd Ed. Academic Press, San Diego, CA.
- Garner C, Matthews H. (1998) The effect of chlorine substitution on the dermal absorption of polychlorinated biphenyls. *Toxicol. Appl. Pharmacol.* 149:150–158.
- Goldstein A, Aronow L, Kalman S. (1974) *Principles of Drug Action: The Basis of Pharmacology*. John Wiley & Sons, Inc., New York.
- Griffin R, Godfrey V, Kim Y, Burka L. (1997) Sex-dependent differences in the disposition of 2,4-dichlorophenoxyacetic acid in Sprague-Dawley rats, B6C3F1 mice and Syrian hamsters. *Drug Metab. Dispos.* 25:1065–1071.
- Hagos Y, Bahn A, Asif AR, Krick W, Sendler M, Burkhardt G. (2002) Cloning of the pig renal organic anion transporter 1 (pOAT1). *Biochimie*. 84(12):1221–1224.
- Hagos Y, Braun IM, Krick W, Burkhardt G, Bahn A. (2005) Functional expression of pig renal organic anion transporter 3 (pOAT3). *Biochimie*. 87:421–424.
- Holland J, Kao M, Whitaker MJ. (1984) A multisample apparatus for kinetic evaluation of skin penetration in vitro: the influence and metabolic status of the skin. *Toxicol. Appl. Pharmacol.* 72:272–280.
- Hunt JAB and Knox MT. (1968) “Enterogastric Reflux and Gastric Emptying Chemical or Mechanical” IN: (Fawcett G, Ed.) *Medications and Drugs in Gastrointestinal Motility*. Springer-Verlag, Heidelberg, pp. 265–290.

- ICH. (2004a) *S3A: Toxicokinetics: The Assessment of Systemic Exposure in Toxicity Studies*. International Conference on Harmonisation, Geneva, Switzerland.
- ICH. (2004b) *S3B: Pharmacokinetics: Guidance for Repeated Dose Tissue Distribution Studies*. International Conference on Harmonisation, Geneva, Switzerland.
- Kearns GL. (2015) Selecting the proper pediatric dose: it is more than size that matters. *Clin. Pharmacol Ther.* 98:238–240.
- Keberle G, Brindle SD, Greengard P. (1971) The route of absorption of intraperitoneally administered compounds. *J. Pharmacol. Exp. Ther.* 178:562–566.
- La Du B, Mandel H, Way E. (1972) *Fundamentals of Drug Metabolism and Drug Disposition*. The Williams & Wilkins Co., Baltimore, MD.
- Lavé T, Parrott N, Grimm HP, Fleury A, Reddy M. (2007) Challenges and opportunities with modeling and simulation in drug discovery and drug development. *Xenobiotica.* 37:1295–1310.
- Levy RH, Thummel KD, Tauger WF, Hansten PD, Meichelbaum E. (2000) *Metabolic Drug Interactions*. Lippincott, Williams & Wilkins, Philadelphia, PA.
- Lobo ED, Hansen RJ, Balthasar JP. (2004) Antibody pharmacokinetics and pharmacodynamics. *J. Pharm. Sci.* 93(11):2645–2668.
- Lu H, Wang J, Chan K, Young D. (1998) Effects of phenobarbital of stereoselective metabolism of ifosfamide in rats. *Drug Metab. Dispos.* 26:476–482.
- Meyer UA. (1994) The molecular basis of genetic polymorphisms of drug metabolism. *J. Pharm. Pharmacol.* 46(Suppl. 1): 409–415.
- Minami H and MacCallum RN. (1984) Gastric emptying in early noninsulin-dependent diabetes mellitus. *Br J Radiology.* 57:689–695.
- Mohler AM, Cook JE, Baumann G. (1992) Binding proteins of protein therapeutics. In: *Protein Pharmacokinetics and Metabolism* (Ferraiolo BL, Mohler MA and Gloff CA, Eds.). Plenum Press, New York/London, pp. 35–71.
- Monostory K, Jemnitz K, Vereczkey L, Czira G. (1997) Species differences in metabolism of panomifene, an analogue of tamoxifen. *Drug Metab. Dispos.* 25:1370–1378.
- Mugfor C, Kidderis G. (1998) Sex-dependent metabolism of xenobiotics. *Drug Metab. Rev.* 30:441–498.
- Myers MJ, Farrell DE, Howard KD, Kawalek JC. (2001) Identification of multiple constitutive and inducible hepatic cytochrome P450 enzymes in the market weight swine. *Drug Metab. Dispos.* 29(6):908–915.
- National Center for Biotechnology Information (NCBI) (2012) *BLAST Assembled Genomes*. National Library of Medicine, Bethesda, MD. Available at: <http://blast.ncbi.nlm.nih.gov/Blast.cgi> (accessed May 4, 2016).
- Nebbia C, Dacasto M, Rossetto Giaccherino A, Giuliano Albo A, Carletti M. (2003) Comparative expression of liver cytochrome P450-dependent monooxygenases in the horse and in other agricultural and laboratory species. *Vet. J.* 165(1):53–64.
- O'Reilly WJ. (1972) Pharmacokinetics in drug metabolism and toxicology. *Can. J. Pharm. Sci.* 7:66–77.
- Parkinson A. (1996) Biotransformation of xenobiotics. In: *Toxicology: The Basic Science of Poisons* (Cassarett LJ, Doull J and Klaassen C, Eds.), 5th Ed. McGraw-Hill, Inc., New York.
- Pratt WB, Taylor P. (1990) *Principles of Drug Action: the Basis of Pharmacology*, 3rd Ed. John Wiley & Sons, Inc., New York.
- Roberts A, Renwick AG. (2014) Toxicokinetics. In: *Principles and Methods of Toxicology* (Hayes AW and Kreuger CL, Eds.), 6th Ed. CRC Press, Boca Raton, FL, pp. 215–246.
- Shapiro J. (1981) *Radiation Protection*, 2nd Ed. Harvard University Press, Cambridge, MA, pp. 12–18.
- Shargel L, Yu ABC. (1999) *Applied Biopharmaceutics & Pharmacokinetics*. Appleton & Lange, New York.
- Shimada T, Mimura M, Inoue K, Nakamura S, Oda H, Ohmori S, Yamazaki H. (1997) Cytochrome P450-dependent drug oxidation activities in liver microsomes of various animal species including rats, guinea pigs, dogs, monkeys, and humans. *Arch. Toxicol.* 71(6):401–408.
- Skaanild MT, Friis C. (1999) Cytochrome P450 sex differences in minipigs and conventional pigs. *Pharmacol. Toxicol.* 85(4):174–180.
- Smith D. (1991) Species differences in metabolism and pharmacokinetics: are we close to an understanding? *Drug Metab. Rev.* 23:355–373.
- Smith DA, Kerns EH. (2010) The effect of plasma protein binding on in vivo efficacy: misconceptions in drug discovery. *Nat. Rev. Drug Discov.* 9:929–939.
- Stevens JC, Shipley LA, Cashman JR, Vandenbranden M, Wrighton SA. (1993) Comparison of human and rhesus monkey in vitro phase I and phase II hepatic drug metabolism activities. *Drug Metab. Dispos.* 21(5):753–760.
- Tahara H, Shono M, Kusuha H, Kinoshita H, Fuse E, Takadate A, Otagiri M, Sugiyama Y. (2005) Molecular cloning and functional analyses of OAT1 and OAT3 from cynomolgus monkey kidney. *Pharm. Res.* 22(4):647–660.
- Tang L, Persky AM, Hochhaus G, Meibohm B. (2004) *J. Pharm. Sci.* 93:2184–2204.
- Tse FLS, Jaffe JM. (1991) *Preclinical Drug Disposition*. Marcel Dekker, Inc., New York.
- US Food and Drug Administration. (October 2002) *Guidance for Industry—Immunotoxicological Evaluation of Investigation New Drugs*. Center for Drug Evaluation and Research, Washington, DC.
- USFDA. (2005) *Guidance for Industry: Safety Testing of Drug Metabolizers*. CDEB, June 2006.
- Vinegar A, Jepson G. (1996) Cardiac sensitization thresholds of halon replacement chemicals in humans by physiologically based pharmacokinetic modeling. *Risk Anal.* 16:571–579.
- Waddell WJ. (1972) Autoradiography in drug disposition studies. In: *Fundamentals of Drug Metabolism and Drug Disposition* (La Du BN, Mandel HG and Way EL, Eds.). Williams & Wilkins, New York, pp. 505–514.
- Wang YM, Reuning R. (1994) A comparison of two surgical techniques for the preparation of rats with chronic bile duct cannulae for the investigation of enterohepatic circulation. *Lab. Anim. Sci.* 44:479–485.
- Yacobi A, Skelly JP, Batera VK. (1989) *Toxicokinetics and New Drug Development*. Pergamon Press, New York.
- Zhang Z, Schwartz S, Wagner L, Miller W. (2000) A greedy algorithm for aligning DNA sequences. *J. Comput. Biol.* 7:203–214.
- Zhao J, Cao Y, Jusko WJ. (2015) Across-species scaling of monoclonal antibody pharmacokinetics using a minimal PBPK model. *Pharm. Res.* 32:3269–3281.

SAFETY PHARMACOLOGY

Safety pharmacology is the evaluation and study of the potentially life-threatening pharmacologic effects of a potential drug which is unrelated to the desired therapeutic effect, and therefore may present a hazard—particularly in individuals who already have one or more compromised or limited organ system functions. Unlike other nonclinical evaluations of the safety of a drug, these evaluations are usually conducted at doses not too much in excess of the intended clinical dose. This topic is another which has had to undergo significant change since the last edition because of a variety changed regulatory environment.

General/safety pharmacology was an emerging discipline within the pharmaceutical industry prior to 2002 (Hite, 1997), when ICH guidance was promulgated, and it became a major area of both concern and activity. It seeks to identify unanticipated effects of new drug candidates on major organ function (i.e., secondary pharmacological effects) and ensure that they are critically assessed in a variety of animal models. A survey was conducted to obtain customer input on the role and strategies of this emerging discipline. Overlooked in importance by all but a few (Zbinden, 1966, 1984) for many years, the Japanese clearly became the leaders in developing and requiring such information, while the United States was in a position behind Japan and the EU in both having formal requirements and implementing industrial programs. While major companies were aware and largely addressing the need by the mid-1990s (Kinter et al., 1994; Murphy et al., 1995; Sullivan and Kinter, 1995), and EMEA promulgated guidelines in 2000, it was only with the ICH S7 guidance (ICH, 2005) that US and global regulatory interest came to play.

While historically companies have conducted evaluations of cardiovascular and CNS functions, few evaluate respiratory, gastrointestinal (GI), and renal functions; a few

conduct a ligand-binding/activity panel as part of their pharmacological profiling. Since 2001, studies to evaluate the cardiovascular, respiratory, pulmonary, and central nervous system (CNS) safety pharmacology aspects of all but a few new drugs have been required (these exceptions will be discussed shortly) before they are evaluated clinically (i.e., had human exposure).

It is recognized to be important that the tests employed detect bidirectional drug effects and that the tests performed be validated in both directions with appropriate reference (control) substances and be sensitive in the acute therapeutic range (Folke, 2000; Redfern et al., 2002). This requirement is less appropriate for multiparameter procedures. Blind testing could be an advantage. Ethical considerations are important, but the ultimate ethical criterion is the assessment of risk for humans. Safety pharmacology studies should not be overinclusive, but should be performed to the most exacting standards, including Good Laboratory Practice (GLP) compliance. Safety pharmacology data is acknowledged to be important to be available during the planning stage for phase I studies, but such is still less often the case than not. Partly this arises from the viewpoint that human tolerance (particularly in a well-designed and executed phase I study in normal volunteers) is, in itself, an adequate assessment of safety pharmacology. This is, of course, backward—such human tolerance is rather, properly, an extension (and expression) of the nonclinical safety pharmacology.

The other point of view in the past has been that properly executed repeated-dose preclinical safety studies meeting the current design will (or could) fill these needs. Recognizing that undesired pharmacological activities of novel drugs or biologics may limit development of a therapeutic agent prior to the characterization of any

toxicological effects in rodent species, general pharmacology assays have traditionally been used to screen new agents for pharmacological effects on the central and peripheral nervous systems, the autonomic nervous system and smooth muscles, the respiratory and cardiovascular systems, the digestive system, and the physiological mechanisms of water and electrolyte balance (Cavero, 2009). In large animal species, such as dogs and nonhuman primates, a smaller number of animals per study limit their use for screening assays, but these species may play an important role in more detailed mechanistic studies. For drugs and biologics that must be tested in nonhuman primates because of species-specific action of the test agent, functional pharmacology data are often collected during acute or subacute toxicity studies. This requires careful experimental design to minimize any impact that pharmacological effects or instrumentation may have on the assessment of toxicity. In addition, with many new therapies targeted at immunological diseases, the pharmacological effect of therapeutics on the immune system presents new challenges for pharmacology profiling. The applications of pharmacology assays by organ system in both rodent and large animal species are discussed, as well as practical issues in assessing pharmacological end points in the context of toxicity studies (Martin et al., 1997; Matsuzawa et al., 1997; Pugsley et al., 2008; Dorato et al., 2015; Pugsley and Curtis, 2015).

Pharmacoepidemiology studies in Europe and the United States show that adverse drug reactions (ADRs) now may account for up to 10% of the admissions of patients to hospitals at a cost of hundreds of millions of US dollars annually (Sjoquist, 2000). This represents a considerable increase compared to 20 years ago. A partial explanation is the many shortcomings of clinical trials and their relevance for health-care. ADRs are often poorly studied and documented in these studies and very seldom included in health economical analyses of the value of new drugs. Pharmacovigilance is product—rather than utilization—oriented and quite invisible in clinical medicine. This is regrettable, since up to 50% of ADRs are dose dependent and thus preventable. Hopefully, the rapid progress in molecular and clinical pharmacogenetics will provide new tools for clinicians to choose and dose drugs according to the individual needs of patients. A good starting point for those not well versed in pharmacology and the range of potential mechanisms of action and of interaction can be found in Goodman and Gilman (Brunton et al., 2011).

18.1 REGULATORY REQUIREMENTS

While the ICH guidelines promulgated in November of 2000 (implemented in Europe in the three regions in June of 2001) are the announced international standards for regulation, the

TABLE 18.1 Cardiovascular System Safety Pharmacology Evaluations

Core

Hemodynamics (blood pressure, heart rate)
Autonomic function (cardiovascular challenge)
Electrophysiology (ECG in dog)

QT prolongation (noncore)

An additional guideline, ICH S7B, is in preparation which will address the assessment of potential for QT prolongation. In the meantime, CPMP 986/96 indicates the following preclinical studies should be conducted prior to first administration to man:
Cardiac action potential *in vitro*
ECG (QT measurements) in a cardiovascular study which would be covered in the core battery
hERG channel interactions (hERG expressed in HEK293 cells)

TABLE 18.2 Respiratory System Safety Pharmacology Evaluation

Respiratory functions

Measurement of rate and relative tidal volume in conscious animals

Pulmonary function

Measurement of rate, tidal volume, and lung resistance and compliance in anesthetized animals

actual situation in different countries remains very mixed (Fujimori, 1999; Olejniczak, 1999).

In the EU, the CPMP issues a draft “Note for Guidance on Safety Pharmacology Studies in Medicinal Product Development” in 1998, but it has not yet been finalized or put in force, and as of the middle of 2001, the US FDA has remained mute on guidelines.

The actual requirements of the initial November 8, 2000, ICH guidelines provided only broad outlines of requirements. They called for the conduct of studies in a core battery to assess effects on the cardiovascular (Table 18.1), respiratory (Table 18.2), central nervous (Table 18.3), and secondary organ systems (Table 18.4). Follow-up studies for the core battery are also required on a case-by-case basis for the three main organ systems:

- *CNS* Behavioral pharmacology, learning and memory, specific ligand binding, neurochemistry, visual, auditory and/or electrophysiology examinations, and so on
- *Cardiovascular system* Behavioral pharmacology, learning and memory, specific ligand binding, neurochemistry, visual, auditory and/or electrophysiology examinations, and so on
- *Respiratory system* Tidal volume, bronchial resistance, compliance, pulmonary arterial pressure, blood gases

TABLE 18.3 Central Nervous System (CNS) Safety Pharmacology Evaluation**Irwin test**

General assessment of effects on gross behavior and physiological state^a

Locomotor activity

Specific test for sedative, excitatory effects of compounds

Neuromuscular function

Assessment of grip strength

Rotarod

Test of motor coordination

Anesthetic interactions

Test for central interaction with barbiturates

Anti-/proconvulsant activity

Potentiation or inhibition of effects of pentylenetetrazole

Tail flick

Tests for modulation of nociception (also hot plate, Randall–Selitto, tail pinch)

Body temperature

Measurement of effects on thermoregulation

Autonomic function

Interaction with autonomic neurotransmitters *in vitro* or *in vivo*

Drug dependency

Test for physical dependence, tolerance, and substitution potential

Learning and memory

Measurement of learning ability and cognitive function in rats

^aUsually a functional observational battery (FOB) is integrated into rodent (rat dose range finding or) repeat-dose toxicity studies to meet this requirement.

TABLE 18.4 Secondary Organ System Safety Pharmacology Evaluation**Renal system**

Renal function—measurement of effects on urine excretion in saline-loaded rats

Renal dynamics—measurement of renal blood flow, GFR, and clearance

GI system

GI function—measurement of gastric emptying and intestinal transit

Acid secretion—measurement of gastric acid secretion (Shay rat)

GI irritation—assessment of potential irritancy to the gastric mucosa

Emesis—nausea, vomiting

Immune system

Passive cutaneous anaphylaxis (PCA)—test for potential antigenicity of compounds

Other

Blood coagulation

In vitro platelet aggregation

In vitro hemolysis

Conditions are also defined under which studies are not necessary:

- Locally applied agents (e.g., dermal or ocular) where systemic exposure or distribution to the vital organs is low.
- Cytotoxic agents for treatment of end-stage cancer patients.
- Biotechnology-derived products that achieve highly specific receptor targeting. Recent regulatory opinions and actions by the FDA have called this exclusion into question, leading recommendations for specific cardiovascular testing strategies for biologics (Vargas et al., 2008). Likewise, specific recommendations for oligonucleotides (which are small molecules) are supported (Berman et al., 2014).
- New salts having similar pharmacokinetics and pharmacodynamics to already well-characterized drugs.

In recent years, there has been increasing support to perform as many safety pharmacology assessments not as separate freestanding studies, but rather more toward incorporating as many of these assessments as possible into existing general toxicology studies, especially while recently available technology makes this more feasible (Redfern et al., 2013).

18.2 STUDY DESIGNS AND PRINCIPLES

As a starting place, unlike older pharmacology studies, safety pharmacology studies are conducted as GLP studies unless being performed as screens. At the same time, unlike other safety assessment studies, these do not need to vastly exceed intended therapeutic doses so as to identify signs of toxicity. In this sense, they are closer to hazard tests.

General guidance for dose (or concentration) section for such studies is as follows (ICH, 2000, 2011; Gad, 2012):

- *In vivo* studies
 - Designed to define the dose–response curve of the adverse effects.
 - Doses should include and exceed primary pharmacodynamic or therapeutic range.
 - In the absence of safety pharmacology parameters, the highest doses equal or exceed some adverse effects (toxic range).
- *In vitro* studies
 - Generally designed to establish an effect–concentration relationship (range of concentrations)

Consideration in the selection and design of specific studies is straightforward.

- The following factors should be considered (selection):
 - Effects related to the therapeutic class
 - Adverse effects associated with members of the chemical/therapeutic class
 - Ligand binding or enzyme data suggesting a potential for adverse effects
 - Data from investigations that warrant further investigation
- A hierarchy of organ systems can be developed:
 - Importance with respect to life-supporting functions:
 - Cardiovascular
 - Respiratory
 - CNS
 - Functions which can be transiently disrupted without causing irreversible harm

The absence of observed activity may represent either a true- or false-negative effect. If an assay is valid for the particular test article and fails to indicate activity, it is an appropriate indicator of future events (Green, 1997). However, if the assay is insensitive or incapable of response, the test represents a form of bias, albeit unconscious. Many biological products demonstrate a specificity of response that limits the utility of commonly employed safety studies. Specificity for many biologics arises from both their physicochemical properties and their similarity to endogenous substances which are regulated in a carefully controlled manner. To overcome the issue of a lack of predictive value, various approaches may be used. For example, a multiple testing strategy of mutually reinforcing studies may be employed, or safety studies may be adaptively fit to the biological circumstance.

A separate issue is how and when to consider isomers, metabolites, and the actual finished product:

- Generally the parent compound and its major metabolite(s) that achieve systemic exposure should be evaluated.
- It may be important to test active metabolites from humans.
- Testing of individual isomers should also be considered.
- Studies with the finished product are only necessary if kinetics/dynamics are substantially altered in comparison to the active substance previously tested.

There are also special considerations as to how to statistically evaluate specific aspects of these studies. Specifically, analysis of time to event becomes very important (Anderson et al., 2000).

18.3 ORGAN SYSTEM-SPECIFIC TESTS

18.3.1 General Considerations in Selection and Design of Safety Pharmacology Studies

The following factors should be considered (selection):

- Effects related to the therapeutic class (e.g., proarrhythmia is a common feature of antiarrhythmic drugs)
- Adverse effects associated with members of the chemical/therapeutic class (e.g., antipsychotics and QT prolongation)
- Receptor/enzyme/ion channel binding data suggesting a potential for adverse effects
- Any data from previous studies that warrant further investigation

A hierarchy of organ systems is considered:

- Importance with respect to life-supporting functions:
 - Cardiovascular
 - Respiratory
 - Central nervous
- Functions which can be transiently disrupted without causing irreversible harm (e.g., urinary system, GI tract)

18.3.2 Studies on Metabolites, Isomers, and Finished Products

Generally the parent compound and its major metabolite(s) that achieve systemic exposure need to be evaluated. This means that either the test species must be metabolically comparable to humans or human metabolites must also be evaluated. The testing of individual isomers also needs to be considered.

18.4 CARDIOVASCULAR

While the initial greatest concern for cardiovascular risks was associated with QT prolongation (first raised by the CAST in 1989), it has become clear since that there are a range of potentially life-threatening cardiovascular pharmacologic drug effects which must be evaluated (see Braunwald, 2008; Gad, 2008) by drugs not intended to have cardiac effects (see Table 18.5). The cardiovascular system is one of the primary vital functions which has to be examined during safety pharmacology studies. Cardiovascular system functioning is maintained by cardiac electrical activity and by pump muscle function, which contribute to hemodynamic efficacy, and is both

TABLE 18.5 Noncardiac Drugs Known to Induce or Worsen Heart Failure according to the Suggested Mechanism(s) Implicated

| Drug Class | Drug |
|---|---|
| Cardiomyopathy | |
| Cytotoxic drugs | Doxorubicin, epirubicin, and other anthracyclines; mitoxantrone, cyclophosphamide, 5-fluorouracil, capecitabine |
| Immunomodulating drugs/antibodies | Trastuzumab, interferon- α 2, interleukin-2, infliximab, etanercept |
| Antifungal drugs | Itraconazole, amphotericin B |
| Antipsychotic drugs | Clozapine |
| Pulmonary hypertension | |
| Antimigraine drugs | Methysergide, ergotamine |
| Appetite suppressants | Fenfluramine, fluramine, phentermine |
| Heart valve abnormalities | |
| Antimigraine drugs | Methysergide, ergotamine |
| Appetite suppressants | Fenfluramine, fluramine, phentermine |
| Antiparkinsonian drugs | Pergolide |
| Fluid overload | |
| NSAIDs, including cyclooxygenase-2 inhibitors | All |
| Antidiabetic drugs | Rosiglitazone, pioglitazone, troglitazone |
| Glucocorticoids | All |
| Herbal drugs | Herbal drugs containing licorice or adulterated with NSAIDs |

complex and the subject of a range of organ-specific toxicities (Gad, 2015a).

The aim of cardiovascular safety pharmacology is to evaluate the effects of test substances on the most pertinent components of this system, in order to detect potentially undesirable effects before engaging in clinical trials (Lacroix and Provost, 2000). In the basic program, a detailed hemodynamic evaluation is carried out in the anesthetized dog. It is completed by cardiac and/or cellular electrophysiology investigations in order to assess the arrhythmogenic risk. The basic program can be preceded by rapid and simple testing procedures, during the early drug discovery stage. It should be completed, if necessary, by specific supplementary studies, depending on the data obtained during the early clinical trials. The current gold standard study is performed using unrestrained radiotelemetrized dogs (Gauvin et al., 2006) and can measure multiple end points continuously.

18.4.1 Hemodynamics, ECG, and Respiration in Anesthetized Dogs or Primates

Anesthetized studies use modular instruments data capture systems to record six-lead ECG (I, II, III, aV_r , aV_l , and aV_f) (Hamlin, 2008), left ventricular pressure variables, arterial blood pressure and respiratory measurement of arterial blood flow in selected vascular beds, cardiac output, and arterial blood gas measurement, ECG intervals are measured from the lead II. ECG, and QT interval can be corrected for heart rate using Bazett's, Fridericia's, or Van de Water's formulas. Different formulas are appropriate for different species (Soloviev et al., 2006).

18.4.2 Cardiac Conduction Studies

In addition to the earlier hemodynamic measurements, intra-ventricular, intra-arterial, and atrioventricular conduction times and velocities can be measured using epicardial electrodes in the anesthetized and thoracotomized dog.

18.4.3 Conscious Dog, Primate, or Minipig Telemetry Studies

Effects on blood pressure, heart rate, lead II ECG, core body temperature, and locomotor activity can be explored using Data Sciences International (DSI) telemetry (or similar) implanted devices in guinea pigs, dogs, pigs, and primates. Effects on behavior can be captured on video using CCTV for dog and primate studies. Repeated administration and interaction studies may also be performed. This approach has the advantages of avoiding the effects of both anesthesia and restraint of the animal. Myocardial contractility has also been suggested for inclusion in these studies, and telemetry methods are now available for this (Markert et al., 2007).

18.4.4 Six-Lead ECG Measurement in the Conscious Dog and Minipig

Conscious studies using Integrated Telemetry System devices for measurement of blood pressure and six chest lead ECG measurements (V2, V4, V6, V10, rV2, and rV4). ECG interval analysis is performed on the V2 lead (RR, PR, QT, QTc intervals, QRS duration). QT dispersion can also be measured in either of these species (Milano, 2012). Locomotor activity can be monitored and behavior captured on video using CCTV:

- In addition to validated systems for automatic measurement of ECG parameters, ECGs can be reviewed by our veterinary cardiology services to detect any transparent abnormalities.

- Colonies of telemetered animals can be set up and maintained for repeat use.
- Respiration rate measurements can be taken from dogs in slings using a pneumograph system.
- An animal-specific correction of QT interval can also be derived for each dog/primate based on individual variability of QT interval with rate using the Framingham equation.

Concerns over the arrhythmogenic effects of a number of marketed compounds resulted in the issue of the “points-to-consider” document, CPMP 986/96, by the EMA (<http://www.fda.gov/ohrms/dockets/ac/03/briefing/pubs/cmpmp.pdf>).

Studies to assess the effects of compound and any known metabolites on ECG and cardiac action potentials are recommended. Changes in action potential duration and other parameters measured are a functional consequence of effects on the ion channels which contribute to the action potential. This *in vitro* test is considered to provide a reliable risk assessment of the potential for a compound to prolong QT interval in man. The guinea pig had also been recommended for such testing (Lacroix, 2002) based on prior use history, but few GLP labs have supported or offer such use.

18.4.5 Systems for Recording Cardiac Action Potentials

These include a range of currently available methodologies, some of which can be incorporated into existing study designs:

- Isolated ventricular Purkinje fibers from dog or sheep
- Isolated right ventricular papillary muscle from guinea pig
- Continuous intracellular recording of action potentials and online analysis of resting membrane potential, maximum rate of depolarization, upstroke amplitude, and action potential duration using NOTOCORD-hem data acquisition system
- Assessment of use-dependent and inverse use-dependent actions by stimulation at normal, bradycardic, and tachycardic frequencies (e.g., see in the succeeding text inverse use-dependent properties of sotalol in dog Purkinje fibers)

18.4.6 Special Case (and Concern): QT Prolongation

Drugs that alter ventricular repolarization (generally recognized as drugs that prolong the QT interval) have been associated with malignant ventricular arrhythmias (especially the distinctive polymorphic ventricular tachycardia called Torsades de pointes) and death. Many of the drugs now

known to alter ventricular repolarization were developed as antiarrhythmics (e.g., dofetilide, sotalol), but others (e.g., cisapride, terfenadine) were developed without the expectation of any effect upon electrically excitable membranes (Fenichel and Koerner, 1999). This has led to ICH promulgating ICH S7B (ICH, 2002) with specific guidance for evaluation.

The QT interval of the ECG is reflected in three main ways. First, electrophysiologically, it reflects the depolarization and repolarization phase of ventricular myocytes. Secondly, mechanically, it represents the time of contraction of the ventricles. And lastly, physiologically, its duration is a function of numerous variables (heart rate, diseases, nutrition, diurnal cycle, etc.). Diagnostically, a prolongation of its duration indicates an enhanced risk for ventricular arrhythmias (Torsades de pointes) and sudden cardiac death.

The association between abnormalities of repolarization and life-threatening arrhythmias is stronger than some other associations between laboratory abnormalities and clinical events. For example, there are drugs (tacrine) and inborn errors of metabolism (Gilbert's syndrome) that cause wild excursions in liver function tests, but with no adverse consequences. In contrast, although the severity or proarrhythmia at a given QT duration varies from drug to drug and from patient to patient, no drug is known to alter ventricular repolarization without inducing arrhythmias,¹ and each of the several congenital long QT syndromes is associated with an elevated incidence of malignant arrhythmias.

With any given repolarization-altering drug, the risk of malignant arrhythmia seems to increase with increasing QT interval, but there is no well-established threshold duration below which a prolonged QT interval is known to be harmless. The extent of QT prolongation seen with a given drug and patient may be nonlinearly related to patient factors (sex, electrolyte levels, etc.) and to serum levels of the drug and/or its metabolites. The actual incidence of malignant arrhythmias, even in association with the drugs most known to induce them, is relatively low, so failure to observe malignant arrhythmias during clinical trials of ordinary size and duration does not provide substantial reassurance.

Abnormal repolarization and the associated arrhythmias are the end results of a causative chain that starts with alterations in the channels of ionic flux through cell membranes. Some cells (e.g., those of the Purkinje system or midmyocardium) seem especially susceptible to these changes. At a substrate level, the links on the chain are alterations in the time course of the action potential, alterations in the

¹ Some QT-prolonging drugs (e.g., amiodarone; see Hohnloser et al. (1994)) are not reported to have caused many arrhythmic deaths, but this observation must be interpreted carefully. In a population with a high incidence of life-threatening arrhythmias, a drug with both proarrhythmic and antiarrhythmic effects might cause a net reduction in arrhythmias, and the arrhythmias that it had induced might not be attributed to it. In a population whose native arrhythmias were not life threatening, the same drug might result in a net decrease in mortality.

propagation of action potentials within a given cell, and alterations in the propagation of action potentials from cell to cell within syncytia and from tissue to tissue within the heart. At a higher level of aggregation, one sees “afterdepolarizations” in the terminal portion of the action potential, spontaneous beats triggered by afterdepolarizations, propagation of these beats to other cells, and reentrant excitation.

With these considerations in mind, the problem of altered repolarization should be integrated into drug development by (Malik and Camm, 2001):

- *In vitro* screening of the drug and its metabolites for effects on ion channels (especially I_{Kr})
- *In vitro* screening of the drug and its metabolites for effects on action potential duration
- Screening of the drug and its metabolites for altered repolarization in animal models
- Focused preclinical studies for proarrhythmia if altered repolarization is seen in preclinical screening or in patients

18.4.7 Some Specific Techniques Which Can Be Employed

18.4.7.1 Cloned Human Potassium Channels

- Assessment of effects on cloned human ether-à-go-go-related gene (hERG) K^+ channels stably expressed in a cell line by measurement of whole-cell K current (I_{Kr}) using voltage clamp.
- Other cloned human ion channels (e.g., $KvLQT1/minK$ -IKs currents) are also possible.

18.4.7.2 Cardiac Action Potential In Vitro: Purkinje Fibers

- Intracellular recording of action potentials from cardiac Purkinje fibers isolated from dog or sheep ventricle
- Measurement of maximum rate of depolarization and action potential duration to detect sodium and potassium channel interactions, respectively, according to recommendations in the EMA “points-to-consider” document—CPMP 986/96 (2000)

18.4.7.3 Monophasic Action Potential in Anesthetized Guinea Pigs

- Epicardial monophasic action potential recording using suction/contact pressure electrodes according to Carlsson et al. (1997)
- Simultaneous measurement of ECG

18.4.7.4 ECG by Telemetry in Conscious Guinea Pigs Lead II ECG recording using DSI telemetry device. Repeated administration and interaction studies can be performed.

18.4.7.5 Hemodynamics and ECG in Anesthetized or Conscious Dogs or Primates

- Conscious studies using DSI telemetry for blood pressure and lead II ECG or the ITS system for blood pressure and six chest lead ECG measurements (including QT dispersion)
- Anesthetized studies using MI^2 data capture system with additional measurement of blood flow in selected vascular beds, cardiac output, and respiratory and left ventricular function

18.4.8 Relevance of hERG to QT Prolongation

Compounds which are associated with ADRs of QT prolongation, arrhythmias such as Torsades de pointes, and sudden death predominantly have a secondary pharmacological interaction with the rapidly activating delayed rectifier potassium channel I_{Kr} . The gene encoding this channel has been identified as *hERG*. Testing of compounds for interactions with the hERG channel allows the identification of potential risk of QT prolongation in humans and can be used as a screen in development candidate selection.

18.4.8.1 Expression and Recording Systems HEK293 cells have been transfected with cDNA for hERG1 to produce a stable expression system. The cell line has been obtained under license for the laboratory of Craig January at the University of Wisconsin (Mohammad et al., 1997).

18.5 CENTRAL NERVOUS SYSTEM

The primary screening tool for CNS safety pharmacology evaluation is the functional observational screen, which seeks to use objective but noninvasive methods for evaluating the pharmacologic effects of a drug on peripheral and nervous system effects.

Initially, the starting basis for such screens was the Irwin screen (Irwin, 1968), which is used to screen for effects in mice and still one of the ICH S7-designated primary screens for fulfilling the regulatory requirements for S7 evaluation of new drugs.

More commonly used is the rat functional observational battery (FOB), initially developed by Gad (1982) and subsequently further modified (Mattsson et al., 1996). Haggerty (1991) and Moscardo et al. (2007) present an excellent description of the rodent FOB as currently performed in the pharmaceutical industry. Other modifications/versions which are included under the ICH screening guidelines cover the use of the dog. Gad has developed, validated, and published FOB designs for dogs (Gad and Gad, 2003) primates and most recently minipigs.

The neurobehavioral screens which serve to meet the primary regulatory requirement for CNS safety pharmacology evaluation all, of necessity, actually evaluate both central

TABLE 18.6 Isolated Tissue Pharmacologic Assays

| Assay System | End Point | Standards (Agonist/Antagonist) | Reference |
|--|-------------------------|--|---------------------------------|
| Rat ileum | General activity | None (side-spectrum assay for intrinsic activity) | Domer (1971) |
| Guinea pig vas deferens | Muscarinic | Methacholine/atropine | Leach (1956) |
| | Nicotinic | Methacholine/hexamethonium | |
| | Muscarinic | Methacholine/atropine | |
| Rat serosal strip | Nicotinic | Methacholine/hexamethonium | Khayyal et al. (1974) |
| Rat vas deferens | Alpha-adrenergic | Norepinephrine/phenoxybenzamine | Rossum (1965) |
| Rat uterus | Beta-adrenergic | Epinephrine/propranolol | Levy and Tozzi (1963) |
| Rat uterus | Kinin receptors | Bradykinin/none | Gecse et al. (1976) |
| Guinea pig tracheal chain | Dopaminergic | Dopamine/none | Domer (1971) |
| Rat serosal strips | Tryptaminergic | 5-Hydroxytryptamine (serotonin)/dibenzylamine or lysergic acid dibromide | Lin and Yeoh (1965) |
| Guinea pig tracheal chain | Histaminergic | Histamine/benadryl | Castillo and De Beer (1947a, b) |
| Guinea pig ileum (electrically stimulated) | Endorphin receptors | Methenkephalin/none | Cox et al. (1975) |
| Red blood cell hemolysis | Membrane stabilization | Chlorpromazine (not a receptor-mediated activity) | Seeman and Weinstein (1966) |
| Frog rectus abdominis | Membrane depolarization | Decamethonium iodide (not a receptor-mediated activity) | Burns and Paton (1951) |

and peripheral nervous system functions. Such evaluations used, to the maximum extent possible, semiquantitative evaluations of a wide range of end points which serve to determine if there are effects on the primary functional domains of the CNS. The methods are noninvasive and use basic instruments to get quantitative data where possible—an electronic thermometer (to measure rectal/core body temperature), a strain gauge with T-bar animal grip (to measure forelimb grip strength), a sand table (to measure hind limb splay), an activity stage (to measure locomotor activity), and rotarod (to evaluate motor coordination).

Other required equipment usually include a ditcher (for auditory startle), a penlight (for pupil response), and a blunt probe (for various touch-based reflexes). The complete screen should be performed in at least two (and preferably three) time intervals after a single dose of the drug.

There are four broad classes of approaches to any subsequent and more detailed assessment of nervous system effects of drugs in animals.

18.5.1 Isolated Tissue Assays

The classic approach to screening for nervous system effect is a series of isolated tissue preparation bioassays, conducted with appropriate standards, to determine if the material acts pharmacologically directly on neural receptor sites or transmission properties. Though these bioassays are normally performed by a classical pharmacologist, a good technician can be trained to conduct them. The required equipment consists of a Magnus (or similar style) tissue bath (Nodine and Seigler, 1964; Turner, 1965; Offermeier and Ariens, 1966), physiograph or kymograph, force transducer, glassware,

stimulator, and bench spectrophotometer. The assays utilized in the screening battery are listed in Table 18.6, along with the original reference describing each preparation and assay. The assays are performed as per the original author's descriptions with only minor modifications, except that control standards (as listed in Table 18.6) are always used. Only those assays that are appropriate for the neurological/muscular alterations observed in the screen are utilized. Note that all these are intact organ preparations, not minced tissue preparations as others (Bondy, 1979) have recommended for biochemical assays.

The first modification in each assay is that, where available, both positive and negative standard controls (pharmacological agonists and antagonists, respectively) are employed. Before the preparation is utilized to assay the test material, the tissue preparation is exposed to the agonist to ensure that the preparation is functional and to provide a baseline dose-response curve against which the activity of the test material can be quantitatively compared. After the test material has been assayed (if a dose-response curve has been generated), one can determine whether the antagonist will selectively block the activity of the test material. If so, specific activity at that receptor can be considered as established. In this assay sequence, it must be kept in mind that a test material may act to either stimulate or depress activity, and therefore the roles of the standard agonists and antagonists may be reversed.

Commonly overlooked when performing these assays is the possibility of metabolism to an active form that can be assessed in this *in vitro* model. The test material should be tested in both original and "metabolized" forms. The metabolized form is prepared by incubating a 5% solution

(in aerated Tyrode's solution) or other appropriate physiological salt solution with strips of suitably prepared test species liver for 30 min. A filtered supernatant is then collected from this incubation and tested for activity. Suitable metabolic blanks should also be tested.

18.5.2 Electrophysiology Methods

There are a number of electrophysiological techniques available which can be used to detect and/or assess neurotoxicity. These techniques can be divided into two broad general categories: those focused on CNS function and those focused on peripheral nervous system function (Seppäläinen, 1975).

First, however, the function of the individual components of the nervous system, how they are connected together, and how they operate as a complete system should be very briefly overviewed.

Data collection and communication in the nervous system occur by means of graded potentials, action potentials, and synaptic coupling of neurons. These electrical potentials may be recorded and analyzed at two different levels depending on the electrical coupling arrangements: individual cells (i.e., intracellular and extracellular) or multiple cells (e.g., EEG, evoked potentials (EPs), and slow potentials). These potentials may be recorded in specific central or peripheral nervous system areas (e.g., visual cortex, hippocampus, sensory and motor nerves, and muscle spindles) during various behavioral states or in *in vitro* preparations (e.g., nerve-muscle, retinal photoreceptor, and brain slice).

18.5.3 CNS Function: Electroencephalography

The electroencephalogram (EEG) is a dynamic measure reflecting the instantaneous integrated synaptic activity of the CNS, which most probably represents, in coded form, all ongoing processes under higher nervous control. Changes in frequency, amplitude, variability, and pattern of the EEG are thought to be directly related to underlying biochemical changes, which are believed to be directly related to defined aspects of behavior. Therefore, changes in the EEG should be reflected by alterations in behavior and vice versa.

The human EEG is easily recorded and readily quantified, is obtained noninvasively (scalp recording), samples several regions of the brain simultaneously, requires minimal cooperation from the subject, and is minimally influenced by prior testing. Therefore, it is a very useful and recommended clinical test in cases in which exposure to drugs produces symptoms of CNS involvement and in which long-term exposures to high concentrations are suspected of causing CNS damage.

Since the EEG recorded using scalp electrodes is an average of the multiple activity of many small areas of cortical surface beneath the electrodes, it is possible that in situations involving noncortical lesions, the EEG may not

accurately reflect the organic brain damage present. Noncortical lesions following acute or long-term low-level exposures to toxicants are well documented in neurotoxicology (Norton, 1980). The drawback mentioned earlier can be partially overcome by utilizing activation or evocative techniques, such as hyperventilation, photic stimulation, or sleep, which can increase the amount of information gleaned from a standard EEG.

As a research tool, the utility of the EEG lies in the fact that it reflects instantaneous changes in the state of the CNS. The pattern can thus be used to monitor the sleep-wakefulness cycle activation or deactivation of the brain stem and the state of anesthesia during an acute electrophysiological procedure. Another advantage of the EEG, which is shared by all CNS electrophysiological techniques, is that it can assess the differential effects of toxicants (or drugs) on various brain areas or structures. Finally, specific CNS regions (e.g., the hippocampus) have particular patterns of afterdischarge following chemical or electrical stimulation which can be quantitatively examined and utilized as a tool in neurotoxicology.

The EEG does have some disadvantages, or, more correctly, some limitations. It cannot provide information about the effects of toxicants on the integrity of sensory receptors or of sensory or motor pathways. As a corollary, it cannot provide an assessment of the effects of toxicants on sensory system capacities. Finally, the EEG does not provide specific information at the cellular level and therefore lacks the rigor to provide detailed mechanisms of action.

Rats represent an excellent model for this as they are cheap, resist infection during chronic electrode and cannula implantation, and are relatively easy to train so that behavioral assessments can be made concurrently.

Depending on the time of drug exposure, the type of scientific information desired, and the necessity of behavioral correlations, a researcher can perform acute and/or chronic EEG experiments. Limitations of the former are that most drugs that produce general anesthesia modify the pattern of EEG activity and thus can complicate subtle effects of toxicants. However, this limitation can be partially avoided if the effect is robust enough. For sleep-wakefulness studies, it is also essential to monitor and record the electromyogram (EMG). Similarly, specific methods exist for assessing pain (Mogil, 2013).

Excellent reviews of these electrophysiology approaches can be found in Fox et al. (1982) and Takeuchi and Koike (1985).

18.5.4 Neurochemical and Biochemical Assays

Though some very elegant methods are now available to study the biochemistry of the brains and nervous system, none has yet discovered any generalized marker chemicals which will serve as reliable indicators or early warnings of

neurotoxic actions or potential actions. There are, however, some useful methods. Before looking at these, however, one should understand the basic problems involved.

Normal biochemical events surrounding the maintenance and functions of the nervous system center around energy metabolism, biosynthesis of macromolecules, and neurotransmitter synthesis, storage, release, uptake, and degradation. Measurement of these events is complicated by the sequenced nature of the components of the nervous system and the transient and labile nature of the moieties involved. Use of measurements of alternations in these functions as indicators of neurotoxicity is further complicated by our lack of a complete understanding of the normal operation of these systems and by the multitude of day-to-day occurrences (such as diurnal cycle, diet, temperature, age, sex, and endocrine status) which are constantly modulating the baseline system. For detailed discussions of these difficulties, the reader is advised to see Johnson (1975) and Damstra and Bond (1980, 1982).

Added to the spectrum of pharmacologic evaluation of CNS-active drugs for safety reasons has been a requirement to assess the addictive (abuse) potential of new drugs (FDA, 2010); starting with a determination that the drug is CNS active (either due to the target therapeutic activity or determination that there is receptor finding in the CNS), GLP studies are required to determine if there is an addictive/abuse potential in the therapeutic range of doses. These study types are (Kallman, 2015) a dependence/withdrawal study, a drug discrimination study, and a self-administration study. These studies are generally required prior to filing of an NDA.

18.6 RESPIRATORY/PULMONARY SYSTEM

The known effects of drugs from a variety of pharmacologic/therapeutic classes on the respiratory system and worldwide regulatory requirements support the need for conducting respiratory evaluations in safety pharmacology. Pharmaceuticals differ from industrial and environmental chemicals in that the scope of concern for their adverse safety effects on the respiratory system extends both to reversible functional degradations and to effects on the respiratory system's functionality due to systemically distributed agents administered by routes other than direct respiration. This is the realm of the relatively new field of safety pharmacology.

As early as 1964, it became apparent that β -adrenergic blocking agents could lead to bronchoconstriction (and possible death) in asthmatics (McNeill, 1964). Since then, many similar adverse effects have been identified. These known effects of drugs from a variety of pharmacologic/therapeutic classes on the respiratory system are summarized in Tables 18.7, 18.8, and 18.9. Resulting worldwide regulatory

TABLE 18.7 Drugs Known to Cause Pulmonary Disease

| | |
|---------------------------------------|--|
| <i>Chemotherapeutic</i> | <i>Analgesics</i> |
| Cytotoxic | Heroin ^a |
| Azathioprine | Methadone ^a |
| Bleomycin ^a | Noloxone ^a |
| Busulfan | Ethchlorvynol ^a |
| Chlorambucil | Propoxyphene ^a |
| Cyclophosphamide | Salicylates ^a |
| Etoposide | |
| Melphalan | <i>Cardiovascular</i> |
| Mitomycin ^a | Amiodarone ^a |
| Nitrosoureas | Angiotensin-converting enzyme inhibitors |
| Procarbazine | Anticoagulants |
| Vinblastine | Beta-blockers ^a |
| Ifosfamide | Dipyridamole |
| Noncytotoxic | Fibrinolytic agents ^a |
| Methotrexate ^a | Protamine ^a |
| Cytosine arabinoside ^a | Tocainide |
| Bleomycin ^a | |
| Procarbazine ^a | <i>Inhalants</i> |
| | Aspirated oil |
| <i>Antibiotic</i> | Oxygen ^a |
| Amphotericin B ^a | |
| Nitrofurantoin | <i>Intravenous</i> |
| Acute ^a | Blood ^a |
| Chronic | Ethanolamine oleate (sodium morrhuate) ^a |
| | Ethiodized oil (lymphangiogram) |
| Sulfasalazine | Talc |
| Sulfonamides | Fat emulsion |
| Pentamidine | |
| <i>Anti-inflammatory</i> | <i>Miscellaneous</i> |
| Acetylsalicylic acid ^a | Bromocriptine |
| Gold | Dantrolene |
| Methotrexate | Hydrochlorothiazide ^a |
| Nonsteroidal anti-inflammatory agents | |
| Penicillamine ^a | Methysergide |
| | Oral contraceptives |
| <i>Immunosuppressive</i> | Tocolytic agents ^a |
| Cyclosporin | Tricyclics ^a |
| Interleukin-2 ^a | L-Tryptophan |
| | Radiation |
| | Systemic lupus erythematosus (drug induced) ^a |
| | Complement-mediated leukostasis ^a |

Source: Touvay and Le Mosquet (2000); Akoun (1989); Dorato (1994); Lalej-Bennis (2001); Mauderly (1989); McClellan and Henderson (1989); Rosnow (1992).

^aTypically causes acute or subacute respiratory insufficiency.

requirements (Tables 18.10 and 18.11) require the conduct of prescribed respiratory evaluations prior to drug in humans. The objective of such studies is to evaluate the potential for drugs to cause nonintended pharmacologic or toxicologic effects that influence respiratory function. Changes in

TABLE 18.8 Drugs Known to Adversely Affect Respiratory Function

| <i>Drugs known to cause or aggravate bronchospasm</i> | <i>Agents associated with pleural effusion</i> |
|---|---|
| Vinblastine | Chemotherapeutic agents |
| Nitrofurantoin (acute) | Nitrofurantoin (acute) |
| Acetylsalicylic acid | Bromocriptine |
| Nonsteroidal anti-inflammatory agents | Dantrolene |
| Interleukin-2 | Methysergide |
| Beta-blockers | L-Tryptophan |
| Dipyridamole | Drug-inducing systemic lupus erythematosus |
| Protamine | Tocolytics |
| Nebulized pentamidine, beclomethasone, and propellants | Amiodarone |
| Hydrocortisone | Esophageal variceal sclerotherapy agents |
| Cocaine | Interleukin-2 |
| Propafenone | |
| <i>Agents associated with acute-onset pulmonary insufficiency^a</i> | <i>Agents that cause subacute respiratory failure</i> |
| Bleomycin plus O ₂ | Chemotherapeutic agents |
| Mitomycin | Nitrofurantoin (chronic) |
| Bleomycin ^a | Amiodarone |
| Procarbazine ^a | L-Tryptophan |
| Methotrexate ^a | Drug-inducing systemic lupus erythematosus |
| Amphotericin B | |
| Nitrofurantoin (acute) ^b | |
| Acetylsalicylic acid ^b | |
| Interleukin-2 ^b | |
| Heroin and other narcotics ^b | |
| Epinephrine ^b | |
| Ethchlorvynol ^b | |
| Fibrinolytic agents | |
| Protamine | |
| Blood products ^b | |
| Fat emulsion | |
| Hydrochlorothiazide | |
| Complement-mediated leukostasis | |
| Hyskon (Dextran-70) ^b | |
| Tumor necrosis factor ^b | |
| Intrathecal methotrexate | |
| Tricyclic antidepressants ^b | |
| Amiodarone plus O ₂ | |
| Naloxone | |
| Onset at less than 48 h | |

Source: McNeill (1964); Borison (1977); Tattersfield (1986); Illum and Davis (1992); Shao et al. (1992); Shao et al. (1992); Fariba et al. (2002).

^aAssociated with hypersensitivity with eosinophilia.

^bUsually reversible within 48–72 h, implying noncardiac pulmonary edema rather than inflammatory interstitial pneumonitis.

TABLE 18.9 Drugs Known to Influence Ventilatory Control

| Depressants | Stimulants |
|------------------------------------|--------------------------------|
| Inhaled anesthetics | Alkaloids |
| Barbiturates | Nicotine |
| Benzodiazepines | Lobeline |
| Diazepam | Piperidine |
| Temazepam | Xanthine analogs |
| Chlordiazepoxide | Theophylline |
| Serotonin analogs | Caffeine |
| Methoxy(dimethyl)tryptamine | Theobromine |
| Dopamine analogs | Analeptics |
| Apomorphine | Doxapram |
| Adenosine analogs | Salicylates |
| 2-Chloroadenosine | Progesterone analogs |
| R-Phenylisopropyladenosine (R-PIA) | Almitrine |
| N-Ethylcarboxamide (NECA) | Glycine analogs |
| β-Adrenergic antagonists | Strychnine |
| Timolol maleate | GABA antagonists |
| GABA analogs | Picrotoxin |
| Muscimol | Bicuculline |
| Baclofen | Serotonin synthesis inhibitors |
| Opiates | p-Chlorophenylalanine |
| Morphine | Reserpine |
| Codeine | |
| Methadone | |
| Meperidine | |
| Phenazocine | |
| Tranquilizers/analgesics | |
| Chlorpromazine | |
| Hydroxyzine | |
| Rompun (xylazine) | |
| Nalorphine | |

Under ICH (S7A) and FDA guidelines, all new drugs (with limited exception—see Chapter 20) must be evaluated for pharmacologic safety in three core organ systems (the central nervous system, cardiovascular system, and respiratory system). Table 22.11 presents the required determinations under these regulations for mandated respiratory system evaluations.

TABLE 18.10 Required Respiratory System Safety Pharmacology Evaluation

| |
|---|
| Respiratory functions |
| Measurement of rate and relative tidal volume in conscious animals |
| Pulmonary function |
| Measurement of rate, tidal volume, and lung resistance and compliance in anesthetized animals |

respiratory function can result either from alterations in the pumping apparatus that controls the pattern of pulmonary ventilation or from changes in the mechanical properties of the lung that determine the transpulmonary pressure (work) required for lung inflation and deflation.

TABLE 18.11 Regulatory Documents Recommending Respiratory Function Testing in Safety Pharmacology Studies

| | |
|----------------|---|
| United States | FDA Guideline for the Format and Content of the Nonclinical Pharmacology/ Toxicology Section of an Application (Section IID, pp. 12, February, 1987) |
| Japan | Ministry of Health and Welfare Guidelines for Safety Pharmacology Studies Required for the Application for Approval to Manufacture (Import) Drugs. Notification YAKUSHIN-YAKU No. 4, January 1991 |
| Australia | Guidelines for Preparation and Presentation of Applications for Investigational Drugs and Drug Products Under the Clinical Trials Exemption Scheme (STET 12, 15) |
| Canada | RA5 Exhibit 2, Guidelines for Preparing and Filing Drug Submissions (p. 21) |
| United Kingdom | Medicines Act 1968, Guidance Notes on Applications for Product Licenses (MAL 2, p. A3F-1) |

The respiratory system is responsible for generating and regulating the transpulmonary pressures needed to inflate and deflate the lung. Normal gas exchange between the lung and blood requires breathing patterns that ensure appropriate alveolar ventilation. Ventilatory disorders that alter alveolar ventilation are defined as hypoventilation or hyperventilation syndromes. Hyperventilation results in an increase in the partial pressure of arterial CO_2 above normal limits and can lead to acidosis, pulmonary hypertension, congestive heart failure, headache, and disturbed sleep. Hypoventilation results in a decrease in the partial pressure of arterial CO_2 below normal limits and can lead to alkalosis, syncope, epileptic attacks, reduced cardiac output, and muscle weakness.

Normal ventilation requires that the pumping apparatus provide both adequate total pulmonary ventilation (minute volume) and the appropriate depth (tidal volume) and frequency of breathing. The depth and frequency of breathing required for alveolar ventilation are determined primarily by an anatomic dead space of the lung. In general, a rapid shallow breathing pattern (tachypnea) is less efficient than a slower deeper breathing pattern that achieves the same minute volume. Thus, any change in minute volume, tidal volume, or the rate of breathing can influence the efficiency of ventilation (Milic-Emili, 1982). The inspiratory and expiratory phases of individual breathing have rates of airflow and durations that are distinct and independently controlled (Boggs, 1992). Thus, by characterizing changes in the airflow rate and duration of each of these phases, mechanisms responsible for changes in tidal volume or respiratory rate can be identified (Milic-Emili, 1982; Indans, 2002). For example, a decrease in airflow during inspiration (the active phase) is generally indicative of a decrease in respiratory drive, while a

decrease in airflow during expiration (the passive phase) is generally indicative of an obstructive disorder.

Mechanisms of ventilatory disorders can also be characterized as either central or peripheral. Central mechanisms involve the neurologic components of the pumping apparatus that are located in the CNS and include the medullary central pattern generator (CPG) as well as integration centers located in the medulla, pons, hypothalamus, and cortex of the brain that regulate the output of the CPG (Boggs, 1992). The major neurologic inputs from the peripheral nervous system that influence the CPG are the arterial chemoreceptors (Boggs, 1992). Many drugs stimulate or depress ventilation by selective interaction with the CNS (Eldridge and Millhorn, 1981; Mueller, 1982; Keats, 1985) or arterial chemoreceptors (Heymans, 1955; Heymans and Neil, 1958).

Defects in the pumping apparatus are classified as hypo- or hyperventilation syndromes and are best evaluated by examining ventilatory parameters in a conscious animal model. The ventilatory parameters include respiratory rate, tidal volume, minute volume, peak (or mean) inspiratory flow, peak (or mean) expiratory flow, and fractional inspiratory time. Defects in mechanical properties of the lung are classified as obstructive or restrictive disorders and can be evaluated in animal models by performing flow-volume and pressure-volume maneuvers, respectively. The parameters used to detect airway obstruction include peak expiratory flow, forced expiratory flow at 25 and 75% of forced vital capacity, and a timed forced expiratory volume, while the parameters used to detect lung restriction include total lung capacity, inspiratory capacity, functional residual capacity, and compliance. Measurement of dynamic lung resistance and compliance, obtained continuously during tidal breathing, is an alternative method for evaluating obstructive and restrictive disorders, respectively, and is used when the response to drug treatment is expected to be immediate (within minutes postdose). The species used in the safety pharmacology studies are the same as those generally used in the toxicology studies (rats and dogs) since pharmacokinetic and toxicologic/pathologic data are available in these species. These data can be used to help select test measurement intervals and doses and to aid in the interpretation of functional change. The techniques and procedures for measuring respiratory function parameters are well established in guinea pigs, rats, and dogs (Amdur and Mead, 1958; King, 1966; Mauderly, 1974; Diamond and O'Donnell, 1977; Murphy, 1994).

The key questions in safety pharmacology of the respiratory system are as follows:

- Does the substance affect the mechanisms of respiratory control (central or peripheral) leading to hypoventilation (respiratory depression) or hyperventilation (respiratory stimulation)?

- Does the substance act on a component of the respiratory system to induce, for example, bronchospasm, obstruction, and fibrosis?
- Does the substance induce acute effects, or can we expect chronic effects?
- Are the effects observed dose dependent or independent?

The classic approach to measuring respiratory function in laboratory animals is plethysmography. It has two basic governing principles (Palecek, 1969; O'Neil and Raub, 1984; Brown and Miller, 1987; Boggs, 1992):

1. The animal (mice, rat, or dog), anesthetized or not, restrained or not, is placed in a chamber (single or double) with pneumotachographs.
2. The variations of pressure in chamber(s) at the time of the inspiration and the expiration make it possible to obtain the respiratory flow of the animal.

There are three main types of body plethysmographs: constant volume, constant pressure, and pressure volume. The constant-volume body plethysmograph is a sealed box that detects volume change by the measurement of pressure changes inside the box. While inside the plethysmograph, inhalation of room air (from outside the plethysmograph) by the test animal induces an increase in lung volume (chest expansion) and thus an increase in plethysmograph pressure. On the other hand, exhalation to the atmosphere (outside the plethysmograph) induces a decrease in the plethysmograph pressure. The magnitude of lung volume change can be obtained via measurement of the change in plethysmographic pressure and the appropriate calibration factor. The plethysmograph is calibrated by injecting or withdrawing a predetermined change in box pressure. To avoid an adiabatic artifact, the rate of air injection or withdrawal is kept the same as that of chest expansion, indicated by the same dP/dt (change in pressure over time).

The constant-pressure body plethysmograph is a box with a pneumotachograph port built into its wall. This plethysmograph detects volume changes via integration of the flow rate, $\int \Delta \text{Flow}$, which is monitored by the pneumotachograph port. There is an outward flow (air moving from the plethysmograph to the atmosphere) during inspiration and inward flow during expiration. Alternatively, in place of a pneumotachograph, a spirometer can be attached to the constant-pressure plethysmograph to detect volume changes. For the detection of plethysmograph pressure and flow rate, sensitive pressure transducers are usually employed. It is important that the transducer be capable of responding to volume changes in a linear fashion within the volume range studied. The plethysmograph should have negligible leaks, and temperature should not change during all respiratory maneuvers. The plethysmograph should also have linear characteristics

with no hysteresis. Dynamic accuracy requires an adequate frequency response. A fast integrated flow plethysmograph, with a flat amplitude response for sinusoidal inputs up to 240 Hz, has been developed for rats, mice, and guinea pigs (Sinnott, 1981). Similar plethysmographs can also be provided for use with large animals.

A third type of pressure-volume plethysmograph has the mixed characteristics of the two types of body box mentioned earlier. For a constant-pressure plethysmograph, the change in volume at first is associated with gas compression or expansion. This fraction of the volume change can be corrected by electronically adding the plethysmographic pressure change to the volume signal. Therefore, the combined pressure-volume plethysmograph has excellent frequency-response characteristics and a wide range of sensitivities (Leigh and Mead, 1974).

If volume, flow rate, and pressure changes are detected at the same time, several respiratory variables can be derived simultaneously from the raw signals. The whole-body plethysmograph method can then be used to measure most respiratory variables, such as tidal volume and breathing frequency, and minute variables, such as tidal volume, breathing frequency, minute ventilation, compliance, pulmonary resistance, functional residual capacity, pressure-volume characteristics, and maximal expiratory flow-volume curves. Table 18.12 defines the parameters that are typically determined by these methods.

Selection of the proper reference values for the interpretation of findings is essential (Drazen, 1984; American Thoracic Society, 1991).

18.6.1 Design of Respiratory Function Safety Studies

The objective of a safety pharmacology evaluation of the respiratory system is to determine whether the drug has the potential to produce a change in respiratory function. Since a complete evaluation of respiratory function must include both the pumping apparatus and the lung, respiratory function safety studies are best designed to evaluate both of these functional components. The total respiratory system is evaluated first by testing for drug-induced changes in ventilatory patterns of intact conscious animals. This is followed by an evaluation of drug-induced effects on the mechanical properties of the lung in anesthetized/paralyzed animals. Together, these evaluations are used to determine (McNeill, 1964) whether drug-induced changes in the total respiratory system have occurred and (Touway and Le Mosquet, 2000) whether these changes are related to pulmonary or extrapulmonary factors.

The time intervals selected for measuring ventilatory patterns following oral administration of a drug should be based on pharmacokinetic data. The times selected generally include the time to reach peak plasma concentration of the drug (T_{\max}), at least one time before and one after T_{\max} , and

TABLE 18.12 Functional Respiratory Responses to Standard Pharmacologic Agents

| | Theophylline | Pentobarbital | Diazepam | Codeine |
|--------------------------------|---------------------------|---------------------------|---------------------------|----------------------------|
| Parameters | 10 mg kg ⁻¹ PO | 35 mg kg ⁻¹ IP | 35 mg kg ⁻¹ IP | 100 mg kg ⁻¹ IP |
| F (breaths min ⁻¹) | +++ | --- | --- | No change |
| TV (mL) | No change | No change | No change | — |
| Ti (s) | -- | ++ | ++ | + |
| Te (s) | -- | +++ | ++ | — |
| PIF (mL s ⁻¹) | ++ | — | — | — |
| PEF (mL s ⁻¹) | ++ | No change | + | — |

Source: Data from Touvay and Le Mosquet (2000).

+ is an increase and — a decrease; F, respiratory rate; PEF, the pulmonary exhalation rate; PIF, the pulmonary inhalation rate; s, seconds; Tc, exhalation time; Ti, inhalation time or duration; TV, tidal volume.

one time that is approximately 24 h after dosing to evaluate possible delayed effects. If the drug is given as a bolus intravenous (IV) injection, ventilatory parameters are monitored for approximately 5 min predose and continuously for 20–30 min postdose; 1, 2, 4, and 24 h time intervals are also monitored to evaluate possible delayed effects. If administered by inhalation or IV infusion, ventilatory parameters would generally be monitored continuously during the exposure period and at 1, 2, 3, and 24 h time intervals after dosing.

The time interval showing the greatest ventilatory change is selected for evaluating lung mechanics. However, if no ventilatory change occurred, the T_{\max} would be used. If the mechanical properties of the lung need to be evaluated within 30 min after dosing, then dynamic measurements of compliance and resistance are performed. Measurements include a predose baseline and continuous measurements for up to approximately 1 h postdose. If the mechanical properties of the lung need to be measured at 30 min or longer after dosing, then a single time point is selected, and the pressure-volume and flow-volume maneuvers are performed.

Supplemental studies including blood gas analysis, end-tidal CO₂ measurements, or responses to CO₂ gas and NaCN can be conducted to gain after the ventilatory and lung mechanical findings have been evaluated. In general, these would be conducted as separate studies.

18.6.2 Capnography

The measurements of rates, volumes, and capacities provided by plethysmograph measurements have a limited ability to detect and evaluate some ventilatory disorders (Murphy, 1994) that markedly affect blood gases.

Detection of hypo- or hyperventilation syndromes requires measurement of the partial pressure of arterial CO₂ (PaCO₂). In humans and large animal models, this can be accomplished by collecting arterial blood with a catheter or needle and analyzing for PaCO₂ using a blood gas analyzer. In conscious

rodents, however, obtaining arterial blood samples by needle puncture or catheterization during ventilatory measurements is generally not practical. An alternative and noninvasive method for monitoring PaCO₂ is the measurement of peak-expired (end-tidal) CO₂ concentrations. This technique has been successfully used in humans (Nuzzo and Anton, 1986) and recently has been adapted for use in conscious rats (Murphy, 1994). Measuring end-tidal CO₂ in rats requires the use of a nasal mask and a microcapnometer (Columbus Instruments, Columbus, OH) for sampling air from the mask and calculating end-tidal CO₂ concentrations. End-tidal CO₂ values in rats are responsive to ventilatory changes and accurately reflect changes in PaCO₂ (Murphy, 1994).

A noninvasive procedure in conscious rats has been developed for use in helping to distinguish between the central and peripheral nervous system effects of drugs on ventilation. Exposure to CO₂ gas stimulates ventilation primarily through a central mechanism (Borison, 1977). In contrast, a bolus injection of NaCN produces a transient stimulation of ventilation through a mechanism that involves selective stimulation of peripheral chemoreceptors (Heymans and Neil, 1958). Thus, to distinguish central from peripheral nervous system effects, our procedure measures the change in ventilatory response (pretreatment vs. posttreatment) to both a 5 min exposure to 8% CO₂ gas and a bolus IV injection of 300 µg kg⁻¹ of NaCN. In this paradigm, a central depressant (e.g., morphine sulfate) inhibits the CO₂ response and has little effect on the NaCN response.

The species selected for use in safety pharmacology studies should be the same as those used in toxicology studies. The advantages of using these species (rat, dog, or monkey) are that (McNeill, 1964) the pharmacokinetic data generated in these species can be used to define the test measurement intervals and (Touvay and Le Mosquet, 2000) acute toxicity data can be used to select the appropriate high dose. Further, the toxicologic/pathologic findings in these species can be used to help define the mechanism of functional change. The rat is the primary choice since

rats are readily available, and techniques for measuring pulmonary function are well established in this species.

18.7 SECONDARY ORGAN SYSTEM

The kidneys are an important target for toxic effects of drug candidates. It is mandatory to select accurate, clinically relevant parameters in order to be in a position to detect putative nephrotoxic effects during the safety pharmacology program. The glomerular filtration rate appears to be of major interest since it is associated with the definition of acute renal failure. Measurement of the renal blood flow, proteinuria, enzymuria, fractional excretion of sodium, etc. is also highly useful to detect any possible renal impact of a new compound. Although the rat is, by far, the most widely used animal species, there are no specific (clinically relevant) reasons to choose it. Various parameters may vary according to the species, sex, strain, age, and so on. Since in most cases acute renal failure occurs following administration of drugs in patients with preexisting risk factors, it is suggested that sensitized animal models be validated and used (salt depletion, dehydration, coadministration of pharmacologic agents, etc.).

The potential effects of new drugs on the digestive system can be examined in a number of model systems of which intestinal motility in the mouse and gastric emptying in the rat are examples recommended for safety pharmacology evaluation (Gad, 2015b). Intestinal motility, assessed by the transit of carmine dye in the mouse, and gastric motility, assessed by stomach weight in the rat, were examined using a range of clinical drugs or potent pharmacological agents known to affect GI function. Assessment of both models in the guinea pig was also evaluated. Activity was demonstrated with codeine, diazepam, atropine, and CCK-8 (all of which inhibited gastric function). However, neither model gave consistent and reliable results with the remaining reference compounds, namely, metoclopramide, bethanechol, cisapride, deoxycholate, carbachol, and domperidone. In conclusion, this investigation questions the usefulness of simple models of GI transport in the rodent as a means of detecting potential effects of a new drug on the digestive system. This finding should be of concern to the pharmaceutical industry as these simple models are routinely used as part of a regulatory safety pharmacology “package” of studies.

A number of classic assays have been designed to examine the effects of a test article on GI function. GI transit rate is most often measured with a test employing a forced meal of an aqueous suspension of activated charcoal (Janssen and Jageneau, 1957). The test article is given via the appropriate route at a preset time prior to the charcoal meal. For example, a compound intended for use via IV injection would be injected intravenously in mice 30 min prior to delivering a charcoal meal by gavage. The distance traveled from the stomach by the black-colored charcoal meal to a specific

anatomic location within the intestine is measured at a fixed time after this meal, usually 20 or 30 min later. In validating this procedure at the Mason laboratories, we tested the ability of a parasympatholytic agent, IV atropine sulfate, to inhibit GI transit. In a dose-dependent fashion, 30 and 50 mg kg⁻¹ atropine sulfate significantly decreased the distance traveled by the charcoal meal.

Another important safety assay of the GI system is the influence of test article on the formation of ulcers (Shay et al., 1945). After overnight fasting, young rats are given the test article and euthanized 4 or 6 h later. The mucosal surface of the stomach and duodenum is scored for the presence of hyperemia, hemorrhage, and ulcers. The dose-dependent ulcerative properties of NSAIDs are clearly demonstrated in this assay, making it important in the development of other NSAIDs that are not as caustic to the GI mucosa (Cashin et al., 1977; Bramm et al., 1981; Darias et al., 1994; Diadone et al., 1994).

Additional digestive system safety pharmacology tests include effects of test articles on gastric emptying rate and gastric secretion. Gastric emptying rate is measured in rats using a solution of phenol red (or Evans blue) delivered via oral gavage a preset time after administration of the test article (Megens et al., 1991). The dilution of phenol red after 30 min in the rat's stomach is determined colorimetrically at 558 nm in a spectrophotometer. This is compared to a group of control rats that are euthanized immediately after phenol red administration. The influence of test articles on gastric juice secretion is accomplished by ligating the pyloric sphincter under anesthesia in rats following a fasting period (Shay et al., 1945; Graf et al., 1982; Takasuna et al., 1992). Immediately after recovery from anesthesia, each rat is given a preset dose of the test article. The fluid content of the rat's stomach is recovered after a set period of time, usually 4 h. The volume and contents of the stomach are measured to determine the effect of the test article on gastric secretions. Electrolyte concentrations, pH, and protein content of gastric secretions can be measured in this assay (Takasuna et al., 1992).

18.7.1 Gastric Emptying Rate and Gastric pH Changes: A New Model

Sometimes new technologies for safety pharmacology can come from clinical settings. The Heidelberg pH capsule (HC) was developed over 30 years ago at Heidelberg University in West Germany. H.G. Noller invented and first tested this device on over 10 000 adult patients over a 3-year period. The HC is a pill-sized device containing an antimony–silver chloride electrode for measuring pH and a high-frequency transmitter operating at an average frequency of 1.9 MHz. The transmitter in the HC is activated by immersion in physiologic saline by a permeable membrane enclosing the battery compartment. Thus, when a patient swallows the HC, the fluid contents of the stomach activate

the transmitter. Transmitted signals are picked up via a belt receiver and can be displayed and recorded. The profile of changes in pH over time correlates with the movement of the HC through the different regions of the GI tract (Mojaverian et al., 1989). The pH of the fasted human stomach is very acidic, on average about pH 1. When the HC moves through the pyloric sphincter and into the duodenum, there is a rapid increase in pH of over 4 pH units. Thus, one can get a fairly precise measure of gastric emptying rate in humans with this noninvasive technique. Additional pH changes have been correlated with transition of the HC through the duodenum, jejunum, and colon.

Mojaverian and colleagues have used the HC extensively to examine the influence of gender, posture, age, and content and frequency of food ingestion on the gastric emptying rate (or gastric residence time) in healthy volunteers (Mojaverian et al., 1989, 1991). While developed for clinical use in people, the HC may be a useful tool for measuring important digestive system parameters in laboratory animals. The size of the HC, approximately the size of a No. 1 gelatin capsule (7 mm diameter, 20 mm long), prohibits its use in small animals (Mojaverian et al., 1989). It may be useful in studies with dogs and possibly in nonhuman primates. In particular, the HC could be used to measure gastric emptying rate in a totally noninvasive manner in dogs (Itoh et al., 1986). Dogs are readily trainable to accept pills and to wear a receiver belt and could be tested after administration of a test compound (Lui et al., 1986; Vashi and Meyer, 1988). This technique for measuring gastric emptying rate in dogs is also advantageous in that it is not a terminal procedure. The influence of test articles on the pH within different portions of the GI system could also be measured with the HC (Youngberg et al., 1985). The major drawback for using the HC for safety pharmacology screening is the price of the capsules and the receiver system.

18.8 RENAL FUNCTION TESTS

There are three major sets of evaluations of the potential adverse effect of a drug on renal function. These are tests to determine glomerular function, tubular function, and hemodynamic function (Pugsley and Curtis, 2015).

“Strong signals” of a potential problem will generally be detected in clinical chemistry/urinalysis data collected as part of the repeat-dose general toxicology studies conducted prior to IND.

It should be noted that while assessment of renal function is generally a secondary safety pharmacology evaluation to be addressed before filing for an NDA or perhaps initiation of phase III trials, in the case where the intended patient population already has impaired renal function, this evaluation should be considered a primary evaluation to be performed before undertaking initial chemical trials in patients.

18.9 SUMMARY

The initiative to all mandated safety pharmacology studies to the drug development process is overdue in arriving. However, its actual implementation and the use of the resulting data in risk/benefit decision will take some time to be worked out.

REFERENCES

- Aghajafari F, Murphy K, Matthews S, Ohlsson A, Amankwah K, Hannah M. (2002) Repeated doses of antenatal corticosteroids in animals: a systematic review. *Am. J. Obstet. Gynecol.* 186(4):843–849.
- Akoun GM. (1989) Natural history of drug-induced pneumonitis. In: *Drug Induced Disorders, Vol. 3: Treatment Induced Respiratory Disorders* (Akoun GM and White JP, Eds.). Elsevier Scientific Publishers B.V., New York, pp. 3–9.
- Amdur MO, Mead J. (1958) Mechanics of respiration in unanesthetized guinea pigs. *Am. J. Physiol.* 192:364–368.
- American Thoracic Society. (1991) Lung function testing: selection of reference values and interpretative strategies. *Am. Rev. Respir. Dis.* 144:1202–1218.
- Anderson H, Spliid H, Larsen S, Dall V. (2000) Statistical analysis of time to event data from preclinical safety pharmacology studies. *Toxicol. Methods.* 10:111–125.
- Berman CL, Cannon K, Cui Y, Kornbrust DJ, Lagrutta A, Sun SZ, Tepper J, Waldron G, Younis HS. (2014) Recommendations for safety pharmacology evaluations of oligonucleotide-based therapeutics. *Nucleic Acid Ther.* 24(4):291–301.
- Boggs DF. (1992) Comparative control of respiration. In: *Comparative Biology of the Normal Lung* (Parent RA, Ed.), Vol. I. CRC Press, Boca Raton, FL, pp. 309–350.
- Bondy SC. (1979) Rapid screening of neurotoxic agents by in vivo means. In: *Effects of Food and Drugs on the Development and Function of the Nervous System: Methods for Predicting Toxicity* (Gryder RM and Frankos VH, Eds.). Office of Health Affairs, FDA, Washington, DC, pp. 133–143.
- Borison HL. (1977) Central nervous system depressants: control-systems approach to respiratory depression. *Pharmacol. Ther.* B. 3:211–226.
- Bramm E, Binderup L, Arrigoni-Martelli E. (1981) An unusual profile of activity of a new basic anti-inflammatory drug, timegadine. *Agents Actions.* 11:402–409.
- Braunwald E. (2008) Biomarkers in heart failure. *N. Engl. J. Med.* 358:2148–2159.
- Brown LK, Miller A. (1987) Full lung volumes: functional residual capacity, residual volume and total lung capacity. In: *Pulmonary Function Tests: A Guide for the Student and House Officer* (Miller A, Ed.). Grune & Stratton, Inc., New York, pp. 53–58.
- Brunton LL, Lazo JS, Parker KL. (2011) *Goodman & Gilman's The Therapeutic Basis of Therapeutics*, 12th Ed. McGraw Hill, New York.
- Burns BD, Paton WDM. (1951) Depolarization of the motor end-plate by decamethonium and acetylcholine. *J. Physiol. Lond.* 115:41–73.

- Carlsson L, Amos GJ, Andersson B, Drews L, Duker G, Wadstedt G. (1997) Electrophysiological characterization of the prokinetic agents cisapride and mosapride *in vivo* and *in vitro*: implications for proarrhythmic potential? *J. Pharmacol. Exp. Ther.* 282:220–227.
- Cashin CH, Dawson W, Kitchen EA. (1977) The pharmacology of benoxaprofen (2-[4-chlorophenyl]- α -methyl-5-benzoxazole acetic acid), LRCL 3794, a new compound with anti-inflammatory activity apparently unrelated to inhibition of prostaglandin synthesis. *J. Pharm. Pharmacol.* 29:330–336.
- CAST. (1989) Preliminary report: effects of encainide and flecainide on morbidity in a randomized trial of arrhythmia suppression after myocardial infarction. *N. Engl. J. Med.* 321:406–414.
- Castillo JC, De Beer EJ. (1947a) The guinea pig tracheal chain as an assay for histamine agonists. *Fed. Proc.* 6:315.
- Castillo JC, De Beer EJ. (1947b) The tracheal chain. *J. Pharmacol. Exp. Ther.* 90:104.
- Cavero I. (2009) Exploratory safety pharmacology: a new safety paradigm to de-risk drug candidates prior to selection for regulatory science investigation. *Expert Opin. Drug Saf.* 8:627–647.
- Committee for Proprietary Medicinal Products (CPMP) (1997) *CPMP/986/96: Points to Consider: The Assessment of the Potential for QT Interval Prolongation by Non-Cardiovascular Medicinal Products*. European Medicines Agency (EMA), London. Available at: <http://www.fda.gov/ohrms/dockets/ac/03/briefing/pubs/cpmp.pdf> (accessed July 5, 2016).
- Committee for Proprietary Medicinal Products (CPMP). (2000) *Note for Guidance on Safety Pharmacology Studies for Human Pharmaceuticals*. European Medicines Agency (EMA), London. Available at: http://www.ema.europa.eu/docs/en_GB/document_library/Scientific_guideline/2009/09/WC500002831.pdf (accessed May 4, 2016).
- Cox BM, Opheim KE, Teschemach H, Goldstein A. (1975) A peptide-like substance from pituitary that acts like morphine 2. Purification and properties. *Life Sci.* 16:1777–1782.
- Damstra T, Bondy SC. (1980) The current status and future of biochemical assays for neurotoxicity. In: *Experimental and Clinical Neurotoxicology* (Spencer PS and Shaumburg HH, Eds.). Williams and Wilkins, Baltimore, MD, pp. 820–833.
- Damstra T, Bondy SC. (1982) Neurochemical approaches to the deletion of neurotoxicity. In: *Nervous System Toxicology* (Mitchell CL, Ed.). Raven Press, New York, pp. 349–373.
- Darias V, Abdallah SS, Tello ML, Delgado LD, Vega S. (1994) NSAID activity study of 4-phenyl-2-thioxo-benzo[4,5]thieno[2,3-d]pyrimidine derivatives. *Arch. Pharm. (Weinheim)*. 327:779–783.
- Diadone G, Maggio B, Raff D, Plescia S, Bajardi ML, Caruso A, Cutuli VMC, Amico-Roxas M. (1994) Synthesis and pharmacological study of ethyl 1-methyl-5-[2-substituted-4-oxo-3(4H-quinazolinyl)-1H]-pyrazole-4-acetates. *Eur. J. Med. Chem.* 29:707–711.
- Diamond L, O'Donnell M. (1977) Pulmonary mechanics in normal rats. *J. Appl. Physiol.* 43:942–948.
- Domer FR. (1971) *Animal Experiments in Pharmacological Analysis*. Charles C. Thomas, Springfield, IL, p. 98 115, 155, 164, 220.
- Dorato MA. (1994) Toxicological evaluation of intranasal peptide and protein drugs. *Drugs Pharm. Sci.* 62:345–381.
- Dorato MA, McMillian CL, Williams TM. (2015) Safety pharmacology. In: *Principles and Methods of Toxicology* (Hayes AW and Kruger CL, Eds.), 6th Ed. CRC Press, Boca Raton, FL, pp. 361–363.
- Drazen JM. (1984) Physiological basis and interpretation of indices of pulmonary mechanics. *Environ. Health Perspect.* 56:3–9.
- Eldridge FL, Millhorn DE. (1981) Central regulation of respiration by endogenous neurotransmitter and neuromodulators. *Annu. Rev. Physiol.* 3:121–135.
- FDA. (2010) *Guidance for Industry on Assessment of Abuse Potential of Drugs*. Draft. Center for Drug Evaluation and Research (CDER), Silver Spring, MD.
- Fenichel RP, Koerner J. (1999) *Development of Drugs that Alter Ventricular Repolarization*. Internet-based draft. Available at: <http://www.fenichel.net/> (accessed June 1, 2016).
- Folke S. (2000) Drug safety in relation to efficacy: the view of a clinical pharmacologist. *Pharmacol. Toxicol.* 86:30–32.
- Fox DA, Lowndes HE, Bierkamper GG. (1982) Electrophysiological techniques in neurotoxicology. In: *Nervous System Toxicology* (Mitchell CL, Ed.). Raven Press, New York, pp. 299–336.
- Fujimori K. (1999) The role of general pharmacological studies and pharmacokinetics in the evolution of drugs (1): the role of general/safety pharmacology studies in the development of pharmaceuticals: international harmonization guidelines. *Folia Pharmacol. Jpn.* 13:31–39.
- Gad SC. (1982) A neuromuscular screen for use in industrial toxicology. *J. Toxicol. Environ. Health.* 9:691–704.
- Gad SC. (2008) Nonclinical safety assessment of the cardiovascular toxicity of drugs and combination medical devices. In: *Target Organ Toxicity: The Cardiovascular System* (Acosta D, Ed.). Taylor & Francis, Philadelphia, PA.
- Gad SC. (2012) *Safety Pharmacology in Pharmaceutical Development: Approval and Post Marketing Surveillance*, 2nd Ed. CRC Press, Boca Raton, FL.
- Gad SC. (2015a) Cardiovascular toxicology and its evaluation. In: *Mammalian Toxicology* (Abou-Donia MB, Ed.). John Wiley & Sons, Inc., Hoboken, NJ, pp. 425–452.
- Gad SC. (2015b) Gastrointestinal toxicology. In: *Mammalian Toxicology* (Abou-Donia MB, Ed.). John Wiley & Sons, Inc., Hoboken, NJ, pp. 539–568.
- Gad SC and Gad SE. (2003) A functional observational battery for use in canine toxicity studies: development and validation. *Int. J. Toxicol.* 22(6):415–422.
- Gauvin DV, Tilley LP, Smith FWK, Baird TJ. (2006) Electrocardiogram, hemodynamics, and core body temperatures of the normal freely moving laboratory beagle dog by remote telemetry. *J. Pharmacol. Toxicol. Methods.* 53:128–139.
- Gecse A, Zsilinsky E, Szekeres L. (1976) Bradykinin antagonism. In: *Kinins; Pharmacodynamics and Biological Roles* (Sicuteri F, Back N and Haberland G, Eds.). Plenum Press, New York, pp. 5–13.
- Graf E, Sieck A, Wenzl H, Winkelmann J. (1982) Animal experiments on the safety pharmacology of lofexidine. *Arzneim.-Forsch./Drug Res.* 32(II)(8a):931–940.
- Green MD. (1997) Problems associated with the absence of activity in standard models of safety pharmacology used to assess biological products. *Int. J. Toxicol.* 16:33–40.

- Haggerty GC. (1991) Strategies for and experience with neurotoxicity testing of new pharmaceuticals. *J. Am. Coll. Toxicol.* 10:677–687.
- Hamlin RL. (2008) How many ECG leads are required for in vivo safety pharmacology? *J. Pharmacol. Toxicol. Methods.* 57:161–168.
- Heymans C. (1955) Action of drugs on carotid body and sinus. *Pharmacol. Rev.* 7:119–142.
- Heymans C, Neil E. (1958) The effects of drugs on chemoreceptors. In: *Reflexogenic Areas of the Cardiovascular Systems* (Heymans C and Neil E, Eds.). Churchill, Ltd, London, pp. 192–199.
- Hite M. (1997) Safety pharmacology approaches. *Int. J. Toxicol.* 16:23–31.
- Hohnloser SH, Klingenhoben T, Singh BN. (1994) Amiodarone-associated proarrhythmic effects. *Ann. Intern. Med.* 121(7):529–535.
- ICH. (2000) *S7A: Safety Pharmacology Studies for Human Pharmaceuticals*. International Conference on Harmonisation, Geneva, Switzerland.
- ICH. (2002) *S7B: Safety Pharmacology Studies for Assessing the Potential for Delayed Ventricular Repolarization (QT Interval Prolongation) by Human Pharmaceuticals*. Draft Consensus Guideline, International Conference on Harmonisation, Geneva, Switzerland.
- ICH. (2005) *S7B: The Non-Clinical Evaluation of the Potential for Delayed Ventricular Prolongation by Human Pharmaceuticals*. International Conference on Harmonisation, Geneva, Switzerland.
- ICH. (2011) *S6(R1): Preclinical Safety Evaluation of Biotechnology-Derived Pharmaceuticals*. International Conference on Harmonisation, Geneva, Switzerland.
- Illum L, Davis SS. (1992) Intranasal insulin. Clinical pharmacokinetics. *Clin. Pharmacokinet.* 23:30–41.
- Indans I. (2002) Non-lethal end-points in inhalation toxicology. *Toxicol. Lett.* 135(1):53.
- Irwin S. (1968) Comprehensive observational assessment: Ia. A systematic, quantitative procedure for assessing the behavioral and physiologic state of the mouse. *Psychopharmacologia (Berlin)*. 13:222–257.
- Itoh T, Higuchi T, Gardner CR, Caldwell L. (1986) Effect of particle size and food on gastric residence time of nondisintegrating solids in beagle dogs. *J. Pharm. Pharmacol.* 38:801–806.
- Janssen PAJ, Jageneau AH. (1957) A new series of potent analgesics: *dextro* 2:2-diphenyl-3-methyl-4-morpholino-butyrylpyrrolidine and related amides. *J. Pharm. Pharmacol.* 9:381–400.
- Johnson MK. (1975) The delayed neuropathy caused by some organophosphorus esters: mechanism & challenge. *Crit. Rev. Toxicol.* 3:289–316.
- Kallman MJ. (2015) Preclinical abuse potential assessment. In: *Principles of Safety Pharmacology* (Pugsley MK and Curtis MJ, Eds.). Springer, Berlin, pp. 115–130.
- Keats AS. (1985) The effects of drugs on respiration in man. *Annu Rev. Pharmacol. Toxicol.* 25:41–65.
- Khayyal MT, Tolba NM, El-Hawary MB, El-Wahed SA. (1974) A sensitive method for the bioassay of acetylcholine. *Eur. J. Pharmacol.* 25:287–290.
- King TKC. (1966) Measurement of functional residual capacity in the rat. *J. Appl. Physiol.* 21:233–236.
- Kinter LB, Gossett KA, Kerns WD. (1994) Status of safety pharmacology in the pharmaceutical industry. *Drug Dev. Res.* 32:208–216.
- Lacroix P. (2002) Electro-cardiographic toxicity in the guinea pig. In: *Current Protocols of Toxicology* (Enna SJ and Williams M, Eds.). John Wiley & Sons, Inc., Hoboken, NJ, pp. 5.29.1–5.29.7.
- Lacroix P, Provost D. (2000) Basic safety pharmacology: the cardiovascular system. *Therapie.* 55:63–69.
- Lalej-Bennis D. (2001) Six month administration of gelified intranasal insulin in 16 type 1 diabetic patients under multiple injections: efficacy vs subcutaneous injections and local tolerance. *Diabetes Metab.* 27(3):372–377.
- Leach GDH. (1956) Estimation of drug antagonisms in the isolated guinea pig vas deferens. *J. Pharm. Pharmacol.* 8:501.
- Leigh DE, Mead J. (1974) *Principles of Body Plethysmography*. National Heart and Lung Institute, NIH, Bethesda, MD.
- Levy B, Tozzi S. (1963) The adrenergic receptive mechanism of the rat uterus. *J. Pharmacol. Exp. Ther.* 142:178.
- Lin RCY, Yeoh TS. (1965) An improvement of Vane's stomach strip preparation for the assay of 5-hydroxy-tryptamine. *J. Pharm. Pharmacol.* 17:524–525.
- Lui CY, Amidon GL, Berardi RR, Fleisher D, Youngberg CA, Dressman JB. (1986) Comparison of gastrointestinal pH in dogs and humans: implications on the use of the beagle dog as a model for oral absorption in humans. *J. Pharm. Sci.* 75:271–274.
- Malik M, Camm AJ. (2001) Evaluation of drug-induced QT interval prolongation: implications for drug approval and labeling. *Drug Saf.* 24:323–351.
- Markert M, Klumpp A, Trautmann T, Mayer K, Stubhan M, Guth B. (2007) The value added by measuring myocardial contractility 'in vivo' in safety pharmacological profiling of drug candidates. *J. Pharmacol. Toxicol. Methods.* 56:203–211.
- Martin LI, Horvath CJ, Wyand MS. (1997) Safety pharmacology screening: practical problems in drug development. *Int. J. Toxicol.* 16:41–65.
- Matsuzawa T, Hashimoto M, Nara H, Yoshida M, Tamura S, Igarashi T. (1997) Current status of conducting function tests in repeated dose toxicity studies in Japan. *J. Toxicol. Sci.* 22:374–382.
- Mattsson JL, Spencer PJ, Albee RR. (1996) A performance standard for clinical and functional observational battery examination of rats. *J. Am. Coll. Toxicol.* 15:239.
- Mauderly JL. (1974) The influence of sex and age on the pulmonary function of the beagle dog. *J. Gerontol.* 29:282–289.
- Mauderly JL. (1989) Effects of inhaled toxicants on pulmonary function. In: *Concepts in Inhalation Toxicology* (McClellan RO and Henderson RF, Eds.). Hemisphere Publishing Corp., New York, pp. 347–401.
- McClellan RO, Henderson RF. (1989) *Concepts in Inhalation Toxicology*. Hemisphere Publishing Corp., New York.
- McNeill RS. (1964) Effect of β -adrenergic blocking agent, propranolol, on asthmatics. *Lancet.* 21:1101–1102.

- Megens AAHP, Awouters FHL, Niemegeers CJE. (1991) General pharmacology of the four gastrointestinal motility stimulants bethanechol, metoclopramide, trimebutine, and cisapride. *Arzneim.-Forsch./Drug Res.* 41(I)(6):631–634.
- Milano S. (2012) The minipig in safety pharmacology. In: *The Minipig in Biomedical Research* (McAnulty PA, Dazan AD, Ganderup NC and Hastings KL, Eds.). CRC Press, Boca Raton, FL, pp. 159–181.
- Milic-Emili J. (1982) Recent advances in clinical assessment of control of breathing. *Lung.* 160:1–17.
- Mogil JS. (2013) Preclinical Pain Testing: Overview, Alternatives and Troubleshooting, McGill Department of Psychology and Centre for Research on Pain, Presented at the Society of Toxicology (SOT) 2013 Annual Meeting and ToxExpo, March 10–14, San Antonio, TX.
- Mohammad S, Zhou Z, Gong Q, January CT. (1997) Blockage of the HERG human cardiac K⁺ channel by gastrointestinal prokinetic agent cisapride. *Am. J. Physiol.* 273:H2534–H2538.
- Mojaverian P, Chan K, Desai A, Vivian J. (1989) Gastrointestinal transit of a solid indigestible capsule as measured by radiotelemetry and dual gamma scintigraphy. *Pharm. Res.* 6:717–722.
- Mojaverian P, Reynolds JC, Ouyang A, Wirth F, Kellner PE, Vlasses PH. (1991) Mechanism of gastric emptying of a nondisintegrating radiotelemetry capsule in man. *Pharm. Res.* 8:97–100.
- Moscardo E, Maurin A, Dorigatti R, Champeraux P, Richard S. (2007) An optimized methodology for the neurobehavioral assessment in rodents. *J. Pharmacol. Toxicol. Methods.* 56:239–255.
- Mueller RA. (1982) The neuropharmacology of respiratory control. *Pharmacol. Rev.* 34:255–285.
- Murphy DJ. (1994) Safety pharmacology of the respiratory system: techniques and study design. *Drug Dev. Res.* 32:237–246.
- Murphy DJ, Joran ME, Grando JC. (1995) A non-invasive method for distinguishing central from peripheral nervous system effects of respiratory depressant drugs in conscious rats. *Gen. Pharmacol.* 26(3):569–575.
- Nodine JH, Siger PE. (1964) *Animal and Clinical Pharmacologic Techniques in Drug Evaluation*. Year Book Medical Publishers, Chicago, IL, pp. 36–38.
- Norton S. (1980) Toxic responses of the central nervous system. In: *Toxicology: The Basic Science of Poisons* (Doull J, Klaassen CD and Amdur MO, Eds.), 2nd Ed. Macmillan, Inc., New York.
- Nuzzo PF, Anton WR. (1986) Practical applications of capnography. *Respir. Ther.* 16:12–17.
- O’Neil JJ, Raub JA. (1984) Pulmonary function testing in small laboratory mammals. *Environ. Health Perspect.* 53:11–22.
- Offermeier J, Ariens EJ. (1966) Serotonin I. Receptors involved in its action. *Arch. Int. Pharmacodyn. Ther.* 64:92–215.
- Olejniczak K. (1999) Development of a safety pharmacology guideline. *Hum. Exp. Toxicol.* 18:502.
- Palecek F. (1969) Measurement of ventilator mechanics in the rat. *J. Appl. Physiol.* 27:149–156.
- Pugsley MK, Curtis MJ. (Eds.) (2015) Principles of safety pharmacology. In: *Handbook of Experimental Pharmacology*, Vol. 229. Springer-Verlag, Berlin.
- Pugsley MK, Authier S, Curtis MJ. (2008) Principles of safety pharmacology. *Br. J. Pharmacol.* 154(2):1382–1399.
- Redfern WS, Wakefield ID, Prior H, Pollard CE, Hammond TG, Valenlin JP. (2002) Safety pharmacology—a progressive approach. *Fundam. Clin. Pharmacol.* 16:161–173.
- Redfern WS, Lorna CE, Lainée P, Pinches M, Robinson S, Valentin J-P. (2013) Functional assessments in repeat-dose toxicity studies: the art of the possible. *Toxicol. Res.* 2:209–234.
- Rosnow EC. (1992) Drug-induced pulmonary disease: an update. *Chest.* 102:239–250.
- van Rossum JM. (1965) Different types of sympathomimetic β -receptors. *J. Pharm. Pharmacol.* 17:202.
- Seeman P, Weinstein J. (1966) Erythrocyte membrane stabilization by tranquilizers and antihistamines. *Biochem. Pharmacol.* 15:1737–1752.
- Seppalainen AM. (1975) Applications of neurophysiological methods in occupational medicine: a review. *Scand. J. Work Environ. Health.* 1:1–14.
- Shao Z, Krishnamoorthy R, Mitra AK. (1992) Cyclodextrins as nasal absorption promoters of insulin-mechanistic evaluations. *Pharm. Res.* 9(9):1157–1163.
- Shay H, Komarox SA, Fels SS, Meranze D, Gruenstein M, Siplet H. (1945) A simple method for the uniform production of gastric ulceration in the rat. *Gastroenterology.* 5:43–61.
- Sinnet EE. (1981) Fast integrated flow plethysmograph for small mammals. *J. Appl. Physiol.* 50:1104–1110.
- Sjoquist F. (2000) Drug-related hospital admissions. *Ann. Pharmacol.* 34:832–839.
- Solovien MV, Hamlin RL, Barrett RM, Chengelis CP, Schaefer GJ. (2006) Different species require different correction factors for the QT interval. *Cardiovasc. Toxicol.* 6:145–157.
- Sullivan AT, Kinter RB. (1995) Status of safety pharmacology in the pharmaceutical industry. *Drug. Dev. Res.* 35:166–172.
- Takasuna K, Kasai Y, Usui C, Takahashi M, Hirohashi M, Tamura K, Takayama S. (1992) General pharmacology of the new quinolone anti-bacterial agent levofloxacin. *Arzneim.-Forsch./Drug Res.* 42(I)(3a):408–418.
- Takeuchi Y, Koike Y. (1985) Electrophysiological methods for the in vivo assessment of neurotoxicology. In: *Neurotoxicology* (Blum K and Manzo L, Eds.). Marcel Dekker, Inc., New York, pp. 613–629.
- Tattersfield AE. (1986) Beta adrenoreceptor antagonists and respiratory disease. *J. Cardiovasc. Pharmacol.* 8(Suppl.4):535–539.
- Touvy C, Le Mosquet B. (2000) Systeme respiratoire et pharmacologie de securite. *Therapie.* 55:71–83.
- Turner RA. (1965) *Screening Methods in Pharmacology*, Vol. I and II. Academic, New York, pp. 42–47 60–68, 27–128.
- Vargas HM, Bass AS, Breidenbach A, Feldman HS, Gintant GA, Harmer AR, Heath B, Hoffmann P, Lagrutta A, Leishman D, McMahon N, Mittelstadt S, Polonchuk L, Pugsley MK, Salata JJ, Valentin J-P. (2008) Scientific review and recommendations on preclinical cardiovascular safety evaluation of biologics. *J. Pharmacol. Toxicol. Methods.* 58:72–76.
- Vashi VI, Meyer MC. (1988) Effect of pH on the in vitro dissolution and in vivo absorption of controlled-release theophylline in dogs. *J. Pharm. Sci.* 77:760–764.

- Youngberg CA, Wlodyga J, Schmaltz S, Dressman JB. (1985) Radio-telemetric determination of gastrointestinal pH in four healthy beagles. *Am. J. Vet. Res.* 46:1516–1521.
- Zbinden G. (1966) The significance of pharmacologic screening tests in the preclinical safety evaluation of new drugs. *J. New Drugs.* 6:1–7.
- Zbinden, G. (1984) *Neglect of function and obsession with structure in toxicity testing*. Proceedings of the 9th International Congress of Pharmacology, Vol. 1, July 30, 1984, London. Macmillan, New York, pp. 43–49.

SPECIAL CONCERNS FOR THE PRECLINICAL EVALUATION OF BIOTECHNOLOGY PRODUCTS

Although many assume that biotechnology is a recent concept, the application of this science has been understood for many years and utilized in its simplest form in the fermentation of beer, wine, and bread by microbial agents. Modern biotechnology can be divided into three research and development areas: recombinant DNA (rDNA) technology, monoclonal antibody (mAb) technology, and bioprocess technology (Mackett, 1993; Malinowski, 1999). The commercialization of these three processes is based on the premise that biotechnology can cost-effectively produce large quantities of a highly purified product. Closely associated and also addressed in this chapter are vaccines and gene therapy products.

Biotechnology as a promising source of new and more efficient source of more targeted therapeutics has been with us since the mid-1980s (Maulik and Patel, 1997). While (as one should expect) some of the early promises of efficacy and safety have not quite been met, biotechnology has turned out to be a valuable source of new and valuable therapeutics and currently accounts for one-third of all new therapeutics entering the marketplace. Table 19.1 lists the proteins approved in 2007 alone that have more than a billion dollars a year in global sales. While the pattern of a consistently increasing pattern of new therapeutics being biotech products has seemingly stabilized, these still are a clear wave of the future.

Protein and other biotechnology-derived therapeutics have some fundamental differences from traditional small (synthetic organic) molecules and so require modified or different approaches to characterize their toxicity and evaluate their safety. Table 19.2 presents a comparative summary of these differences as now practiced.

What has become clear in recent years, however, is that while almost all safety concerns for biologic therapeutics are over- or undesired expressions of their intended pharmacology (“on-target hits” as opposed to off-receptor target hits), such

effects can be both distal to the intended therapeutic target tissue and may have significant “lag” periods (perhaps even years) before they are seen in patients (Giezen et al., 2008). A prime example would be progressive multifocal leukoencephalopathy (PML), a brain infection primarily related to immunosuppressive drugs for MS. Herein lies a challenge for nonclinical safety assessment that has not yet been addressed.

The decades of ongoing profits (past the time of patent expiry) have set these molecules apart from their small-molecule cousins, but the reality of “biosimilar” competition seems to finally be arriving (though as of mid 2016, only three such has actually been approved). Mid 2016 also brought the news that nine out of ten biotechnology drugs fail to show efficacy in clinical trials (Weisman, 2016)—though this is not significantly worse than for new small molecules.

Among all the other aspects of increasing understanding of what is involved in the evaluation and development of biologically derived therapeutics has been a continuing evolution of what is needed to evaluate the safety of these products. In the beginning, there was a stark duality of expectations. On one side, early advocates of biotechnology held that there were unlikely to be any safety concerns other than those due to hyperpharmacology overactivity at the target receptors (Vallbracht et al., 1982; Weissinger, 1989; Thomas and Myers, 1998) and contamination (such as occurred with the Cutter product early on in the early history of the polio vaccine; Offit, 2005). On the other hand, there were those that cautioned against the possibility of extreme and unforeseeable toxicities (such as occurred with TGN-412). The truth, as is usually the case, has turned out to be in between. We have come to understand that the toxicity of protein moieties primarily arises from either overexpression of their desired therapeutic effects (i.e., largely disordering of the immune system) such as seen with interferons (IFNs) and interleukins (Fent

TABLE 19.1 Blockbuster Biotech Approvals (2007)

| Company | Product | Indication | 2007 Approval | Status | Peak Sales Potential (\$ Billions) |
|------------------|-------------------------------|---|---------------|------------------|------------------------------------|
| Cephalon | Nuvigil (armodafinil) | Daytime sleepiness | Approved | Approved June 18 | \$1–2 |
| Onyx/Bayer | Nexavar (sorafenib) | Hepatocellular carcinoma | Likely | PDUFA Dec. 28 | \$1+ |
| ImClone | Erbitux (cetuximab) | Refractory colorectal cancer (survival claim) | Approved | Approved Oct. 2 | \$1 |
| Genzyme | Renvela (sevelamer carbonate) | Serum phosphorus control in dialysis patients with chronic kidney disease | Approved | Approved Oct. 21 | \$1 |
| Alexion | Soliris (eculizumab) | Paroxysmal nocturnal hemoglobinuria | Approved | Approved Mar. 19 | \$0.5–1+ |
| Speedel/Novartis | Tektura | Hypertension | Approved | Approved Mar. 6 | \$0.5–1 |

TABLE 19.2 Comparison of Protein Therapeutic Agents with Small-Molecule Drugs

| Parameter | Proteins | Small molecules |
|-----------------------|--|--|
| Drug substance | Heterogeneous mixture Broad specifications during development Specifications may change during development | Single entity; high chemical purity Exception: racemic mixtures Specifications well defined early in development |
| Drug product | Usually intravenously or subcutaneously | Generally oral; few formulations during development |
| Impurities | Difficult to standardize | Purity standards well established |
| Bridging requirements | Significant for drug substance | Bioequivalence procedures |
| Biological activity | May mimic naturally occurring molecules Primary mechanism of toxicity Predictive based on mechanism | Less predictive |
| Nonspecificity | Variable significance | Usually significant Drug–drug interactions |
| Chronic toxicity | Lack of models because of species-determined biological specificity and antigenicity | Models sometimes relevant |
| Impurities | Toxicity not a major issue May impact immunogenicity | May be significant Purity standards well established |

and Zbinden, 1997) or due to “off-target effects” that is at tissues other than the intended target(s) (Cavagnaro, 2008). Additionally, there are cases of immune responses to therapeutic and of antibody neutralization products (Vallbracht et al., 1982; Weissinger, 1989).

Additionally, as “biosimilar” products enter the marketplace (still a slow process in the United States with only one approval as of late 2015, but undoubtedly an increasing occurrence due to cost considerations), there are new concerns raised by doctors and patients (Alkon, 2015).

The principal purposes of preclinical safety evaluation in this context remain:

- To detect harmful (toxic) effects
- To exclude other potentially harmful effects
- To determine their relationship to dose and duration of treatment
- If possible to discover their mechanism or at least pathogenesis

The information from (a) to (d) should be used to predict possible adverse actions in the target species, in order to:

- Warn clinicians about unacceptable risks
- Warn clinicians about risks that should be monitored
- Remind themselves and others of the possibility of toxic effects not detected because they could not be displayed by the test systems used (e.g., headache in a nonhuman species or carcinogenicity in a 1-month experiment) or were not sought

In addition, the toxicologist, as a general biological scientist, should always be alert to physiological and pharmacological effects manifested in her or his experiments, because they may illuminate mechanisms of health and disease both of academic and practical importance (Folb, 2006; Cavagnaro, 2008; Ellis et al., 2013). Table 19.3 points out the common differences between current small- and large-molecule safety assessment program.

The objectives of the preclinical safety studies on biologically derived therapeutics are to identify the pharmacodynamic (PD) and toxicodynamic actions (and comparative response curves) that are likely to be encountered throughout the course of clinical development and use. The selection

TABLE 19.3 Differences between Chronic Use Nonclinical Safety Assessment Plans for Large and Small Molecules

| Biopharmaceuticals | Small-Molecule Pharmaceuticals |
|--|--|
| "Case-by-case" approach to nonclinical safety evaluation | Similar nonclinical safety studies for most products |
| Pharmacology used to select species | Metabolism used to select species |
| Single species common for repeat-dose studies beyond 28 days | Rodent and nonrodent typical |
| Dosing sequences may not be done | Dosing daily unless justified |
| Immunogenicity important | Generally not immunogenic |
| Genetic toxicology not required | Genetic toxicology required |
| Metabolism studies not appropriate | Metabolism studies required |
| Longest repeat dose typically 6 months nonrodent | Nonrodent may be up to 12 months |
| Two-year carcinogenicity rare | Two-year carcinogenicity required |
| Developmental toxicology in relevant species may be required | Developmental toxicology in two species required |

TABLE 19.4 Classification of Bioengineered Products on Practical Grounds

| Type | Bioengineering Involvement | Pharmacological Properties | Physicochemical Characterization | Example |
|--|---|-----------------------------------|----------------------------------|--|
| 1. Low molecular weight substance | New production route | Well known | Rigorous | Amino acid 6-APA |
| 2. High molecular weight substance | New production route | Fairly well known | Extensive | Human hormones, for example, hGH, hPTH |
| 3a. Endogenous high molecular weight substance | First ever production | Some knowledge | Moderate | IFN |
| 3b. Endogenous high molecular weight substance | First ever production perhaps gene splicing to make hybrid molecule | Scanty to limited knowledge | Moderate | Other lymphokines, tumor necrosis factor, etc. |
| 4. Engineered antigen | Partly or totally synthetic ag + rDNA production | Probably predictable | Rigorous | Synthetic vaccine for poliomyelitis or hepatitis B |
| 5. Monoclonal antibody (or component) | Hybridoma human cell line | Probably predictable antigenicity | Moderate | Antitumor antibody for imaging anti-idiotypic antibody as vaccine |
| 6. Living organism | Removal of pathogenicity by genetic manipulation | Uncertain | Limited | As vaccine immunogen, for example, <i>Salmonella typhimurium</i> TY21a, modified herpes or to carry antigen, for example, vaccinia |

and design of such studies should first consider what may be known about other products which are structurally and/or pharmacologically similar. The nonclinical program and study designs should then proceed to consider:

1. Intended manner of use, including dose, route of administration, and particulars of dosing regimen
2. Age and physiologic condition of intended patient population
3. Selection of relevant model species
4. Stability of the formulated drug substance under the conditions of use
5. Physiologic (disease) state of intended patients

In the area of bioengineered products, many of which are complex proteins of potent but sparsely studied activities in living systems, the investigative responsibilities of the

toxicologist are likely to be very important, because the toxicologist may be the first observer able to study the effects of repeated administration of a range of doses on a living system (Griffiths and Lumley, 1998). It is now possible to frame a classification of the types of biologically derived therapeutic products (Table 19.4).

Each type of product has some specific considerations. The range of materials is enormous. The deciding factors for the toxicologist should be the precision with which the material can be characterized by physicochemical or other means, as that should be inversely related to the burden of repeated biological testing necessary to assure safety, and the extent of prior knowledge of its biological properties and what tissues they are expressed in. The greater our ignorance of the latter, the more far reaching should be the toxicologist's studies in order to discern the biological (PD) properties of the substance. Exposure of the individual must also be considered, as different criteria may apply to

deliberate administration of a living organism, which could spread in the community. Another challenge is to understand and then interpret the specific risk–benefit considerations.

As in any safety evaluation, the planned work should be related to the intended use and treatment of the target species (almost always humans), for example, one dose in a few gravely ill patients or multiple doses of the entire healthy community as prophylaxis against a trivial condition. Contrast, say, what might be appropriate for tumor necrosis factor, as in an experimental trial in a few sufferers from late-stage cancer, with the requirements for a candidate vaccine against dental caries to be widely administered to healthy children.

Because of the rapid development of new biotechnology products, toxicology and safety assessment departments at pharmaceutical companies are presently or soon will be confronted with the development of testing protocols for the assessment of safety evaluation for long-term clinical use of rDNA products. Routine toxicology assessment as performed in the past using standard protocols may not apply and may in fact represent unnecessary or inappropriate studies. Because of the relatively nontoxic nature and species specificity of many of the new biotechnology products, less evaluation in rodent species may be required than for some of the chemicals of the past. What is needed in dealing with the products of these new technologies is the rethinking of traditional toxicology testing approaches.

19.1 REGULATION

The regulation of biologically derived therapeutics actually has a long history and also continues to evolve (see Table 19.5). This history led to the PHS Act providing a somewhat mixed description of the products under its authority, which in turn serves to define biologics for CBER: “[A biologic is] any virus, therapeutic serum, toxin, antitoxin, vaccine, blood, blood component or derivative, allergenic product, or analogous product, or arsphenamine or its derivative (or any other trivalent organic arsenic compound), applicable to the prevention, treatment, or cure of disease, or injuries in man....” 21 CFR 600.3 states that “(h) Biological product means any virus, therapeutic serum, toxin, antitoxin, or analogous product applicable to the prevention, treatment or cure of diseases or injuries in man.” Confusion of authority and responsibility between the three human health product centers of the Food and Drug Administration (FDA) (CDER, CBER, and CDRH) led in 1992 to the promulgation of three intercenter agreements. The agreement between CDER and CBER states that the following biological products require licensure and come under CBER’s authority (Mathieu, 1997):

- Vaccines, regardless of the method of manufacture (vaccines were defined as agents administered for the

TABLE 19.5 Historical Perspectives of Biologic Therapeutics Regulation

| | |
|------|---|
| 1902 | Federal Virus, Serum and Toxin Act/PHS (after tetanus contaminated diphtheria antitoxin led to deaths of 10 children). Intent was to ensure safety, purity, and potency |
| 1906 | Pure Food and Drug Act (Upton Sinclair’s <i>The Jungle</i>) |
| 1937 | Division of Biologics Control/NIH |
| 1938 | FD&C Act (sulfanilamide elixir), biologics exempt! |
| 1955 | Division of Biological Standards Established (poliovirus) |
| 1962 | FD&C Amendments (thalidomide) |
| 1972 | Bureau of Biologics/FDA |
| 1978 | Good Manufacturing Practices Regulations |
| 1979 | Good Laboratory Practices Regulations |
| 1982 | Bureau of Biologics merged with Bureau of Drugs ⇒ National Center for Drugs and Biologics |
| 1983 | Orphan Drug Act |
| 1984 | Drug Price Competition and Patent Term Restoration Act |
| 1992 | User fees |
| 1997 | ICH harmonized guidelines |
| 2002 | Responsibility for regulation of biological substances divided between CDER and Center for Biologics Evaluation and Research (CBER). FDA CBER |

purpose of eliciting an antigen-specific cellular or humoral response) (CBER, 1997)

- *In vivo* diagnostic allergenic products and allergens intended for use as “hyposensitization agents”
- Human blood or human blood-derived products, including placental blood-derived products, animal-derived procoagulant products and animal- or cell culture-derived hemoglobin-based products intended to act as red blood cell substitutes
- Immunoglobulin products
- Products composed of, or intended to contain, intact cells or intact microorganisms
- Proteins, peptides, or carbohydrate products produced by cell culture, excluding antibiotics, hormones, and products previously derived from human or animal tissue regulated as drugs
- Protein products made in animal body fluid by genetic alteration of transgenic animals, animal venoms, or constituents of venoms

Other classes of products identified as CBER-regulated products include:

- Synthetically produced allergenic products intended to specifically alter the immune response to a specific antigen or allergen
- Certain drugs used in conjunction with blood banking or transfusion

As will be seen at the end of this chapter in Section 19.7 (where the National Institutes of Health (NIH) also has some regulatory role), there is still some ambiguity of authority in areas where technology has outrun regulatory foresight.

Within the United States, the regulation of therapeutics is split on relatively arbitrary grounds. This is presented in Table 2.5. This chapter reflects both current FDA practices and the ICH guidelines. EMA guidances started with that of CPMP (1989) but subsequently have come from EMEA (now EMA). Most recently, subsequent to the TGN1412 near disaster, EMA has issued new guidances meant to issue a more conservative approach to clinical trials (EMA, 2007). With harmonization has come regulatory guidance from ICH (such as ICH S6(R1), 2014).

19.2 PRECLINICAL SAFETY ASSESSMENT

Because of the complexity and wide diversity of biologic products, their safety is evaluated on a case-by-case basis (in accordance with CBER's promulgated points to consider as summarized in Table 19.6) until such time as enough data on either specific products or a class of products are available.

Generally, *in vivo* nonclinical studies should be designed to include a sufficient number of animals per group to permit a valid estimation of a drug's toxicologic, pharmacologic, and immunogenetic effects in terms of incidence, severity, and the dose–response relationships involved. The latter point requires, as pointed out throughout this text, thoughtful selection of doses (and, for biologics, regimens of treatment). Comparable formulation, routes, and regimens of administration duration of exposure and suitable time to allow expression of expected response are also important proper design features. Tables 19.6, 19.7, 19.8, and 19.9 summarize the basic testing requirements for each of three subsets of biologically derived drugs.

The number of species necessary in preclinical testing programs varies. However, there is no specific requirement for the routine use of two species (e.g., one rodent and one nonrodent) in toxicology studies of biological products. Proper species selection for use in these trials is essential. Table 19.10 summarizes the key considerations for species selection.

In each stage of product development, it is important to determine exposure by measuring pharmacokinetic (PK) (including ADME) or PD end points. This includes the following: (i) measurements of the biologic in plasma or target organs, (ii) the distribution and persistence of cells for cellular therapies, (iii) measurements of viral shedding and recovery of certain values, (iv) localization of targeted novel delivery systems, and (v) tissue tropism, including germline tissue, of vectors used in gene therapies.

Such studies provide important information for a better interpretation of the toxicity observed in animals and aid in

TABLE 19.6 Points to Consider in the Preclinical Safety Assessment of Biologics

| |
|--|
| Rationale |
| <i>In vitro</i> or <i>in vivo</i> studies |
| Potency assays |
| Receptor characteristics (across species) |
| Physiological modeling |
| Scientific literature |
| Scientific speculation |
| Indication |
| Replacement therapy (long term) |
| Nonpharmacodynamic treatment (prophylactic or diagnostic) |
| Pharmacodynamic treatment (short term or long term) |
| Pharmacological activity (pharmacodynamics) |
| Primary end points |
| Secondary end points |
| <i>In vivo</i> model selection |
| Species-specific effects |
| Effects independent of species |
| Animal model of disease |
| Pharmacokinetics and ADME: correlation with pharmacodynamics |
| Low dose |
| High dose |
| General toxicity |
| Single dose (acute) |
| Repeated dose (subacute or subchronic) |
| Specific toxicity (may include one or more of the following studies) |
| Local irritation (local reactogenicity) |
| Antigenicity |
| Chronic toxicity |
| Reproduction toxicity, including teratogenic potential |
| Mutagenicity |
| Tumorigenicity |
| Carcinogenicity |
| Other toxicity concerns (e.g., neurotoxicity, immunotoxicity, etc.) |

the selection of not only the proposed initial human dose but of the dose-escalation scheme and the frequency of dosing in the clinical trial(s). Further, once such exposure data are available in humans, the data can be used to better correlate the human and animal findings. Toxicity studies should be performed in the same species used to assess efficacy (pharmacodynamics). Often times, exposure, toxicity, and immunogenicity are measured in the same study, particularly when nonrodents are used.

Toxicity studies should be designed not only to identify a safe dose but also a toxic dose threshold (and the spectrum of toxicodynamics) to anticipate the product's safety and to better define the therapeutic index in humans. Specific product considerations that may complicate the process of defining a toxic dose may include limits based on formulation, lack of significant systemic absorption, or the amount of the product available. The lack of significant toxicity in

TABLE 19.7 Biotechnology-Derived Drug Test Matrix

| Test Requirement | Species |
|--|---------|
| Develop bioanalytical for three species (man/rodent/nonrodent) | NA |
| <i>Initial clinical trial/IND requirement</i> | |
| Pilot single dose toxicity in rodents (IV ^a) | R |
| Pilot single dose toxicity in nonrodents (IV) | D |
| Genotoxicity only if appropriate (species cases) – check guidance for SRNA drug | |
| Safety pharmacology: CV <i>in vivo</i> (IV) | |
| Safety pharmacology: respiratory – rodent (IV) | R |
| Pivotal/repeat dose in rodents (match clinical plan ^b) | R/M |
| Pivotal/repeat dose in nonrodents (14–28 days) (match clinical plan ^b) | |
| Develop bioanalytical for three species (man/rodent/nonrodent) | N/A |
| Allergenicity (rodent) | |
| <i>To support continued clinical development</i> | |
| Pivotal/repeat dose in appropriate species to support marketing approval | |

Note: Species: B, rabbit; D, dog; M, mouse; P, primate; R, rat; S, pig; TBD, to be determined. All studies above must be performed GLP.

^aRecommended.

^b(for example) if proposed course of treatment is weekly, then 4 weekly administrations.

TABLE 19.8 Vaccine Test Matrix

| Test Requirement | Species |
|---|---------|
| Initial clinical trial/IND requirement | |
| 1. Acute toxicity in rodents (oral and IV ^a) | R/M |
| 2. Acute toxicity in nonrodents (oral) | D/S/P |
| 8. Pivotal/repeat dose in rodents (14–28-day oral) ^b | R/M |
| 9. Pivotal/repeat dose in nonrodents (14–28-day oral) | D/P/S |
| To support continued clinical development and to support marketing approval | |
| 13. Immunotoxicity | TBD |

Note: Species: B, rabbit; D, dog; M, mouse; P, primate; R, rat; S, pig; TBD, to be determined. All studies described above must be performed GLP.

^aRecommended.

^bMay be required.

animals does not necessarily mean that the product is safe. The margin of safety for the initial starting dose, however, will likely be adequate.

Acute studies should be irrelevant for biologics, except when acting as “canaries in the coal mine” for drugs such as TGN-412.

Studies are often 1–2 weeks in duration and routinely include body weight determinations, clinical observations, and gross necropsy findings. Additional antemortem studies may be performed as appropriate, especially in large animals (e.g., observation of local reactogenicity, PK evaluations,

TABLE 19.9 Biologics Test Matrix

| Test Requirement | Species |
|---|-----------------|
| Initial clinical trial/IND requirement | |
| 1. Acute toxicity in rodents (oral and IV ^a) | R/M |
| 2. Acute toxicity in nonrodents (oral) | D/S/P |
| 3. Seven-day DRF toxicity in rodents (oral) | R/M |
| 4. Seven-day DRF toxicity in nonrodents (oral) | D/S/P |
| 5. Genotoxicity: bacterial mutagenicity (Ames) ^b | <i>In vitro</i> |
| 6. Pivotal/repeat dose in rodents (28-day oral) | R/M |
| 7. Pivotal/repeat dose in nonrodents (28-day oral) | D/P/S |
| 8. Develop bioanalytical for three species (man/rodent/nonrodent) | NA |
| To support continued clinical development | |
| 9. Developmental tox (Seg II)—rat and rabbit pilots and rat and rabbit studies ^b | R/B |
| 10. Neoantigenicity | B |
| To support marketing approval | |

Note: Species: B, rabbit; D, dog; M, mouse; P, primate; R, rat; S, pig; TBD, to be determined. All studies described above must be performed GLP.

^aRecommended.

^bMay be required.

hematological and/or clinical biochemistry measurements). Histological evaluations may also be performed.

The duration of repeat-dose studies should be at least as long as the proposed dosing exposure in clinical study. These studies are designed to establish a dose–response relationship, define target organ(s) of toxicity, and determine whether observed toxicities are reversible. Evaluation parameters should include not only those routinely performed in the acute studies but those performed in the additional studies as well. Special tests such as ophthalmoscopic, electrocardiograph, allergenicity, body temperature, and blood pressure monitoring are often included. Depending on the study duration, PK sampling at multiple time points may be necessary to better characterize the kinetics of response. As mentioned, a group of animals will be examined at term, and some may be reserved for at treatment-free or recovery period to evaluate the reversibility of any findings.

Specific (local tissue tolerance) toxicity studies may be necessary due to special characteristics of the product or the clinical indication. Adjuvanted vaccines are routinely evaluated for local (injection site) reactions, and cellular therapies are routinely screened for tumorigenic potential. Research is also needed to better predict the sensitizing potential of biological products and to determine the relevance of serum antibody levels following repeat dosing in animals and humans.

While carcinogenicity studies have not been performed routinely for biological products, they may be appropriate for products proposed for chronic use. Reproductive toxicology studies will probably become more common, especially as more women of childbearing potential participate in early clinical trials. In the past, such studies have not been

TABLE 19.10 Factors to Consider in Species Selection for Protein Therapeutic Development

| Cross-Reactivity | Immunogenicity |
|--|--|
| Cross-reactivity alone is not sufficient to identify a relevant species. Suitable affinity and potency to give valid results are also necessary. How close does the level of potency have to be to the situation in humans to be meaningful? | An emerging issue as more mAbs are developed for chronic use is the impact of neutralizing antibodies on repeat-dose studies. This is a significant scientific problem that might be partially overcome by the use of the surrogate antibodies. What is the potential for the regulatory acceptability of surrogate antibodies for immunogenicity reasons? |
| How can alternative approaches be “front-loaded” into the development pathway without leading to a standard two-species approach? | To what extent is immunogenicity considered in species selection for safety and toxicology studies? Should this be given more priority in the selection of relevant species for long-term toxicity studies for mAbs that are intended for chronic indications? |
| What is the possibility of assigning greater value to nonconventional preclinical studies, such as genetically altered rodents and surrogate antibodies that are scientifically relevant? | |
| If, after a variety of cross-reactivity testing that includes binding studies, functional activity in cell-based systems, sequence homology, and tissue cross-reactivity studies, the only relevant species is chimpanzees, is it justified, scientifically and ethically, to use the chimpanzees to study the effects of the monoclonal antibody (mAb)? | |

conducted for biologics. Reproductive toxicology studies have been performed with many of the recently approved therapeutics (e.g., IFNs, interleukins, cytokines, growth factors, etc.). Such studies also have been conducted in the development of AIDS vaccines intended for use in pregnant women. The standard protocol designs were modified to address specific vaccine-related concerns, including dosing in relationship to immunologic effects.

The development of biologics to treat various nervous system diseases (immunologically based and otherwise) has involved additional, specific neurotoxicological studies on these products. However, despite the fact that most products regulated as biologics have an immune component or impact directly or indirectly on the immune system, standardized immunotoxicity tests that are potentially useful in screening large numbers of chemicals for their ability to adversely affect the immune system have not proven essential in assessing the safety of biological products.

Throughout the various phases of product development, additional preclinical safety studies may be necessary due to unexpected toxicity, significant changes in the manufacturing process or the final formulation, or changes in the clinical indications. In some cases, the ideal assessment of the safety of novel biological therapies may require alternative approaches, such as *in vitro* or *in vivo* organogenesis model systems, animal models of tolerance, animal models of disease, or transgenic animal models.

19.3 RECOMBINANT DNA TECHNOLOGY

The concept of rDNA technology is based on the premise that a gene sentence may be taken from an animal or human

gene responsible for the production of a particular protein and inserted into the DNA of *Escherichia coli*, a single-cell bacterium. The bacterial cells then divide very rapidly, making billions of copies of themselves, including a replica of the gene that has been inserted.

There are unique ways to insert human genes into bacteria. In addition to chromosomal DNA, bacteria have numerous copies of extrachromosomal circular DNA called plasmids, which are not attached to the bacterial chromosome. These plasmids can be transferred from one bacterium to another by conjugation (e.g., mating) and can be isolated from bacteria and easily purified. Through the use of restriction enzymes (i.e., a family of enzymes that can cut DNA at specific base sequences), the gene sentence to be inserted can be isolated, and the plasmid DNA can be opened. While in the open state, the desired piece of animal or human DNA can be inserted. Through the use of ligase enzymes, complementary ends of the plasmid can be connected, thereby producing a recombinant plasmid recombined by joining two heterologous pieces of DNA. This recombinant plasmid can then be put back into the bacteria, and the bacteria will express the new gene function that has been inserted.

A unique characteristic of plasmids is that thousands are produced within each bacterium to the point that up to 40% of the total DNA of the bacterium may in fact be plasmid DNA. Hence, a single piece of human DNA that heretofore could only be obtained in low concentrations can be recombined with a plasmid, and the DNA sequence multiplied by a million- or a billion-fold (i.e., cloned). Use of cloning techniques may produce many grams of a particular human protein, instead of the few molecules that are produced in normal cells.

Examples of the early application of rDNA technology in medicine are the development of recombinant human growth hormone (HGH); human insulin; human IFNs, thought to have anticancer activity in addition to antiviral activity; interleukins (regulatory proteins from lymphocytes that are believed to be important in the treatment of immunodeficiency diseases and cancer); tumor necrosis factor; epidermal and bone marrow progenitor cell growth factors; and the production of vaccines (Table 19.1).

HGH and insulin produced by rDNA technology are already registered with the US FDA for therapeutic use. The applications of rDNA technology agriculture should improve the quality of domesticated animals through the production of new and improved vaccines, growth-promoting hormones, and less expensive food additives. Seed crops will be produced that offer improved yields and better resistance to environmental conditions. Further applications may include the insertion of genes into plants or bacteria for production of toxins that can act as biochemical pesticides or allelopathic agents (chemicals that act as natural herbicides to prevent the growth of other plant species in the same geographical area).

19.3.1 General Safety Issues

rDNA technology represents one of the most innovative achievements in biology in the last century. Although the new technology has generated much enthusiasm for its potential applications, it has also raised concerns among both scientists and the public in general. Many of the early fears of the inadvertent development of an "Andromeda strain" during the genetic engineering of a specific microbe have long since vanished. However, other concerns remain. Can a gene cloned for toxic production from an rDNA microbe be transferred into normal bacterial flora? Could antibiotic-resistant genes be cloned and inadvertently inserted into clinically relevant pathogens not presently antibiotic resistant? To reduce these possibilities, the NIH has only certified nonconjugative plasmids (e.g., nomating) for use in rDNA microbes.

Studies by Levine et al. (1983) have addressed the issue of plasmid mobilizations, the movement of plasmids between different host cells. Human volunteers fed tetracycline along with *E. coli* HS-4 (typical of the normal intestinal flora of humans) bearing highly mobilized plasmids (e.g., pJBK5) that carried resistance to chloramphenicol and tetracycline became cocolonized with *E. coli* HS-4 bearing the antibiotic-resistant plasmid. However, the use of a poorly mobilizable plasmid (pBR325) did not result in plasmid transfer.

Taken as a whole, these studies establish the safety of rDNA research when poorly mobilizable cloning vectors are used while supporting the rationale for biologic containment of highly mobilizable plasmids. They also point out the need to protect laboratory workers on antibiotic therapy from

potential exposure to rDNA organisms carrying any sort of antibiotic-resistant genes. A reassuring point is the relatively poor survival of rDNA strains of *E. coli* in the intestinal environment. For example, in most successful studies, 50 billion *E. coli* HS-4 organisms were required to ensure survival within the harsh environment of a human's stomach.

19.3.2 Specific Toxicological Concerns

While rDNA techniques offer exciting possibilities, there are many unanswered questions about the potential toxicity that each new product represents. For example, acute clinical toxicities of IFNs include flu-like syndrome, fever, chills, malaise, anorexia, fatigue, and headache. Chronic dose-limiting toxicities include neutropenia, thrombocytopenia, impairment of myeloid maturation, reversible dose-related hepatotoxicity, some neurological toxicity (stupor, psychosis, peripheral neuropathy), and gastrointestinal toxicity. Some of these toxicities would be difficult to ascertain in rodents and, in fact, may be species specific.

A particular toxicity associated with the administration of IFN to human and experimental animals has been depression of the cytochrome P450 monooxygenase (MFO)-metabolizing enzymes. As a consequence of MFO inhibition following treatment with IFN, the sleep time of mice treated with hexobarbital is increased, as is the toxicity of acetaminophen (Stebbing and Weck, 1984). Possible effects on the metabolism of chemotherapeutic agents or other drugs processed by the P450 MFOs should be anticipated.

The *in vivo* antitumor effects of IFNs are believed to be related to both augmentation of natural killer cell activity and antiproliferative effects. Antiproliferative activity probably also accounts for the bone marrow suppression observed in some individuals given IFN and could potentially produce effects in a routine preclinical reproduction or teratology evaluation. Dosing studies performed in newborn mice with homologous IFN have resulted in death at high doses and a marked wasting syndrome when given over an extended period (Gresser and Bourali, 1970). Both effects were attributed to the antiproliferative activity of IFNs. Inhibition of proliferation and metabolism represent potential dose-limiting toxicities of this family of rDNA molecules.

19.4 IMMUNOGENICITY/ALLERGENICITY

Human biopharmaceuticals are commonly immunogenic (elicit and antibody response) in nonhuman species. Immunogenicity should be evaluated in repeat-dose nonclinical safety studies to help determine whether antibody may have influenced pharmacology, toxicity, or exposure. If immunogenicity decreases exposure or neutralizes the activity of the biopharmaceutical, it may not be appropriate to continue the study or conduct studies of longer duration in

that species. It is important to confirm biological activity of administered biological in “no-effect” studies where the top dose is the NOEL. The immunogenicity of a biopharmaceutical in nonclinical species is not necessarily predictive of a potential for antibody formation in humans, but the potential consequences (summarized in Table 19.11) must be considered. Table 19.12 summarizes the strategies to conduct a bioanalytical evaluation of immunogenicity responses.

19.5 MONOCLONAL ANTIBODY TECHNOLOGY

Offering an impressive potential for human therapy, mAbs have become the first commercialized products of the new biotechnology. They have become widely used in diagnostic medicine and are very successful as therapeutic agents in cancer (Oldham, 1983). In clinical diagnostic medicine, they have provided us with the sensitivity not heretofore available for specific and rapid diagnosis of a particular drug level or infectious disease process.

TABLE 19.11 Consequences of Immunogenicity

| | |
|--------------------|----------------------------------|
| Loss of efficacy | Enhancement of efficacy |
| Insulin | Growth hormone |
| Salmon calcitonin | |
| Factor VIII | Neutralization of native protein |
| Interferon alpha 2 | MDGF |
| Interferon beta | EPO |
| IL-2 | |
| GnRH | General immune effects |
| TNFR55/IgG1 | Allergy |
| HCG | Anaphylaxis |
| GM-CSF/IL3 | Serum sickness |

Note: Immunogenicity in humans may be a major safety concern.

Antibodies are important in the body as defense against infectious agents. They are extremely specific proteins that are produced in response to a foreign material, or antigen, by lymphoid cells of the immune system and share the property of being able to bind specifically to the inducing antigenic epitope (a single antigenic determinant; that portion of the antigen which combines with the antibody paratope). Unfortunately, under most conditions of antigenic stimulation, a family of antibodies is produced, each with a slightly different antigenic specificity.

In 1975, Kohler and Milstein observed that if an antibody-producing cell was fused with a myeloma tumor cell; a rapidly dividing hybrid was produced that synthesized a monospecific antibody. Each hybridoma formed then became a “factory,” producing antibodies monospecific to a particular sensitizing antigenic epitope. Cell cloning allows selection of hybrids producing antibody with the desired characteristics.

mAbs are thought to represent a major advance in cancer therapy because they have a very high therapeutic-to-toxic index when compared with anticancer drugs or radiation therapy, and should provide a greater degree of specificity for the tumor cell than other forms of therapy. The conjugation of toxins with mAbs is theoretically very exciting because a high specific toxin activity could be achieved at the tumor target cell.

Clinically, mAbs are also proposed as drug delivery vehicles in certain tumors where specific tumor-associated antigens are expressed. In this context, investigators have found that by conjugating toxins such as the A chain polypeptide of the plant protein ricin or the bacterial toxin from *Corynebacterium diphtheriae* to mAbs specific for certain tumor type, as few as one or two molecules of antibody-toxin conjugate can destroy a tumor cell *in vitro*. Some success has also been obtained in clinical trials with mAb-toxin conjugates.

TABLE 19.12 Immunogenicity Bioanalytical Strategy for Animal Studies

| Bioanalytical Scheme for Lower-Risk Products | Bioanalytical Scheme for Medium- and Higher-Risk Products |
|--|---|
| <i>Frequency of sampling within study</i> | |
| Planned assessment of baseline and an appropriate, drug-free, end-of-study immunogenicity sampling time point. Contingent analysis of dosing phase samples if required to support pharmacokinetic profiles | Planned assessment of baseline and an appropriate, drug-free, end-of-study immunogenicity sampling time point. Contingent analysis of dosing phase samples if required to support pharmacokinetic profiles |
| <i>Assessment of ADAs</i> | |
| Detection of ADAs through screen and confirmatory immunoassays | Detection of ADAs through screen and confirmatory immunoassays |
| Consideration of characterization of titer/relative concentration of ADAs | Characterization of titer/relative concentration of ADAs |
| | Detection of cross-reactivity to endogenous counterpart through screen and confirm (and maybe titer) immunoassays |
| | If antibody reactivity to endogenous counterpart is detected, characterization of neutralizing ability using target binding inhibition-based neutralizing antibody immunoassay or cell-based neutralizing antibody bioassay |

ADAs, antidrug antibodies.

mAbs have also been proposed for detoxification of individuals suffering from drug overdose or chemical intoxication, as well as for radioimaging of tumor burden or metastatic foci. In veterinary medicine, mAbs are already being used to develop new rapid methods for diagnosis of infections in poultry, cattle, and other animals.

19.5.1 Toxicological Concerns with Monoclonal Antibodies

It has been clear for a while that there are certain problems implicit in the use of mAbs in therapeutic uses. Hansel et al. (2010) overviewed the scope of the recognized issues, and the issues (and incidences of problems) have only expanded with ever broadening uses of products (see Table 19.13). For example, there may be modulation of the antigenic determinant on the target cell, so that the mAb cannot recognize its appropriate antigenic epitope. Secondly, the tumor cell may release free antigens so that the mAb is effectively neutralized before it can reach the target cell. Third, antibodies to mouse epitopes on the mAb could be induced (this may be overcome in the future by the use of human–human hybrids or the use of immunosuppressive agents to prevent the development of antibodies). Fourthly, mAbs have an extremely short half-life in systemic circulation, which would require that they be intermittently infused to provide the beneficial effect. Lastly, there may be an unwanted release of the toxin from its conjugate; or specificity problems may develop, whereby the antibody–toxin conjugates end up in an inappropriate organ.

The clinical toxicology findings associated with the use of mAbs in therapeutic trials have included fever, chills, flushing, dyspnea, hypotension and tachycardia, anaphylactic and anaphylactoid reactions, urticaria, rash, nausea, elevated creatinine levels, headache, bronchospasm, and serum sickness (Oldham, 1983). Few of these reactions can potentially be predicted from safety evaluation in rodents. A major problem with using the intact ricin or diphtheria toxin molecule, containing both the A and B polypeptide chains, has been the dissociation of the parent molecule from the mAb, leading to toxicity of the reticuloendothelial system. A promising solution to this problem comes from separating the A chain (toxic moiety) from the B chain (cell association moiety) and preparing only A chain conjugates. This results in much lower toxicity if the A chain should become dissociated from the antibody conjugate because cellular association does not occur.

The FDA website provides a list of mAbs that have been reviewed and were found to have evidence of fetal risk.

Since an mAb is a fusion product of a malignant mouse cell and an antibody-producing cell, there is some concern about the safety of the production process itself (Petricciani, 1983). Methods for the production of mAbs raise two general safety issues: (i) the theoretical risk of transferring in the

product factors associated with malignancy (e.g., oncogene factors) and (ii) the use of animals for antibody production that are known to harbor a number of microbial agents, some of which can produce diseases in humans.

Preclinical studies should address the potential toxicity due to inappropriate release of the conjugated toxin. Preclinical toxicology of mAbs may not require extensive animal studies but should examine for cross-reactivity with antigenic epitopes present on normal cells *in vitro* and for the presence of human or rodent viruses. Early clinical trial should involve biodistribution studies with radiolabeled material.

The level of regulatory concern with the safety of the products of biotechnology underwent a sea change in early 2006 with the near catastrophe with TGN1412, a superagonist anti-CD28 mAb that induces the production of anti-inflammatory cytokines by directly stimulating T cells.

A German company, TeGenero, had first-in-man clinical trial of the product initiated by a clinical research organization (CRO) at a clinic within a hospital in London. Eight healthy young males were enrolled, and six received a dose of the drug, while two received placebo. Within 90 min after receiving a single intravenous (IV) dose of the drug, all six volunteers had a systemic inflammatory response characterized by a rapid induction of proinflammatory cytokines and accompanied by headache, myalgias, nausea, diarrhea, erythema, vasodilation, and hypotension. Within 12–16 h after infusion, they became critically ill, with pulmonary infiltrates and lung injury, renal failure, and disseminated intravascular coagulation. Severe and unexpected depletion of lymphocytes and monocytes occurred within 24 h after infusion. All six patients were transferred to the care of the authors at an intensive care unit at a public hospital, where they received intensive cardiopulmonary support (including dialysis), high-dose methylprednisolone, and an anti-interleukin-2 (IL-2) receptor antagonist antibody. Prolonged cardiovascular shock and acute respiratory distress syndrome developed in two patients, who required intensive organ support for 8–16 days. Despite evidence of multiple cytokine release syndrome (CRS), all six patients survived (Goodyear, 2006; Suntharalingam et al., 2006). This response has been characterized as a “cytokine storm,” similar to that seen earlier with OKT3.

The preclinical evaluation which occurred before the trial was conducted was a set piece approach to the then guideline requirements. Studies were conducted in rats, mice, and primates, with the repeat-dose studies in the rats and mice being once a week for 4 weeks at doses up to 50 mg kg⁻¹, with these doses being characterized as well tolerated with transient increases in CD4 and CD8 (which were expected) and of IL-2, IL-5, and IL-6 but with no signs of first-dose CRS. There was cross-reactivity in stained lymphoid tissue and astrocytes in both human and primate tissues, but no histopathology signs of CNS toxicity were seen in safety studies.

TABLE 19.13 Side Effects of Licensed Monoclonal Antibodies

| Target | mAb | Type | FDA Approval | Indications ^a | Selected Side Effects |
|---|--|--|-------------------------------------|--|---|
| Platelet glycoprotein IIb/IIIa | Abciximab (ReoPro; Centocor Ortho Biotech, Eli Lilly) | Chimeric antibody fragment: c7E3 Fab | 1994 | Prevention of ischemic cardiac complications of percutaneous coronary interventions and unstable angina | Hypersensitivity and immunogenicity Increased risk of bleeding Thrombocytopenia |
| Tumor necrosis factor- α | Adalimumab (Humira; Abbott) Certolizumab (Cimzia; UCB) Infliximab (Remicade; Centocor Ortho Biotech) | Fully human Humanized PEGylated Chimeric | 2002 2008 1998 | Rheumatoid arthritis Ankylosing spondylitis Psoriasis Psoriatic arthritis Crohn's disease Ulcerative colitis | Infusion reactions and immunogenicity Hypersensitivity reactions Immunosuppression and infections (tuberculosis) Anemia, leukopenia, and thrombocytopenia Worsening heart failure Malignancy, lymphoma, and lymphoproliferative disorders Elevated liver transaminases Increased nuclear-specific antibodies |
| CD52 on mature B, T, and natural killer cells | Alemtuzumab (Campath; Genzyme) | Humanized | 2001 | B-cell chronic lymphocytic leukemia Graft-versus-host disease Multiple myeloma Multiple sclerosis Vasculitis Behçet's disease | Infusion reactions Hypersensitivity and immunogenicity CRS Tumor lysis syndrome Immunosuppression and opportunistic infections Cytopenias: pancytopenia, lymphopenia, and thrombocytopenia Autoimmune hemolytic anemia Thyroid disorders Cardiotoxicity |
| Interleukin-2-receptor- α on activated lymphocytes | Basiliximab (Simulect; Novartis) Daclizumab (Zenapax; Roche) | Chimeric Humanized | 1998 1997 discontinued in Europe | Prophylaxis of renal transplant allograft rejection | Severe acute hypersensitivity reactions CRS and immunogenicity Immunosuppression and infections Local skin reactions Warnings when combined with other immunosuppressives |

(Continued)

TABLE 19.13 (Continued)

| Target | mAb | Type | FDA Approval | Indications ^a | Selected Side Effects |
|------------------------------------|--|---|---|--|--|
| Vascular endothelial growth factor | Bevacizumab (Avastin; Genentech) | Humanized | 2004 | Metastatic colorectal cancer Non-small-cell lung carcinoma Metastatic breast carcinoma Metastatic renal carcinoma | Infusion reactions and immunogenicity Local complications at tumor site Arterial and venous thromboembolic events Hemorrhage Severe hypertension Cardiac failure Reversible posterior leukoencephalopathy Slower wound healing and GI perforation |
| | Ranibizumab (Lucentis; Genentech, Novartis) | Humanized (Fab fragment from bevacizumab) | 2006 | Injected intravitreally for neovascular (wet) age-related macular degeneration | Conjunctival hemorrhage Intraocular inflammation Increased intraocular pressure Retinal detachment Endophthalmitis |
| Complement C5 | Eculizumab (Soliris; Alexion) | Humanized | 2007 | Paroxysmal nocturnal hemoglobinuria | Meningococcal and <i>Neisseria</i> infection Intravascular hemolysis |
| CD11a | Efalizumab (Raptiva; Genentech) | Humanized | 2003 recently discontinued | No longer licensed for chronic plaque psoriasis | First-dose reaction complex Immunosuppression Serious opportunistic infections PML Guillain-Barré syndrome, encephalitis, meningitis Immune hemolytic anemia Immune thrombocytopenia |
| CD3 antigen on T cells | Muromonab-CD3 (Orthoclone OKT3; Ortho Biotech) | Mouse | 1986 (no European Medicines Agency authorization) | Acute resistant allograft rejection in renal, cardiac, and hepatic transplant patients | Severe acute infusion reactions Immunosuppression and infections Immunogenicity Cardiovascular side effects Hepatitis |
| α 4-Integrin | Natalizumab (Tysabri; Biogen Idec, Elan Pharmaceuticals) | Humanized | 2004 | Highly active relapsing–remitting multiple sclerosis | Infusion and hypersensitivity reactions Immunogenicity PML (0.1%) with immunosuppressives Hepatotoxicity |
| Immunoglobulin E (IgE) | Omalizumab (Xolair; Genentech, Novartis) | Humanized | 2003 | Severe allergic asthma unresponsive to conventional therapy and with acute exacerbations | Anaphylaxis (0.1%) Injection site reactions Immunogenicity URTI Churg–Strauss syndrome (rare) |
| Fusion protein of RSV | Palivizumab (Synagis; MedImmune) | Humanized | 1998 | Prevention of RSV complications in high-risk infants | Anaphylaxis and apnea (rare) Fever, injection site reactions |

| | | | | | |
|----------------------|---|-------------|------|--|--|
| CD20 on B cells | Rituximab (Rituxan/ MabThera; Genentech, Biogen Idec) | Chimeric | 1997 | Follicular non-Hodgkin's lymphoma CD20 ⁺ diffuse large B-cell non- Hodgkin's lymphoma Autoimmune hematological disorders | Prominent acute infusion reactions CRS Tumor lysis syndrome Transient hypotension Immunogenicity Serum sickness Severe mucocutaneous reactions Immunosuppression Hepatitis B reactivation with fulminant hepatitis PML Renal toxicity Cardiac arrhythmias |
| EGFR | Panitumumab (Vectibix; Amgen) | Fully human | 2006 | Monotherapy for EGFR-positive metastatic colorectal carcinoma with nonmutated (wild-type) <i>KRAS</i> after failure of conventional chemotherapy | Infusion reactions Skin rashes in most patients (90%) Diarrhea (60%), nausea, and vomiting Hypomagnesemia (2%) |
| | Cetuximab (Erbix; Bristol-Myers Squibb, ImClone Systems, Merck Serono) | Chimeric | 2004 | EGFR-positive metastatic colorectal cancer Squamous cell carcinoma of head and neck | Severe infusion reactions IgE against oligosaccharide and HAMA Urticaria and dermatological toxicity Bronchospasm and pulmonary toxicity Hypomagnesemia |
| | Trastuzumab (Herceptin; Genentech) | Humanized | 1998 | ERBB2-positive breast carcinoma | Hypersensitivity and infusion reactions Cardiotoxicity with anthracyclines Skin reactions Pulmonary toxicity Hypomagnesemia |
| Interleukin receptor | Tocilizumab (Actemra; Roche, Chugai) | Humanized | 2009 | Unresponsive active rheumatoid arthritis Castleman's disease | Anaphylaxis and anaphylactoid reactions UTRI Headache Serious infections Abnormal liver function, neutropenia, and lipid deregulation |

Source: Adapted from Hansel et al. (2010, pp. 325–338).

CRS, cytokine release syndrome; EGFR, epidermal growth factor receptor; ERBB2, also known as HER2/neu; FDA, Food and Drug Administration; GI, gastrointestinal; HAMA, human antimouse antibodies; *KRAS*, v-Ki-ras2 Kirsten rat sarcoma viral oncogene homolog; PML, progressive multifocal leukoencephalopathy; RSV, respiratory syncytial virus; UTRI, upper respiratory tract infection.

^a Some of these indications are not currently licensed.

The NOAEL was set at 50 mg kg⁻¹ in the primate (cynomolgus monkey), and with FDA human equivalent dose (HED)-style allometric scaling and safety factor of 10, then additional safety factor applied proposed 0.1 mg kg⁻¹ for first-in-man clinical trials, which was the dose used in the trial.

What was not done for this highly humanized mAb was to use the appropriate model (either a knockout mouse responsive to the humanized molecule or evaluation of a suitable homolog molecule).

The EMEA response has been to put in place the minimum active biological effective level (MABEL) approach to setting first doses in clinical trials (EMA, 2007). In the case of TGN1412, based on biologic activity in the rat, a safe dose would have been set at 0.005 mg kg⁻¹.

19.6 BIOPROCESS TECHNOLOGY

In the chemical and pharmaceutical industry, rDNA technology will allow the synthesis of chemical that can only be practically achieved through a bioprocess. For example, methylation of a particular carbon in a chemical structure might be done quite easily with a recombinant-engineered bacteria. This technology will allow the synthesis of a family of isomers and the development of a synthetic process that cannot be achieved by strict physical chemical processes.

This area of rDNA technology also has application in the degradation of solid waste materials: in wastewater recovery, in leaching minerals from ore-containing rock, in improved oil recovery, and in the decontamination of chemical waste dumps through the engineering of microorganisms that can destroy specific toxic contaminants.

19.7 GENE THERAPY PRODUCTS

Gene therapy products, while holding tremendous promise, have so far delivered but limited (two cases as of this writing) positive outcomes. The concept involved—inserting functioning genes in place (or places) where nonfunctional or malfunctioning genes have produces a disease state—is stunning. But the public outcry over the death of Jesse Gelsinger, an 18-year-old in a clinical trial at the University of Pennsylvania's Institute of Human Gene Therapy in 2000, has lead to a significant slowdown in the rate of evaluation. This tragic event, probably due to an innate immune response to a protein in the vector's protein coat (Stephenson, 2001), has led to increased restrictions.

Five aspects specific to gene therapy need to be evaluated to assess the safety of a therapy:

1. DNA/RNA biodistribution
2. Gene transfer and biological activity

3. Risk of vertical transmission of the gene
4. The safety of the vector (the means of delivering the gene to the intended site)
5. The safety of the product protein
 1. Evaluate DNA/RNA biodistribution
 - Radiolabeling
 - Southern blot
 - PCR
 - Real-time PCR
 - *In situ* PCR
 2. Evaluate gene transfer and biological activity
 - Immunohistochemistry
 - Western blot
 - Enzyme-linked immunosorbent assay (ELISA)
 - Flow cytometry
 3. Evaluate the risk of vertical transmission
 - To gonads
 - If yes, then look at semen/germ cells (probably required anyway)
 - To circulating blood
 - If so, how long (persistence)?
 4. Assess the safety of the vector.
 - Identify a suitable model species.
 - Assess the acute toxicity of the vector particle (in rabbits):
 - At high dose, potential for an anaphylactic response (not seen in mice)
 - Also look for neutrophil proliferation
 5. Assess the safety of the product protein
 - Using data from preclinical pharmacology toxicology ("safety") studies to support the safety of clinical trials
 - Ones not appropriate for generator (instead, assess integration/insertion frequency in a mammalian cell)
 - Standard rodent carcinogenicity *probably* not appropriate

Regulatory authority for gene therapy products is unique in that overlapping responsibilities extend to both CBER and NIH.

| | |
|------|--|
| CBER | Division of Cell and Gene Therapies |
| | Manufacturing |
| NIH | Division of Clinical Trial Design and Analysis |
| | Preclinical pharmacology and toxicology |
| | Clinical trial design, safety, and efficacy |
| | 20% of CBER clinical protocols are now for gene therapy |
| | Recombinant DNA Advisory Committee (RAC) |
| | No authority for approval for clinical trials, but all adverse events must be reported |

This duplicating authority has led to both misunderstanding and problems in trials, as investigators must report adverse responses to both. But each has a different definition of what constitutes a reportable adverse response.

Currently, gene therapy is restricted to life-threatening and severely disabling diseases, but when a larger safety database has been accumulated, there should be expanded opportunities for therapy. It is not possible or desirable to identify a uniform “recipe” for the safety studies that should be conducted with gene therapy products, to support either the first dose in humans or extended clinical evaluation. Each product should be treated on a case-by-case basis, taking into consideration a number of important factors, such as the clinical indication, the duration of expression of the gene, and whether DNA transfer will be *in vivo* or *ex vivo*. For example, elimination of a tumor may require short-term treatment such as the transient expression of a suicidal gene. On the other hand, treatment may be long term, such as the replacement of a missing enzyme in the liver, where the goal may be lifetime expression. Gene therapy is currently an area of limited but rapidly advancing knowledge, and study design should be based on previous experience, together with ongoing feedback from the clinic throughout development. Considerable early thought should also be given to appropriate assays and their sensitivities. The choice of assay will need to be justified and the basic “toolkit” of assays properly considered and evaluated in advance.

19.7.1 Vectors

In gene therapy, genes typically are delivered using a vector which may be nonviral or viral. The complete construct should be tested; separate safety evaluation studies of vectors *per se* are not generally recommended except to explore mechanisms of action if potentially harmful effects have been demonstrated in a previous investigation, for example, red cell agglutination on IV administration. If a novel nonviral vector is to be used, evidence of its lack of toxicity and information on its basic pharmacokinetics will be an essential component of the preclinical package. The interaction between the vector and the gene is quite important, perhaps more so with the nonviral gene therapeutics, where the physicochemical properties of the particles themselves very much determine which tissues take up the gene. For viral vectors it is important to have sufficient knowledge of how they replicate, how to render the viruses replication incompetent, any inherent pathogenicity and immunogenicity, and any risk of recombination with wild-type virus.

Conventional pharmaceutical quality assurance procedures should be applied to gene therapy products as well as appropriate infectivity tests for self-replicating/living vectors.

19.7.2 Studies to Support the First Dose in Man

The most scientific approach is to replicate, in an appropriate animal, the type of dosing that would be expected to be used in humans, employing the “dose for dose/animal to human” principle. A single, suitable animal species should suffice. If viral vectors are used, the animal species should be sensitive to infection by the wild-type virus. Studies should not automatically be done in primates, but, initially, the commonly used laboratory species should be utilized. Only if those are demonstrated to be unsuitable should the next step be to consider the use of a primate.

Based on these general principles, the first dose in humans should be supported by a single-dose study in an appropriate animal species by the intended clinical route. Several dose levels should be explored, as some gene therapy expression products have a narrow therapeutic index.

There are many circumstances when a single-dose IV study can provide useful information; for example, if the intended treatment route is intraperitoneal (IP), or if the product will be administered to an open wound or injected into a muscle or a tumor, it might accidentally enter a blood vessel, and the knowledge gained from an IV study would be of value as well as one by the clinical route. Hence, if the intended clinical route is not IV, the absence of an additional study with IV dosing would require specific justification.

The physiological consequences of the gene product should be explored in these studies, particularly with totally novel gene products. In addition, all the standard toxicological evaluations should be carried out, including examination of functional end points *in vivo* including cardiovascular and respiratory effects.

19.7.3 Distribution of the Gene and Gene Product

The distribution of the gene must be evaluated carefully in time-point assays. The choice of assay used, with regard to their specificity and sensitivity, must be justified. The objectives are to identify the tissues in which the gene is present, to demonstrate whether or not the gene product is expressed in particular tissues, and to demonstrate the time course of gene expression, that is, how long it persists.

Some regulatory authorities, including the United Kingdom, are particularly concerned that the possibility of alteration to germline cells should be excluded. Both male and female gonads should therefore be examined. If the gene is found there, then it is necessary to examine the gonads at a more detailed level to ascertain whether the gene is present in the actual germ cells. Where gene persistence is short as, for example, in nonintegrating nonreplicating vectors, assay of the gonads at an appropriate time point will minimize false-positive findings.

19.7.4 Studies to Support Multiple Doses in Humans

The animal studies should parallel the intended treatment regime in humans. At their present stage of development, a

single dose of a gene therapeutic may not be totally curative, so multiple cycles of treatment may be used clinically rather than a single period of administration. It is appropriate to explore this in animal safety work, that is, the cycle regimen should be paralleled in the animal, up to a maximum of three cycles. The duration of follow-up in the test animals after completion of the last test cycle should be based on the duration of gene expression, up to a maximum of 6 months. There may be situations, particularly if long-term gene expression is the goal, when there could be an argument for a longer follow-up but that should be considered on a case-by-case basis.

There should not be blind adoption of a checklist of assays and observations, but appropriate investigations should be selected, based on earlier findings in the single-dose studies. Increasing the number of doses raises more concern about the immune response. There may be indications of this, such as lymphocyte infiltration at the site of administration, and there are many markers from single-dose studies that would indicate when it might be appropriate to examine the immune response to the gene product and selection markers, for example, immunity to adenoviral vectors, or to expressed proteins, resulting in accelerated loss of the transgene.

19.7.5 Unnecessary Studies

Genotoxicity and carcinogenicity studies are not generally recommended or required. Eliminating the possibility of the gene being inserted in the germline ensures that some elements of concern about reproduction toxicology have already been addressed; hence, classical reproduction and developmental toxicity studies are not generally recommended. They should be considered on a case-by-case basis; for example, if the treatment were to be intended to manage a long-term metabolic disease, the patients would then survive to reach reproductive competence.

Drug interaction studies are not generally appropriate, with the exception of gene-directed enzyme prodrug therapy (involving a gene expressing an enzyme that activates a prodrug given subsequently). In the latter case, it is necessary to demonstrate the presence of the prodrug as well as the gene in the animal and to consider the potential toxicity of the active metabolite(s) both locally (which is the desired pharmacological effect) and systemically (the undesirable effects).

19.7.6 *Ex Vivo* Procedures

Ex vivo procedures involve removing cells and transfecting or transducing them. The cells should be checked to confirm that they are all healthy and that they are still expressing their normal surface markers; observations of normal growth characteristics can also be reassuring. Animal studies are of limited value to test the safety of transfected or transduced human cells.

19.7.7 Change of Gene or Vector

Currently, only a limited number of vectors are available, although there is a large array of inserted genes. If the therapy involves developing a construct of a new gene in a well-characterized vector, it is important to use existing information on the vector. Rather than regenerating data on the vector itself, bridging studies of the construct should be carried out, that is, additional studies involving a limited toxicology evaluation to specifically characterize the nature of the new gene. Since the safety of the vector is already known, this should drive more exploration of the effects of the gene and the gene product rather than the vector itself.

If the vector is changed, a full safety evaluation may be required. However, if the changes are minor compared to the structure of a fully evaluated vector, it is appropriate for safety to be addressed by bridging studies. For example, if there is only a minor change on one of the condensation peptides of a nonviral, self-assembling vector, then some simple bridging work, rather than a full evaluation, may be appropriate.

The possibility of abbreviated testing is referred to in FDA (2012b), a guidance document on somatic cell and gene therapy published by the US FDA. According to that guideline, if changes are made to the vector backbone which do not alter the safety properties of the vector and the same route of administration and a similar dosing regimen is used to that employed previously, then truncated testing may be appropriate, depending upon the gene being expressed. When a promoter sequence or a targeting sequence in a viral or nonviral vector is changed, or the expression of viral gene products is considerably altered, the vector should be considered as a new vector, even though it may have the same gene as the previous version of that vector.

19.7.8 Change of Route

It is quite possible that a treatment may be initiated using, for example, intratumoral injections to deliver a gene, which may subsequently be administered systemically. As a considerable amount of relevant information will have already been generated to support the intratumoral route, this would be another case for doing bridging studies. Comparative distribution studies will help to identify how much more safety evaluation may be required.

19.7.9 Insertional Mutagenesis

The long-term management of genetic disorders will require integration of the therapeutic DNA into the host genome, or the maintenance of a stable episomal gene. The target of homologous recombination is still not achievable, and until that time, the problem of insertional mutagenesis—that is, inappropriate insertion of DNA into the host genome—must

be addressed. How can the risk be quantified? Characterization of gene expression over time is more important than the copy number of the gene. An increased copy number can equate to an increased risk of insertional mutagenesis, but on the other hand it also equates to an increase in the desired product. Insertional mutagenesis is a safety problem, and it is important to advise and warn patients who receive genes which will become integrated into the genome of this potential risk associated with their treatment.

19.8 VACCINES

Vaccination against viral and bacterial diseases has been one of the success stories of human and veterinary medicine. Probably the most outstanding example of the effectiveness of vaccination is the eradication of smallpox. In 1967 between 10 and 15 million cases of smallpox occurred annually in some 33 countries. By 1977 the last naturally occurring case was reported in Somalia. Polio too has been controlled in developed countries, for example, the number of cases in the United States was reduced from over 40 000 per year in the early 1950s, before a vaccine was available, to only a handful of cases in the 1980s. Diphtheria is now almost unheard of yet over 45 000 cases in 1940 led to 2 480 deaths from diphtheria in the United Kingdom (similar numbers to those who died from AIDS in the United Kingdom in the entire 1980s). This has been reduced in the United Kingdom to only 13 cases and no deaths from the bacterium between 1986 and 1991. The scale of the problem is enormous—over 10 million deaths worldwide per year are due to infectious disease. The WHO figures suggest that cancers, circulatory problems, and injuries cause fewer deaths in developing countries than infectious diseases.

The process of developing vaccines is becoming increasingly complex due to the nature of the infections being protected against, the nature of the cultures in which the effected individuals live, and societal concerns as to vaccine safety (Plotkin et al., 2008). Kaufmann (1996) provided an excellent addressal of this process and the inherent problems.

19.8.1 Approaches to Vaccination

There are two classical strategies for vaccination. One involves vaccination with either killed pathogenic organisms or subunits of the pathogenic organism. The other utilizes live attenuated viruses or bacteria that do not cause disease but have been derived from the pathogenic parent organism.

Inactivated vaccines are made from virulent pathogens by destroying their infectivity usually with β -propiolactone or formalin to ensure the retention of full immunogenicity. Vaccines prepared in this way are relatively safe and stimulate circulating antibody against the pathogens surface proteins, thereby conferring resistance to disease. Two or

three vaccinations are usually required to give strong protection, and booster doses are often required a number of years later to top up flagging immunity.

Subunit vaccines can be seen as a subcategory of inactivated vaccines because similar considerations apply to subunits and whole organisms. Doses, routes, duration of immunity, and efficacy of these vaccines are all very comparable. In this case a part of the pathogen, such as a surface protein, is used to elicit antibodies that will neutralize the infectivity of the pathogenic agent. The widespread use of hepatitis B virus (HBV) surface antigen purified from the blood of carriers (or more recently from recombinant yeast) shows that this can be a very effective way to immunize. HBV surface antigen, the product of a single gene, assembles into a highly antigenic 22 nm particle which if used in three 40 μ g doses at 0, 1, and 6 months gives virtually complete protection against infection with HBV.

Another example that can be included in the subunit vaccine class is the use of bacterial toxoids. Many bacteria produce toxins which play an important role in the development of the disease caused by a particular organism. Thus, vaccines against some agents, for example, tetanus and diphtheria, consist of the toxin inactivated with formaldehyde conjugated to an adjuvant. Immunization protects from disease by stimulating antitoxin antibody which neutralizes the effects of the toxin.

A further type of vaccine included in the subunit category is the capsular polysaccharide vaccines, for example, those against *Haemophilus influenzae* and meningococcal meningitis. In this case an extract of the polysaccharide outer capsule of the bacterium is used as a vaccine and is sometimes conjugated to protein to improve immunogenicity. Antibody persists for several years and is able to protect against the bacterium.

About half of all vaccines have traditionally been from live attenuated mutants of parent pathogenic organisms (Walker and Gingold, 1993). In effect live vaccines mimic natural infection, yet produce subclinical symptoms and elicit long-lasting immunity often giving rise to resistance at the portal of entry. Most of today's attenuated vaccine strains have been derived by a tortuous often empirical route involving passage in culture until the pathogen is found to lose its virulence. This loss of virulence is tested in animal model systems before being tested in human volunteers. For example, the vaccine used to immunize against tuberculosis was derived after 13 years' passage in bile-containing medium by Calmette and Guérin (hence the name BCG—*bacillus Calmette-Guérin*).

There has been much debate over the past 40 years as to the relative merits of live and killed vaccines often generating more heat than light! The evidence is that both routes will give adequate vaccines that can be used to protect against disease under the appropriate conditions. Table 19.14 summarizes the major points of debate. Many factors

TABLE 19.14 Relative Merits of Live versus Killed Vaccines

| | | Live | Killed/Subunit |
|----------------|-------------------------------------|----------------------------------|-----------------------|
| Production | Purification ^a | Relatively simple | More complex |
| | Cost | Low ^b | Higher |
| | Route | Natural or injection | Injection |
| | Dose | Low, often single | High, multiple |
| Administration | Adjuvant | None | Required ^c |
| | Heat lability | Yes | No |
| | Need for refrigeration ^d | Yes | Yes |
| | Antibody response | IgG; IgA | IgG |
| Efficacy | Duration of immunity | Many years | Often less |
| | Cell-mediated response | Good | Poor |
| | Interference | Occasional OPV only ^e | No |
| | Reversion to virulence | Rarely ^f | No |
| Safety | Side effects | Low level ^g | No |

^aIncreasing safety standards mean that for new vaccines some of the older methodologies would not be acceptable.

^bThe price for new vaccines will approach that of killed subunit vaccines as safety standards are increased.

^cVery few adjuvants for human use are acceptable.

^dThe need for refrigeration increases the costs significantly.

^eEspecially in the Third World.

^fAt very low levels (<1 case per 10⁶ vaccinations).

^gThis varies from occasional mild symptoms with rubella and measles vaccines to possible brain damage with pertussis vaccine.

including cost, safety, number of immunizations, ease of access to vaccines, politics, and social acceptance will determine whether there is a high uptake of a particular vaccine and whether it is ultimately successful in eradicating the target disease. Even if a perfectly viable, relatively safe vaccine is available, uptake may be limited. For example, it has been estimated that vaccination against measles within the WHO EPI has prevented over 60 million cases and 1.37 million deaths. Despite these efforts there are still some 70 million cases of measles annually, resulting in nearly 1.5 million deaths; consequently a recent WHO congress adopted the following goals:

1. Increasing immunization coverage
2. Improving surveillance
3. Developing laboratory services and improving vaccine quality
4. Training
5. Promoting social mobilization
6. Developing rehabilitation services
7. Research and development

This again also serves to illustrate the importance of factors other than the efficacy of the vaccine itself in disease prevention.

The single most important issue in developed countries is the safety of a vaccine; a single death in a million vaccinations for a new vaccine would be unacceptable (except possibly if it were an effective AIDS vaccine). While this is obviously important in a Third World country, other issues

such as cost and how to deliver the vaccine are of paramount importance.

19.8.2 Genetic Engineering and Vaccine Development

Not all protective antigens are as simple to identify, clone, and express as the surface antigen gene of hepatitis. The entire sequence of the HBV genome became available, and as it is less than 10kb, it was relatively simple to establish which open reading frame to express. It has been known for many years that irradiated malarial sporozoites protect against malaria. As the sporozoite stage in the life cycle of the malarial parasite can only be grown in small quantities, it was left to rDNA technology to identify, clone, and express components of the sporozoite that might be of use in vaccine production. The genome of the malarial parasite is many thousands of times larger than the genome of HBV and therefore provides a different scale of problem; not only was there little sequence data available, but there was also no idea of which gene products may be protective.

The starting point of any rDNA work is to generate a library of DNA in *E. coli* which is representative of the organism under study. Once having a cDNA bank or a genomic library, there are three basic ways of identifying and isolating a gene of interest.

19.8.2.1 DNA/Oligonucleotide Hybridization If there is some preexisting knowledge of the nucleic acid sequence, or where purified mRNA is available, it is possible to detect recombinant clones by hybridization of ³²P-labeled DNA or RNA to bacterial colonies or bacteriophage plaques. Often a

protein has been purified, and some amino acid sequence is available which allows a corresponding nucleic acid sequence to be synthesized. Due to the degeneracy of the genetic code, a complex mixture of oligonucleotides is required to ensure that all possible sequences are represented. Labeling this mixture of oligonucleotides yields a probe that can be used to screen a cDNA (or possibly genomic) library that might be expected to contain the gene of interest.

19.8.2.2 Hybrid Selection and Cell-Free Translation A second approach is to use hybrid selection of mRNA coupled with cell-free translation. DNA clones from a library, either individually or in pools of clones, can be immobilized by binding to a solid support and mRNA hybridized to them. Only the mRNA that corresponds to the clones will bind, and this can then be eluted and translated to protein in a cell-free system. The protein can then be immunoprecipitated with antisera to the gene product of interest or assayed for activity. An example that encompasses both this approach and the sequence route is in the development of a vaccine for Epstein-Barr virus (EBV). It had been known since 1980 that antibody to the major membrane antigen of the virus (gp350/220) would neutralize the virus. Around 1983 a fragment of the virus genome was cloned and sequenced; using computer predictions, the gp340/220 gene was identified. The experimental evidence that confirmed this prediction was published in 1985 and came from experimental work that managed to hybrid select EBV mRNA using genomic DNA clones. This was followed by cell-free translation of the eluted mRNA and immunoprecipitation of gp350/220 with a high-titer antibody. The DNA clone that hybridized with the gp340/220 mRNA was the one predicted to encode the gp340/220 gene by computer analysis. The hybrid selection approach is rather labor intensive and has for the most part been superseded by one of the forms of expression cloning.

19.8.2.3 Expression Cloning This approach is invaluable when the only means of identification are antisera against the protein or pathogen of interest.

Probably the most laborious form of this approach is its use in conjunction with a biological assay. cDNA libraries are cloned into a plasmid that will allow expression in eukaryotic cells, for example, SV40 or EBV vectors. Clones or pools of clones are then transferred to appropriate cell types, for example, COS cells for SV40 vectors and cell extracts or cell supernatant are assayed for biological activity. If a pool of clones gives the biological activity, then the individual clones can be reassayed and the desired cDNA clone identified. This methodology although tedious has allowed many of the interleukin genes to be cloned probably because the assays for these proteins are very sensitive.

Other gene products or vaccine antigens may require an enrichment step. For example, many genes expressed on the cell surface, that is, receptors, adhesion molecules, etc., have been cloned by “panning” techniques where the cells expressing the gene of interest are selected out either with antibody or by interaction with other cells. cDNA libraries are constructed in *E. coli* and the library is transferred to eukaryotic cells. Those cells expressing the gene of interest are enriched for and the library transferred back to *E. coli*; this can be done for several rounds of expression, and eventually individual clones conferring the selected phenotype will be isolated.

The most extensively used form of expression cloning involves the use of plasmid or bacteriophage vectors in *E. coli* and identification of DNA clones using antisera to the protein of interest. Here a vector such as the bacteriophage λ gt11 is set up so that when cDNA fragments are cloned into sites adjacent to the β -galactosidase, gene bacteria will express a β -galactosidase fusion protein containing epitopes present in the cDNA. Recombinant phage is detected with antisera. The cDNA insert is then sequenced and the whole gene can then be isolated in a more traditional way. The antisera used can be mAbs, polyclonal monospecific antisera, or even polyclonal antisera with many antibody specificities present. A variation on this method allowed the initial cloning of the malarial sporozoite surface antigen. Malarial sporozoite stage cDNAs were introduced into the ampicillin resistance gene of the plasmid pBR322. Low levels of expression of the sporozoite surface antigen were detected by solid-phase radioimmunoassay using an mAb specific for the protein. In this way a cDNA clone coding for the antigen was isolated and subsequently sequenced. This information was then used to design peptide vaccines which have already been tested in humans.

The λ gt11 system is a more sophisticated bacteriophage version of the plasmid system described earlier and has been used to isolate many different antigens from various stages in the life cycle of the human malarial parasite using human immune sera as well as antigens from pathogens.

19.8.2.4 Expression of Potential Vaccine Antigens In general, in the future eukaryotic cell culture is likely to be the method of choice for the production of subunit vaccine antigens where the organism to be vaccinated against replicates in eukaryotic cells. *E. coli* are unable to post-translationally modify some vaccine candidates, for example, bacterial systems cannot add carbohydrate which is important in the antigenicity and structure of many protective antigens.

Since 1986, the US FDA has approved 22 vaccines (Table 19.15), half of them from a genetic engineering (and all, of course, from a biotechnology source). The cells used for such genetic engineering production of vaccine can be mammalian, insect, or bacterial.

TABLE 19.15 Vaccines Approved by US FDA Since 1986

| Vaccine | Indication | Date Approved | Company |
|--|---|---------------|--|
| Recombivax HB | Hepatitis B | June 23, 1986 | Merck, Chiron |
| ProHIBIT | <i>Haemophilus influenzae</i> B | 1988 | Connaught |
| Pedvax | HIB | 1989 | Merck |
| Engerix-B | Hepatitis B | 1989 | SmithKline Beecham |
| Tetramune, Hib TITR <i>Haemophilus b</i> , diphtheria CRM 197 protein conjugate | Bacterial meningitis | Jan. 1991 | Lederle-Praxis Biologicals/ American Cyanamid |
| IPOL | Poliovirus vaccine inactivated/injected | 1991 | Inst. Merieux |
| Acel-Imune | Diphtheria, tetanus toxoids, and acellular pertussis vaccine | Jan. 6, 1992 | Takeda Chemical Industries/ American Cyanamid |
| Tripedia | Diphtheria, tetanus toxoids, and acellular pertussis | Aug. 20, 1992 | Connaught Laboratories Inc. |
| JE-VAX | Japanese encephalitis | Dec. 18, 1992 | Connaught/Biken |
| Enzon | Bubonic plague | 1994 | Green Labs |
| Typhim Vi | Typhoid | 1994 | Laboratories Inc. |
| Havrix | HAV | Mar. 1995 | SmithKline Beecham |
| Varivax, Varicella Virus Vaccine Live | Chickenpox | Apr. 10, 1995 | Merck |
| VAQTA | Hepatitis A | Mar. 29, 1996 | Merck |
| COMVAX | <i>Haemophilus b</i> and hepatitis B | Oct. 2, 1996 | Merck |
| Infanivir | Diphtheria, tetanus, and pertussis (DTP) | Jan. 27, 1997 | SmithKline Beecham |
| RabAvert | Rabies (pre- and postexposure) | Oct. 27, 1997 | Chiron Behring |
| Certina | DTP | Jul. 29, 1998 | North American Vaccine |
| RotaShield ^a | Rotavirus | Aug. 31, 1998 | Wyeth |
| LYMERix | Lyme disease | Dec. 21, 1998 | SmithKline Beecham |
| Prenar | Pneumococcal disease | Feb. 17, 2000 | Lederle |
| TWINRIX | Hepatitis A and B | May 11, 2001 | SmithKline Beecham |

^aSubsequently withdrawn.

The CBER has provided broad guidelines on the evaluation and production of vaccines (CBER, 2000). In general, the center's requirements have paralleled those for other biotechnology products. Beyond establishing sterility and lack of pyrogenicity and of viral contaminants, a single GLP toxicity study in an appropriate species (one that has been established, if possible, to be immune responsive to the vaccine) is required. If the vaccine is to be used in pregnant women or women of childbearing potential, a segment I style reproductive study should be performed in an appropriate animal species. A common concern that may require specific evaluation is dose site reactions (local site responses). It must be remembered that it is not possible to evaluate the safety of a vaccine uncoupled with its intended adjuvant (Brennan and Dougan, 2005).

Regulatory guidance for the conduct of clinical trials on vaccines is specific. Traditional phase I trials in normal volunteers are not conducted. Rather, all trials assess not only safety but also efficacy (or at least immunogenicity). Trials may well be challenge trials—that is, after immunization subjects are purposely challenged with exposure to the infective agent of concern.

In any case, injection site responses (erythema, edema, pain, and tenderness) and systemic responses are both evaluated in subjects (Mathieu, 1997).

The US FDA also has specific guidance on the tracking and reporting of adverse clinical responses to vaccines.

Any adverse events of product problems with vaccines should not be sent to MedWatch but to the Vaccine Adverse Event Reporting System (VAERS), operated jointly by the FDA and the national Centers for Disease Control and Prevention. For a copy of the VAERS form, call 1-800-822-7967, or download the form (in PDF format) from www.fda.gov/cber/vaers/vaers1.pdf on the FDA's website.

19.9 SPECIAL CHALLENGES

The problem with using a classical toxicological approach for evaluating an rDNA product or species-specific protein is that standard protocols are probably inappropriate and nonrelevant in most cases. In the traditional approach to toxicology, a standard protocol or battery of tests is performed, followed by an estimation of the types of and degree of hazard or risk to humans. For example, conventional toxicity testing of a new rDNA product might lead to evaluation at excessively high doses in two rodent species. The production of antibody in the test species during preclinical toxicology testing may inactivate the test compound and thus invalidate the toxicity evaluation.

This approach appears somewhat irrational and without much scientific merit, since many of these new molecules are minimally toxic or nontoxic by this sort of acute evaluation. As in the case of IFNs or mAbs, the toxic effects observed in humans might not be predicted from safety assessments in rodents. An appropriate test species should be selected. Is the rat or mouse the appropriate species to evaluate a species-specific rDNA protein such as HGH or IFNs, or would nonhuman primates be more suitable? Does the nonhuman primate really offer any advantages? There is some consensus that the nonhuman primate may be a more appropriate species for testing some rDNA human proteins.

In contrast, in the "pharmacological approach" to toxicology, the potential targets of toxicity are first identified (Zbinden, 1986). The criteria for relevant effects are established, usually based on experience with reference substances, and appropriate *in vivo* or *in vitro* experimental models are selected to assess the pertinent toxicological responses.

Doses should be selected that are reasonable multiples of the proposed therapeutic dose to be employed, especially since in many cases the amount of material available for testing may be limited and not available in kilogram amounts. Preclinical rodent or primate studies should merely provide the flags to monitor during phase I clinical trials. Reason should prevail not only in the selection of methods and models for assessing the potential toxicity of the new agents but also in the use of these data for extrapolation to man. Whether US industry succeeds or fails in the biotechnology arena will depend on the quick resolution of issues such as selection of appropriate toxicologic tests, fermentation scale-up of the rDNA microbe, product purity, and expedition of regulatory pathways.

19.9.1 Purity and Homology

Major concerns in the production of a species-specific protein by rDNA technology are the purity of the product, the amount and type of contaminants present, and the homology of the product to the native molecule (Table 19.16). The toxicologist should be concerned about the acceptability

TABLE 19.16 Issues in the Safety Evaluation of Species-Specific rDNA Products

| |
|---|
| Purity |
| Homology to native molecule (amino acid sequence, extra amino acids, three-dimensional structure) |
| Type and amounts of contaminants (chemicals, <i>E. coli</i> proteins, fermentation products, foreign DNA) |
| Stability of clone |
| Immunogenicity |
| Toxicities (direct or secondary to therapeutic effect) |

and toxicity of intentional or inadvertent contaminants introduced during fermentation or isolation of the product (e.g., DNA, chemicals, *E. coli* proteins). Other issues concern the introduction of amino acid residues that might alter the three-dimensional structure or antigenicity of the molecule, partial denaturation of the product during isolation and recovery, genetic stability of the rDNA clone during production (mutation could result in altered amino acid sequence), and the level of foreign DNA present. Although these are issues of analytical biochemistry, their impact on the potential toxicity and overall safety of the finished product is of some toxicological concern.

19.9.2 Immunogenicity

The problem of the immunogenic nature of many human rDNA proteins, and the potential to generate antibodies to a normal human protein, is of special interest to the immunotoxicologist. For example, 3 of 16 patients administered the rDNA-derived IFN- α (clone A) developed antibodies of the IgG class that were undetectable prior to or during therapy (Guterman et al., 1982). These antibodies were capable of *in vitro* neutralization of IFN activity, although *in vivo* neutralization of IFN was not documented. Since there are several different subtypes of IFN- α s containing epitopes not present on their own IFN subtype. Similarly, two patients treated with IFN- β for many months developed high-titered antibody, which on one case was correlated with an inability of the patient's fibroblasts to produce IFN (Vallbracht et al., 1982).

Virtually all patients treated with conventional porcine insulin develop circulating anti-insulin antibodies (Klaff et al., 1978) that are less frequent and in lower titer in individuals treated with more highly purified porcine (Falholt, 1982) or rDNA human insulin (Fineberg et al., 1983). In the study by Fineberg and associates, 44% of the patients developed antibodies to rDNA human insulin over a 12-month period compared to a 60% antibody frequency with porcine insulin. HGH (Genentech, Inc.) prepared by rDNA technology was observed to produce a frequency of immunogenicity similar to that seen with human insulin (~40% of the children developed antibody, according to the product insert). The ultimate goal is to develop rDNA products that will be less immunogenic than purified animal sources of these therapeutic agents.

The exact mechanism of the immunogenicity of species-specific rDNA proteins is unknown, but is believed to be attributable to (i) the addition of extra amino acid residues during synthesis, which the host reads as foreign; (ii) denaturing of the native molecules; or (iii) contamination by *E. coli* polypeptides or lipopolysaccharides.

A second unanswered concern is whether the antibody induced by the recombinant protein has any discernible health effect. Other than some reports of neutralization of biological activity, little pathology has been attributed to the

presence of antibodies in patients given recombinant protein therapy. It should also be noted that the question of antibody specificity has not been well studied, so that it is entirely conceivable that autoimmune pathology or even an anaphylaxis response could be induced. Equally important is the concern that induced antibody might neutralize the endogenous hormone or protein that it is intended to replace or supplement.

A third consideration is that certain routes of administration may favor immunogenicity of recombinant proteins. In early trials, rDNA proteins introduced by subcutaneous or intramuscular injections (procedures known to improve the immunogenicity of proteins) resulted in a higher frequency of antibody responses than in the IV route.

In summary, these are the clinically relevant questions about the immunogenicity of rDNA species-specific proteins: will antibody be induced in the recipient that will neutralize the therapeutic effect or lead to immune complex disease? What is the class (e.g., IgG or IgE) and specificity (i.e., reactivity against specific protein or contaminant) of the antibody induced? The former antibody type could potentially neutralize the product and produce immune complex disease, while the latter could result in an anaphylaxis response. It is possible that the antibody induces is of insignificant health consequence, and its presence is known only because of improvements made in the sensitivity of detection methods with the introduction of the ELISA.

19.10 PLANNING A SAFETY EVALUATION PROGRAM

Safety evaluation of a candidate product should start with a consideration of specific nature and consequential hazards of the three Ps:

- The producing system
- The process
- The product

The need under each heading is to decide what data are required, then how to obtain them with the greatest efficiency and economy, and last, whether the toxicologist is necessarily the person with the appropriate skills and experimental techniques to do so. There will often be a trade-off between precise control by other means and possibly cheaper or more familiar, old-fashioned toxicological studies (Cavagnaro, 2008; Dorato et al., 2014). The inventor of a new product or process, too, may often have to do a great deal of work to show safety by excluding hypothetical hazards, which subsequent manufacturers can afford to ignore.

19.10.1 The Producing System

The questions of particular concern here are the nature of the system used to manufacture the desired substance and the precision with which it is controlled.

If the system consists of prokaryotic cells, then how well defined is their provenance and how is their consistency demonstrated? If mammalian cells are employed, their lineage must be considered. In both instances, it is important to ensure that extraneous virus, infections, DNA, and less well-defined factors such as “slow viruses” are excluded by the origins and history of the producer strain, or because the physical (e.g., filtration) or chemical (pH, solvents, affinity separation) nature of the production process can be relied upon to exclude passage of an infectious agent.

If the degree of safety arising from these factors is weak, the toxicologist should consider appropriate studies *in vivo* to exclude contaminating agents, oncogenic factors, etc., but there is no point in doing short-term or prolonged animal experiments or other types of test unless the desired end point has first been clearly defined.

The aspect to which far more attention has been directed is the nature of the inserted gene(s) and promoters in rDNA products. Again, the toxicologist should ask how well the nucleotide sequence is known, whether there is only one reading frame, and how are any introns handled. Again, toxicity-type testing would appear to be an inefficient and expensive way to study molecular biology and biochemistry.

Last under this heading, for intact infectious organisms to be used directly in man, is assessment of pathogenicity to the range of individuals that make up our populations, the possibility of reversion to a wild and more dangerous strain, the hazard of an allergenic reaction to the organism (e.g., vaccinia), and the possibility of spread from subject to subject in a naïve or incompletely immune population.

There may be some role for animal experimentation here, if there is a suitable model, because it gives the chance to study the organism under intense pressure from commensals and the rising immune response. Possible hazards in the manufacturing plant also need to be evaluated.

In general, conventional toxicity procedures seem to have little to offer here, except in specific instances of helping to exclude certain infections factors, perhaps ruling out oncogenicity, and examining the stability of engineered organisms for direct infection of man.

19.10.2 The Process

The toxicologist has the least to offer here. In fact, only the toxicologist’s intellectual analysis and review of the literature should be required to assess the manufacturing process and any residues of its chemicals and so to set analytical limits on purity and residues in the final preparation.

19.10.3 The Product

There are two distinct and probably divergent forces affecting the way in which the toxicologist regards the final product. One is whether the toxicologist's opportunity is the first opportunity and therefore need to learn about its biology and pharmacology, and the other is, what do scientific concern, clinical caution, and industrial pace define as the minimum it is reasonable to do before clinical trial or marketing?

19.10.4 Biology of Bioengineered Products

This may not be a useful concept scientifically, but it represents the practical point that the pace of development often forces the rapid sequence—interesting biological property, identification of responsible molecule in very small amounts (e.g., tumor necrosis factor or erythropoietin), cloning, etc.—resulting in large-scale production, perhaps even before the full structure is known. The clinical interest in administering the substance to man for investigative or therapeutic purposes must be balanced against the total lack of knowledge of its general effects on the body or of the consequences of prolonged high-level exposure to it of the function of interest.

As examples, there is the history of IFN, discovered through its antiviral actions, subsequently found to modulate mitosis and certain immune functions, capable of producing fever and probably ECG and EEG changes as evidence of membrane effects in excitable tissues. If an IFN were a novel discover, now just being produced for the first time, then investigation of its general biological effects on repeated administration to responsive animals would be important prior to study in man. The same arguments apply to other lymphokines, for example, α -IFN, IL-2, and so on.

The planning of this type of investigation as an empirical open study of responses must be carefully related to the nature and what is known of the product concerned:

1. It is necessary to work in a species capable of responding to the principal activity. IFNs are notorious for their species specificity, but most other lymphokines at least are more generally active. Work in a primate may be required, but it depends on the substance to be tested. There may be no point in working in using more than one species in pivotal studies.
2. Any test should be as broad and as general as possible, that is, monitor many variables clinically, in the laboratory and by pathology, until enough is known for there to be confidence in a focused approach.

Relate any testing to the clinical circumstances of probable use.

Thus, if a new synthetic antigen or engineered antigen is for testing, for administration only a few times to man, there would be no point in a multiple-dose

experiment. It should suffice to show that it was antigenic in the intended preparation. Unless there were a prior reason to do so, a special search for, say, autoimmune reactions, seems unnecessary. Similarly, testing an mAb for activity is likely to be difficult if not impossible, because of species specificity and the antigenicity of the preparation.

3. The toxicologist should be prepared to do nothing if the material is well known, its properties are understood, and there is adequate characterization of the nature of the preparation supplied; for example, human insulin or growth hormone produced by genetic engineering should not be submitted to prolonged safety tests in animals, provided that the molecular forms present are sufficiently well understood.

It may be useful, however, to consider limited animal studies to examine the pharmacokinetics and duration of action even of a well-known material made by a new route, unless physiochemical analyses show that to be pointless.

19.10.5 Animal Models

Species selection is probably one of the most important considerations when designing a preclinical safety program; for a biotechnologically derived pharmaceutical, it requires an understanding of the biology of the product. Since most of these either are human proteins or target human receptors, they tend to be species specific. Studies in rodents and dogs, the species commonly used in traditional toxicity studies, may not provide scientifically meaningful data. However, nonhuman primates are not necessarily the most appropriate species either, despite their phylogenetic similarity to human beings.

Some approaches that offer guidance in selecting relevant species include a literature review; determining the extent of homology between the endogenous animal protein and the human recombinant protein; determining the activity of the protein in pharmacological models; and *in vivo* assays of the receptor/tissue binding.

A literature review may provide useful information about the physiological properties of the protein in animals and how they compare with those of the human protein in man. For example, prior to rDNA technology, growth factors and/or hormones were purified from biological fluids. Although the quantities obtained were limited, they were nevertheless sufficient to allow investigation of the physiological properties of these proteins. Computer programs are now available for online searching of databases which hold information not only on the sequences of various animal and human proteins but also the extent of homology between an animal protein and its human equivalent, including common amino acid sequences. It should be remembered, however, that a protein showing a high degree of homology to the human protein may not necessarily share similar pharmacological

activities. Evaluation of activity or lack of activity in pharmacological animal models, if available, certainly would aid species selection. Finally, *in vitro* assays which analyze receptor and/or tissue binding are commonly used to determine the appropriate species for preclinical safety evaluation.

Some biotechnologically derived pharmaceuticals will cross-react with species that can be evaluated toxicologically, while others cross-react only with nonhuman primates such as the chimpanzee—a protected species. In this case, a well-designed “safety” or “phase 0” study at doses higher than the proposed clinical dose may provide valuable safety information. However, a lack of cross-reactivity with any nonhuman species does not necessarily make preclinical safety evaluation impossible, nor does it limit toxicity testing to species in which the protein lacks relevant pharmacological activity. Some alternative possibilities are summarized in Table 19.17.

Toxicity studies traditionally are conducted using “normal” animals. However, studies in animal disease models may provide additional safety information regarding the possibility of disease exacerbation. For example, the administration of human recombinant erythropoietin was associated with hypertension in patients with chronic renal failure, and also in uremic dogs, but not in normal dogs.

Species differences must be considered when choosing a model and, in particular, species-specific immunological differences between the human and the test animal. For example, in man, an anti-CD4 mAb will bind to CD4 expressed on monocytes, with subsequent fixing of complement and destruction of antigen-presenting cells. However, since CD4 molecules are not expressed on murine monocytes, these effects would not be evident in a murine model.

For highly humanized proteins, the approach to proper safety evaluation starts with the identification of an appropriate nonclinical safety model.

First, evaluate for comparative tissue binding; then if no appropriate species is identified, then one can:

- Either prepare and test a homolog
- Or test molecule itself in a humanized mouse model

A tissue cross-reactivity study is required by the FDA “points-to-consider document” for mAbs prior to the first clinical trial. Its usefulness is debated, but its purpose is to evaluate the potential for binding to nontarget tissues. Positive and negative controls are important for interpreting results. Tissues from humans and all nonclinical species used in safety studies should be evaluated. At minimum, the evaluation should include the 32 tissues recommended by the FDA “Points to Consider in the Manufacturing and Testing of Monoclonal Antibody Products for Human Use.”

When is it appropriate to use a homolog? According to ICH S6 guidance, “When no relevant species exists, the use of homologous proteins should be considered. While useful information may also be gained from the use of homologous proteins, it should be noted that the production process range of impurities/contaminants, pharmacokinetics, and exact pharmacological mechanism(s) may differ between the homologous form and the product intended for clinical use.”

Comparability of the homolog with the clinical candidate is critical:

- Characterize pharmacology
 - Literature: What’s known about the target in the test species compared to humans?
 - *In vitro* binding—similar affinity or neutralization?
 - Functional assays
 - *In vitro* cells
 - *In vivo* bioassays (if possible relevant)
 - Similar tissue distribution (tissue cross-reactivity for mAbs)
- Pharmacokinetics
- Is Fc activity important and similar?

Challenges of homologs

- They represent a second test article.
 - The decision must be made early in development.
- Sometimes, it is not possible to make a homolog.

If possible, months to years needed to develop construct, clone, manufacture material, characterize pharmacology, and/or establish bioanalytical support

TABLE 19.17 Alternative Models for Toxicity Assessment

| Model | Example | Caveat |
|------------------------------------|--|--|
| Nontraditional animal model | Transgenic mice carrying appropriate human receptor | Antibody formation would need to be monitored, as it is probable that a large human protein would produce an immune response |
| Homologous proteins and/or systems | Testing purified animal protein in the same species or, for monoclonal antibodies, testing an antibody directed against the receptor in the animal | Data should be interpreted with caution as the biological properties of the animal protein may differ from those of the human protein |
| <i>In vitro</i> methods | Tissue binding assays | If no <i>in vivo</i> models are available, <i>in vitro</i> methods combined with <i>in vivo</i> testing in a pharmacologically nonreactive may suffice |

May be immunogenic, thus limiting usefulness

- How do you interpret the data?
 - No “validation” can be performed that homolog is predictive of human toxicities.
 - What if findings are different from the clinical candidate in an appropriate toxicology species?
 - How do we extrapolate safety margins to the clinical candidate?

Regulatory challenges of homologs

- No common criteria for what’s expected
 - How much comparison with the clinical candidates is enough?
 - Expectations for analytical characterization and does this need to be comparable?
 - Do all aspects of testing need to be GLP?
- Can studies with a homolog replace studies with the clinical candidate?
 - Development and reproductive testing
 - What if results are more severe than with the clinical candidate?
- Are negative findings meaningful?
 - It’s not your clinical candidate, so does the data impact risk assessment?

Homologs have been used to support registration for infliximab (Remicade®) (anti-TNF) and efalizumab (Raptiva®) (anti-CD11a). In both cases, there was not an appropriate species for the clinical candidate. Chimpanzees were the only pharmacologically responsive species, but they are not acceptable for toxicity testing due to humane reasons:

“Humanizing” mice

- Isolate CD34+ stem cells (hHPCs) from human cord blood
- Breed NOD. Cg-Prkdc^{scid}//2rg^{tmWjl}
- PND1 pup irradiated 1 cGy
- hHPCs (10⁵) injected in 25 µL PBS into the liver
- Assay tail blood for reconstitution of human immune cells

Humanized mice can be used to:

- Assess the *in vivo* influence of stressors and drugs on the development of immune cells
- Evaluate how toxicants can modify *in vivo* human immune responses
- Investigate whether hematopoietic stem cells from cord blood can be induced to develop into nonimmune cell types
- Determine the influence of mouse effects on human immune cells longevity and human immune cells on mouse longevity

19.10.6 Study Design

It is questionable whether traditional toxicological paradigms are applicable to biological or protein agents. If they are not, then how can the clinician gain reassurance to administer the first dose to man, to move into multidose trials, and to even assess the agent in combination with other established medicines or biological agents? mAbs, soluble cytokine receptors, and growth factors all have been used in patients for nearly a decade, providing a wealth of experience in this area from which to learn. One of the most striking lessons is that PD effects may appear long after dosing of the agent has been discontinued.

As a class, biotechnologically derived pharmaceuticals share certain characteristics which have influenced their preclinical development. They are proteins and therefore toxicity was expected to be minimal and limited to an exaggeration of their desired pharmacological effects, a myth which was ultimately exploded. These agents are designed to perturb specific molecular or cell-to-cell interactions, sometimes with minimal effect on the pathophysiology of the target disease. Owing to the species-specific nature of these agents, preclinical toxicology is usually limited. For example, if a primatized anti-CD4 mAb cross-reacts only with chimpanzee and human CD4, the species of choice for toxicity tests is the chimpanzee, the use of which is restricted by its limited availability.

These characteristics of protein agents give rise to problems in clinical development, such that the traditional paradigm for preclinical testing may not be appropriate. The dose in animals may not be predictive of an appropriate starting dose for man. A surrogate marker, for example, CD4 cell counts in the preclinical chimpanzee model, may be useful in setting the initial human dose but may only serve to indicate a no-effect dose. Once in the clinic, trials conducted early in development are usually not sufficiently powered to distinguish effects due to the toxicity of the test agent, from those due to, for example, the underlying disease and concomitant or previous medications. Finally, short-term (3–6 months) preclinical studies do not necessarily predict the long-term effects of these agents.

The long-term toxicities of concern are opportunistic infections, lymphoproliferative disorders, and immunogenicity, manifesting as tachyphylaxis and/or allergic reactions. Preclinical approaches which serve to identify these as potential hazards to man of a biologic drug moiety are thus needed.

The choice of toxicity studies, and the design of individual studies, will depend on the proposed clinical program. Important issues to consider are:

- The frequency and route of administration, including the use of novel delivery systems
- The duration of dosing
- Special toxicity testing

19.10.7 Frequency and Route of Administration

Clinical trials for biotechnologically derived pharmaceuticals may be more complex than those for conventional pharmaceuticals, and so the route and frequency of test drug administration should, if possible, mirror the proposed clinical use, even if that route employs a novel delivery system.

19.10.8 Duration

Traditionally, the duration of a toxicity study depends on the intended clinical use and disease duration. The potential immunogenicity of the human protein is a significant issue since antibody binding can partially or completely inhibit the biological activity of that protein, affect its catabolism, or alter its distribution and clearance. Any multiple-dose study therefore should include evaluation of the impact of antibody formation, including their neutralizing capacity. However, antibody formation in itself should not be a reason for termination of a toxicity study, particularly if the antibodies are not neutralizing or do not alter the pharmacodynamics of the protein.

Multiple-dose toxicity studies are usually conducted before single-dose administration to volunteers. Many of the clinical trials for biological agents target life-threatening illnesses, and it has therefore been suggested that single-dose toxicity studies are sufficient to support single-dose “proof-of-concept” clinical studies. While this approach promotes faster introduction into the clinic, it may be of limited use since, once in many, there is a tendency to overlook the preclinical data. Clinical development may not progress without interruption if relevant preclinical data are missing.

19.10.9 Special Toxicity Testing

In addition to multiple-dose studies, information on potential functional changes—as obtained from safety pharmacology studies—and the potential for genotoxicity, reproductive toxicity, and carcinogenicity may be required for registration. Once again, the species specificity of recombinant proteins may preclude the use of traditional animal species such as rodents and/or rabbits for safety pharmacology, reproductive toxicity, and carcinogenicity studies. Functional evaluations of cardiovascular and pulmonary systems could be incorporated into a nonhuman primate multiple-dose toxicity study. If appropriate, potential reproductive toxicity can be evaluated in a nonhuman primate.

There may be situations which warrant an assessment of carcinogenic potential, but immunogenicity and species specificity may preclude a 2-year rodent bioassay. It may be necessary to develop *in vitro* assays to address a particular concern. For example, growth factors which may have the

potential to support or stimulate the growth of transformed cells should be assessed for their ability to promote growth of either malignant or normal cells.

High molecular weight compounds are unlikely to react with DNA or other chromosomal material, and therefore a genotoxicity evaluation may be of little value. However, genotoxicity studies may provide useful information about the safety of products containing organic linkers.

19.10.10 Program Design Considerations

The standard toxicological data package for any new drug entity typically evaluates:

- Potential toxicity following single and multiple dosing
- Genotoxic potential
- Functional changes, that is, safety pharmacology studies

In addition, depending on the proposed clinical plan, the following may need evaluation:

- Toxicity following chronic dosing
- Carcinogenic potential
- Possible reproductive toxicity

Although biotechnologically derived pharmaceuticals often need customized preclinical development programs, certain issues are common to all. These include species specificity, potential immunogenicity and its impact on the duration of dosing, and the need for special toxicity testing.

19.11 CHALLENGES: BIOSIMILARS

For small molecules, the availability of generic versions after patent expiration has significantly contributed to keeping the cost of many (particularly chronic use) medications down. Such generics have identical chemical structures to those of the marketed active molecules, and demonstrating that structural equivalence is readily done by available analytical techniques.

For biologics, however, where the drug entity is produced by modified cellular processes and the precise complex molecular structure cannot precisely be defined and therefore an identical structure cannot be either achieved or proven. So it is not possible to produce “generic” biologic drugs, even after patent expiry. Though guidelines and laws have been enacted in the EU and the United States (the “Biologics Price Competition and Innovation Act” (BPCI) enacted in 2009) to provide a path to production (biologically sourced therapeutics that are equivalent in terms of efficacy and safety, “biosimilars”) and a specific guidance was enacted by the FDA (2012a), progress has been slow.

TABLE 19.18 Nonclinical Studies with Marketed Biosimilars

| Biosimilar versus Innovator Name/Identity/ Action/Authorization Date/Reference ^a | Biological Activity/ Pharmacology | Kinetics | Toxicology |
|---|--|--|---|
| Growth factor | | | |
| Omnitrope versus Genotropin/recombinant human growth somatropin/growth disorders/2006 | Yes (<i>in vivo</i>) | TK from toxicology work | Repeat-dose rat toxicology + rabbit local tolerance studies |
| Valtropin versus Humatrope/recombinant human growth somatropin/growth disorders/2006 | Yes (<i>in vivo</i>) | PK in rabbit + TK from toxicology work | Repeat-dose rat toxicology study |
| Epoetin | | | |
| Retacrit and Silapo versus Eprex/Erypo/epoetin alfa/anemia/2007 | Yes (<i>in vitro</i> and <i>in vivo</i>) | TK from toxicology work | Repeat-dose rat and dog toxicology + rabbit local tolerance studies |
| Binocrit, epoetin alfa Hexal and Abseamed versus Eprex/Erypo/epoetin alfa/anemia/2007 | Yes (<i>in vitro</i> and <i>in vivo</i>) | PK in dog + TK from toxicology work | Repeat-dose dog toxicology + rabbit local tolerance studies |
| Granulocyte colony-stimulating factor | | | |
| Biograstim, Ratiograstim, Tevagrastim, and Filgrastim ratiopharm versus Neupogen/human G-CSF/neutropenia/2008 | Yes (<i>in vitro</i> and <i>in vivo</i>) | PK in rat and monkey + TK from toxicology work | Repeat-dose rat and dog toxicology + rabbit local tolerance studies |
| Filgrastim Hexal and Zarxio versus Neupogen/human G-CSF/neutropenia/2009 | Yes (<i>in vitro</i> and <i>in vivo</i>) | TK from toxicology work | Repeat-dose rat toxicology + rabbit local tolerance studies |
| Nivestim versus Neupogen/human G-CSF/neutropenia/2010 | Yes (<i>in vitro</i> and <i>in vivo</i>) | TK from toxicology work | Repeat-dose rat toxicology + rabbit local tolerance studies |

^aData from European Public Assessment Report for each biosimilar (EMA, 2016).

As of the end of 2015, only one drug (Zarxio, a biosimilar for Amgen's Neupogen which was first approved in 1991) has been so approved in the United States. The process of approval requires:

1. Analytical studies demonstrating that the biologic product is highly similar to the marketed ("reference") product.
2. Animal studies showing equivalence of safety and therapeutic activity.
3. One or more clinical studies to demonstrate safety, purity, and potency under reference drug label conditions of use. Such studies must assess and show equivalence for immunogenicity and either pharmacokinetics or pharmacodynamics.

The animal toxicity studies should be designed to provide a comparative evaluation of the reference and biosimilar product in a relevant animal species. As such, either PK or PD measures should be obtained to support a determination of equivalence (Baldrick and Donniger, 2012).

Examples of such nonclinical studies used to support marketing approval of biosimilars in Europe are presented in Table 19.18. The biosimilar approval process in the EU has moved along very quickly.

REFERENCES

- Alkon C. (2015) Doctors Unwilling to Trust Similarity of Biosimilars. *Oncology Business News*, June 2, 2015. Available at: <http://www.onclive.com/publications/oncology-business-news/2015/june-2015/doctors-unwilling-to-trust-similarity-of-biosimilars/1> (accessed June 1, 2016).
- Baldrick P, Donniger R. (2012) Nonclinical aspects of biosimilar development: practical considerations from an industry perspective. *Regul. Rapporteur*. 9:15–19.
- Brennan FR, Dougan G. (2005) Non-clinical safety evaluation of novel vaccines and adjuvants, new strategies. *Vaccine*. 23:3210–3222.
- Cavagnaro JA. (2008) *Preclinical Safety Evaluation of Biopharmaceuticals*. John Wiley & Sons, Inc., Hoboken, NJ.
- Center for Biologics Evaluation and Research (CBER). (1997) *Guidance for Industry for the Evaluation of Combination Vaccines for Preventable Diseases: Production, Testing and Clinical Studies*. CBER, Washington, DC.
- Center for Biologics Evaluation and Research (CBER). (2000) *Guidance for Industry: Consideration for Reproductive Toxicity Studies for Preventive Vaccines for Infectious Disease Indications*. CBER, Washington, DC.
- Committee on Proprietary Medicinal Products (CPMP). (1989) Guidelines on the preclinical biological safety testing of medicinal products derived from biotechnology. *TIBTECH*. 7:613–617.

- Dorato MA, McMillian CL, Williams TM. (2014) Toxicological assessment of pharmaceutical and biotechnology products. In: *Principles and Methods of Toxicology* (Hayes AW and Kruger CL, Eds.), 6th Ed. CRC Press, Boca Raton, FL, pp. 333–372.
- Ellis CE, Hartsough MT, Green MD. (2013) Nonclinical safety assessments of biotechnology-derived pharmaceuticals. In: *Nonclinical Safety Assessment* (Brock WJ, Hastings KL and McGowan KM, Eds.). John Wiley & Sons, Ltd, Chichester, UK, pp. 347–380.
- European Medicines Agency (EMA). (2007) *Guideline on Strategies to Identify and Mitigate Risks for First-in-Human Clinical Trials with Investigational Medicinal Products*. EMA, London.
- European Medicines Agency (EMA) (2016) *European public assessment reports*. European Medicines Agency, London. Available online: http://www.ema.europa.eu/ema/index.jsp?curl=pages/medicines/landing/epar_search.jsp&mid=WC0b01ac058001d125 (accessed May 6, 2016).
- Falholt K. (1982) Determination of insulin specific IgE in serum of diabetic patients by solid phase radioimmunoassay. *Diabetologia*. 22:254–257.
- Fent K, Zbinden G. (1997) Toxicity of interferon and interleukin. *TIPS*. 8:100–105.
- Fineberg SE, Galloway JA, Fineberg NS, Rathbun MJ, Hufferd S. (1983) Immunogenicity of recombinant DNA human insulin. *Diabetologia*. 25:465–469.
- Folb PT. (2006) Animal tests as predictors of human response. In: *Clinical Trials of Drugs and Biopharmaceuticals* (Lee C-J, Lee LA, Wu CL, Lee BR and Chen M-L, Eds.). Taylor & Francis, Boca Raton, FL.
- Giezen TJ, Mautel-Teenwisse AK, Straus SMJM, Schellekens H, Leufkens HGM, Egberts ACG. (2008) Safety-related regulatory actions for biologicals approved in the United States and the European Union. *JAMA*. 300:1887–1896.
- Goodyear M. (2006) Editorial: learning from the TGN1412 trial. *BMJ*. 332:677–678.
- Gresser I, Bourali C. (1970) Development of newborn mice during prolonged treatment of interferon. *Eur. J. Cancer*. 6:553–556.
- Griffiths SA, Lumley CE. (1998) *Safety Evaluation of Biotechnology-Derived Pharmaceuticals: Facilitating a Scientific Approach*. Kluwer Academic, Boston.
- Guterman JU, Fine S, Quesada J, Horning SJ, Levine JF, Alexanian R, Bernhardt L, Kramer M, Speigal H, Colburn W, Trown P, Merigan T, Dziewanawska Z. (1982) Recombinant leukocyte A interferon: pharmacokinetics, single-dose tolerance, and biological effects in cancer patients. *Ann. Intern. Med.* 96:549–556.
- Hansel TT, Kropshofer H, Singer T, Mitchell JA, George AJT. (2010) The safety and side effects of monoclonal antibodies. *Nat. Rev. Drug Discov.* 9:325–338.
- International Conference on Harmonization (ICH). (2014) *S6(R1): Preclinical Safety Evaluation of Biotechnology-Designed Pharmaceuticals*. International Conference on Harmonisation, Geneva, Switzerland.
- Kaufmann GHE. (1996) *Concepts in Vaccine Development*. Walter de Gruyter, New York.
- Klaff LJ, Vinik AI, Berelowitz M, Jackson WPU. (1978) Circulating antibodies in diabetics treated with conventional and purified insulins. *S. Afr. Med. J.* 54:149–153.
- Kohler G, Milstein C. (1975) Continuous culture of fused cells secreting antibody of predefined specificity. *Nature*. 256:495–497.
- Levine MM, Kaper JB, Lockman H, Black RE, Clements ML, Falkow S. (1983) Recombinant DNA risk assessment studies in man: efficacy of poorly mobilizable plasmids in biologic containment. *J. Infect. Dis.* 148:699–709.
- Mackett M. (1993) Vaccination and gene manipulation. In: *Molecular Biology and Biotechnology* (Walker JM and Gingold EB, Eds.), 3rd Ed. Royal Society of Chemists, Cambridge, UK.
- Malinowski MJ. (1999) *Biotechnology: Law, Business and Regulation*. Aspen Law & Business, New York.
- Mathieu M. (1997) *Concepts in Vaccine Development*. Walter de Gruyter, New York.
- Maulik S, Patel SD. (1997) *Molecular Biotechnology: Therapeutic Applications and Strategies*. Wiley-Liss, New York.
- Offit PA. (2005) *The Cutter Incident*. Yale University Press, New Haven, CT.
- Oldham RK. (1983) Monoclonal antibodies in cancer therapy. *J. Clin. Oncol.* 1:582–590.
- Petricciani JC. (1983) An overview of safety and regulatory aspects of the new biotechnology. *Regul. Toxicol. Pharmacol.* 3:428–433.
- Plotkin SA, Orenstein WA, Offit PA. (2008) *Vaccines*, 5th Ed. Saunders, Philadelphia, PA.
- Stebbing N, Weck PK. (1984) Preclinical assessment of biological properties of recombinant DNA-derived human interferons. In: *Recombinant DNA Products: Insulin, Interferon and Growth Hormone* (Bollon AP, Ed.). CRC Press, Boca Raton, FL.
- Stephenson J. (2001) Studies illuminate cause of total reaction in gene-therapy trial. *JAMA*. 285:2570–2571.
- Suntharalingam G, Perry MR, Ward S, Brett SJ, Castello-Cortes A, Brunner MD, Panoskaltsis N. (2006) Cytokine storm in a phase 1 trial of the anti-CD28 monoclonal antibody TGN1412. *N. Engl. J. Med.* 355:1018–1028.
- Thomas JA, Myers LA. (1998) *Biotechnology and Safety Assessment*. Taylor & Francis, Philadelphia, PA.
- U.S. Food and Drug Administration (FDA) (2012a) *Scientific Considerations in Demonstrating Biosimilarity to a Reference Product*, April 2015. Center for Drug Evaluation and Research (CDER) and Center for Biologics Evaluation and Research (CBER), Silver Spring, MD.
- U.S. Food and Drug Administration (FDA) (2012b) *Preclinical Assessment of Investigational Cellular and Gene Therapy Products*. Center for Biologics Evaluation and Research (CBER), Rockville, MD.
- Vallbracht A, Treuner J, Manncke KH, Niethammer D. (1982) Autoantibodies against human β interferon following treatment with interferon. *J. Interferon Res.* 2:197.
- Walker JM, Gingold EB. (1993) *Molecular Biology and Biotechnology*. Royal Society of Chemists, Cambridge, UK.
- Weisman R. (2016) Now what? How biotechs deal with failure. *Boston Globe*, July 12, 2016.
- Weissinger J. (1989) Nonclinical pharmacologic and toxicologic considerations for evaluating biologic products. *Regul. Toxicol. Pharmacol.* 10:255–263.
- Zbinden G. (1986) *A toxicologist's view of immunotoxicology*. Proceedings of an International Seminar in the Immunological System as a Target for Toxic Damage. CIJ, Evreux, France.

SAFETY ASSESSMENT OF INHALANT DRUGS AND DERMAL ROUTE DRUGS

While the two most common routes for administering a drug are either orally or parenterally (by injection), there are many other routes that can be used. The two next most common are inhalation (with all its variations) and dermal route, which share with the oral route both the role of major entry path into the body and (because they are the first line of defense against potentially hazardous external agents) specialized barrier functions and immune capabilities. Because such tissue interactions are dependent on the effects of the formulation (and means of administration) that the drug is given to humans, the relevant Food and Drug Administration (FDA) regulatory offices within CDER/Office of New Drugs both require that the GLP systemic (general) toxicity studies be performed with the administration being as it will be to humans.

Each of these routes has special considerations as to the evaluation of their safety—both regulatory and technological—which this chapter will consider in turn.

20.1 INHALED THERAPEUTICS

Drugs and medicinal agents administered by the inhalation route include the gaseous and vaporous anesthetics, coronary vasodilators, the aerosols of bronchodilators, corticosteroids, mucolytics, expectorants, antibiotics, and peptides and proteins where there is significant nasal absorption (Cox et al., 1970; Williams, 1974; Paterson et al., 1979; Hodson et al., 1981; Newman et al., 1981a, b; Lourenco and Cotromanes, 1982). Concerns with the environmental effects of chlorofluorocarbons have also led to renewed interest in dry powder inhalers, which have additionally shown promise for better tolerance and absorption for some new drugs. Excessive inhalation of a drug into the pulmonary system

during therapy or manufacturing may result in adverse local and/or systemic effects. Consequently, safety assessment of inhaled medicinal preparations with respect to pulmonary toxicity and the therapeutic-to-toxicity ratio is essential. The data generated is essential for charting the course of evaluation and development of a potential therapeutic agent.

20.2 THE PULMONARY SYSTEM

An average man inhales approximately 7.5, 28.6, and 42.9 L of air per minute during resting, light work, and heavy work periods, respectively, as well as the corresponding mean tidal volumes of 750, 1673, and 2030 mL (National Academy of Sciences, 1958). Each breath is distributed between 300 and 400 million alveoli, where gas exchange takes place. The total alveoli surface area is approximately 75 m², which is penetrated by approximately 200 km of capillary blood vessels (Hatch and Gross, 1964). The high vascularity and large surface area of the lung ensure rapid gas exchange and entry of an inhaled drug into the bloodstream. A drug is then quickly carried to the heart and brain before reaching the liver, where first-pass metabolism occurs. The pulmonary system is, therefore, a very effective portal through which gases, vapors, and aerosols can enter the body to exert desirable therapeutic effects and undesirable side effects locally and/or systemically.

Anatomically, the pulmonary system is divided into extrathoracic and thoracic regions. The extrathoracic, or head, region includes the nasal and pharyngeal passages. The thoracic region is subdivided into tracheobronchial (TB) and alveolar (AL) regions. The TB region consists of the trachea, primary and secondary bronchi, and primary through-ciliated bronchioles. The AL region consists of

| Regions | | | Generations | Total cross section (cm ²) | Particle size penetration limits (μm) | Distribution | | | |
|---------------|-------------------|-------------------------|-------------|--|---------------------------------------|-------------------------------------|---------------|---------------|---------------|
| | | | | | | Cell types | Mucous glands | Smooth muscle | p-sym. innerv |
| Extrathoracic | Head | Nasopharynx | | | 60 | | | | |
| Intrathoracic | Tracheo bronchial | Trachea | 0 | 2.54 | | Ciliated | | | |
| | | Primary bronchi | 1 | 2.33 | 20 | Goblet | | | |
| | | Secondary bronchi | 2 | 2.13 | 10 | Brush | | | |
| | | | 3 | 2.00 | | Border | | | |
| | | | 4 | 2.48 | | Squamous | | | |
| | | | 10 | 13.4 | | Columnar | | | |
| | | Bronchioles | 11 | 19.6 | 6 | | | | |
| | | | 15 | 113 | | | | | |
| | | Terminal bronchioles | 16 | 180 | 4 | Cuboidal Less ciliated | | | |
| | Alveolar | Respiratory bronchioles | 17 | 300 | 3 | Less goblet nonciliated clara | | | |
| | | | 18 | 534 | | | | | |
| | | | 19 | 944 | | | | | |
| | | Alveoli ducts | 10 | 1.60K | <3 | Type I | | | |
| | | | 21 | 3.22K | | | | | |
| | | | 22 | 5.88K | | | | | |
| | | Alveoli | 23 | 11.8K | <3 | Type II | | | |

FIGURE 20.1 The distribution of cell types in the respiratory tract and lungs.

nonciliated terminal bronchioles, AL ducts, and alveoli (Lippman, 1970; Lippman et al., 1980). The anatomical structure of the pulmonary system maximizes gas exchange but minimizes the penetration of extraneous particulate matter into the lungs. The formalized anatomy (Davis, 1961; Weibel, 1963; Horsfield and Cunnig, 1968; Parent, 2015) of the branchings, the dimensions of the airways, the penetrability by particles of certain sizes, and the distribution of cell types in the respiratory tract and lungs are summarized in Figure 20.1.

20.3 PENETRATION AND ABSORPTION OF INHALED GASES AND VAPORS

Pulmonary dynamics, the dimension and geometry of the respiratory tract, and the structure of the lungs, together with the solubility and chemical reactivity of the inhalants, greatly influence the magnitude of penetration, retention, and absorption of inhaled gases, vapors (Dahl, 1990), and aerosols (Raabe, 1982; Phalen, 2009). The quantity of an inhalant effectively retained in the pulmonary system constitutes the inhaled “dose” that causes pharmacotoxic responses.

Highly reactive and soluble gaseous or vaporous drugs react and dissolve readily in the mucosal membrane of the

nasopharynx and the upper respiratory tract (URT), thereby exerting pharmacological effects or causing local irritation and/or adverse effects on the ciliated, goblet, brush border columnar, and squamous cells of the epithelium (Weibel, 1983). The dissolved drug is also absorbed into the bloodstream and transported to the target organ where it exerts systemic effects. Less reactive and less soluble gaseous or vaporous drugs are likely to penetrate beyond the URT and reach the bronchial and AL regions causing local and systemic effects. The unabsorbed gases or vapors are then exhaled. For example, ammonia gas generated from a 10% ammonia water may be inhaled for reflex respiratory stimulation purposes (Budavari, 1989). Ammonia is extremely soluble in water at a concentration of 715 mL of ammonia per milliliter (mL) of water (Phalen, 2009) and is readily solubilized in the mucous lining causing URT irritation. By contrast, oxygen is only sparingly soluble in water (0.031 cm³ of oxygen per milliliter of water) and capable of penetrating deeply into the alveoli where gas exchange takes place. Oxygen that binds reversibly with the hemoglobin of erythrocytes is unloaded at the target tissues, while the unbound oxygen is exhaled. Inhalation of properly humidified oxygen is life supporting, but inhalation of unhumidified oxygen may cause a reduction in the mucociliary clearance of secretions in the trachea of animals (Pavia, 1984) and humans

(Lichtiger et al., 1975; Gamsu et al., 1976). Gases or vapors of low lipid solubility are also poorly absorbed in the lungs, with much of the inhaled vapor exhaled. Other pharmacological gases and vapors, such as the anesthetics (nitrous oxide, halothane, enflurane, isoflurane, etc.) and the coronary vasodilators (amyl nitrite), likewise affect the epithelium of the respiratory tract and the lungs. The absorbed drugs exert local effects on various types of epithelial cells of the respiratory tract and on type I and II cells and the alveolar macrophages (AMs) in the alveoli. Repeated inhalation of some halogenated hydrocarbon anesthetics will result in accumulation of the vapors and systemic toxicity (Chenoweth et al., 1972). By contrast, vapors such as the fluorocarbons (FC 11 and FC 12), which are used extensively as propellants for bronchodilator and corticosteroid aerosols, are absorbed rapidly but are not accumulated in the body even upon repeated inhalation (Aviado and Micozzi, 1981).

In general, dissolved gases or vapors at a nontoxic concentration are absorbed and metabolized locally by the lungs and systemically by the liver. The unchanged parent drug and its metabolites may be excreted to some extent via exhalation but mainly via the renal system. A dissolved gas or vapor at a toxic concentration, however, is likely to exert local effects such as altering the surface tension of the alveoli linings or disrupting the normal functions of the epithelial cells, the pneumocytes, and the AMs. The disrupted AMs in turn release their intracellular enzymes, causing destruction of the AL septa and contributing to histopathologic changes of the respiratory tract and the lungs. Again, the magnitude of the adverse effects is dependent on pulmonary dynamics and the solubilities of the inhalants in the mucous membrane of the URT and in the plasma or lipids of the erythrocytes.

20.4 DEPOSITION OF INHALED AEROSOLS

For inhaled aerosols, particle size is the major factor affecting the penetration, deposition, and hence the “dose” and site of pharmacological action (Dautrebande, 1962a, b; Agnew, 1984) of administered aerosols (liquid or solid). Particle size is expressed in terms of *aerodynamic diameter* (AD), defined as the diameter of a spherical particle of unit density (1 g cm^{-3}) that has the same terminal settling velocity as the particle in question, regardless of its shape and density (Marple and Rubow, 1980). The unit for AD is micrometers (μm). A sample of aerosol particles having ADs within a narrow size range is considered to be a monodisperse aerosol, whereas a sample of aerosols with a wide range of ADs is a heterodisperse, or polydisperse, aerosol. The pattern of particle-size distribution is usually bell shaped, with smaller and larger particles on both sides of the mean AD. An aerosol sample with a high proportion of particles of similar size has a narrow particle-size distribution or small

geometric standard deviation (GSD). An aerosol sample with a GSD of less than 2 is considered to be a monodisperse aerosol. Thus, both the AD and GSD of 2 or less are considered to be optimal for pulmonary penetration and distribution in the respiratory tract and the lungs. For example, in nose breathing, aerosol particles with ADs $> 15 \mu\text{m}$ are likely to be trapped in the nasopharynx (extrathoracic, or head, region) by filtration and impaction. Particles deposited in the nasopharynx are considered to be “noninhalable” (Lippman, 1977; Miller et al., 1979).

In mouth breathing, only 10–15% of $15 \mu\text{m}$ particles penetrate through the larynx to the intrathoracic TB region. Particles reaching the TB region are considered to be “inhalable” (Lippman, 1977; Miller et al., 1979).

In natural nose and mouth breathing, only a negligible proportion of aerosol particles of $\text{AD} > 10 \mu\text{m}$ reach the lungs (Swift and Proctor, 1982). Aerosol particles of 3–4 μm in AD are considered to be optimal sizes for TB deposition. The mechanisms of deposition are by impaction along the trachea and at bronchial branchings where the direction of airflow changes and by gravity settlement in the fine airways in amounts proportional to the particle settling velocity and the time available for settlement (Hatch and Gross, 1964; Heyder et al., 1980). Aerosol particles of 1–2 μm in AD, however, decrease in TB deposition because the particles are too small for effective impaction and sedimentation (Lippman, 1977; Chan and Lippmann, 1980; Stahlhofen et al., 1980). Consequently, the majority of the very fine particles are exhaled. However, the deposition of the ultrafine particles of approximately 0.5 μm in AD on the walls of the finest bronchioles and the alveoli increases again due to molecular diffusion processes. Even so, some 90% of the inhaled 0.5 μm particles will still be exhaled during quiet tidal breathing and much more under forced exhalation (Davis et al., 1972; Taulbee et al., 1978). Those fine particles reaching the finest bronchioles and alveoli are considered to be “respirable” (Lippman, 1970).

In general, particles of $\text{AD} > 10 \mu\text{m}$ deposit mainly in the URT, whereas particles of 1–5 μm AD, with a GSD of less than 2, are likely to reach the lower respiratory tract, which includes the TB region and the alveoli, with small oropharyngeal loss.

The proportion of an aerosol sample suitable for inhalation can also be determined on the basis of mass median aerodynamic diameter (MMAD), which is defined as the percentage (50%) by weight of an aerosol sample having ADs equal to or less than the stated median AD. For example, a sample with an MMAD of 5 μm means that 50% by weight of that sample has ADs of 5 μm and smaller. The MMAD is, therefore, a good index for determining the proportion of an aerosol sample that is “noninhalable,” “inhalable,” or “respirable.” An aerosol sample with an MMAD of 5 μm and a GSD of less than 2 is considered to be optimal for pulmonary deposition and retention (Task Group on Lung Dynamics, 1966).

In addition to AD and GSD, the pulmonary dynamics of a subject also greatly influence the distribution of aerosol particles in various regions of the respiratory tract (Agnew, 1984). For example, the velocity of airflow in the respiratory tract significantly influences the pattern of TB deposition. An increase in airflow velocity in the airways increases the effectiveness of particle impaction at the bifurcations of the large airways (Dennis, 1961; Hatch and Gross, 1964; Parent, 2015). As a result, spots impacted with a high concentration of particles (hot spots) are frequently present at the carina and the bifurcations of the airways (Lee and Wang, 1977; Bell, 1978; Stahlhofen et al., 1981). Furthermore, the depth of each breath (tidal volume) also influences the distribution of aerosols. A small tidal volume permits greater impaction in the proximal conducting airways and less sedimentation in the distal airways.

In general, slow, deep inhalation followed by a period of breath holding increases the deposition of aerosols in the peripheral parts of the lungs, whereas rapid inhalation increases the deposition in the oropharynx and in the large central airways. Thus, the frequency of respiration (the flow velocity) and the depth of breath (tidal volume) influence the pattern of pulmonary penetration and deposition of inhaled aerosols. Therefore, an aerosol of ideal size will penetrate deeply into the respiratory tract and the lungs only when the aerosols are inhaled in the correct manner.

20.5 ABSORPTION AND CLEARANCE OF INHALED AEROSOLS

Soluble aerosols deposited on the epithelial linings of the respiratory tract are absorbed and metabolized in the same way as soluble gases and vapors.

Insoluble medicinal aerosols are few in number. Sodium cromoglycate (SCG) is probably the only insoluble powder to be administered as a prophylactic antiasthmatic (Wanner, 1979). Insoluble particles deposited on the ciliated linings of the URT are removed by a mucociliary clearance mechanism (Lauweryns and Baert, 1977). Particles deposited on a terminal airway devoid of ciliated cells may be endocytosed into the epithelial cells (Jones, 1984; Newhouse et al., 1976). At a toxic concentration, the cells die and the debris is then phagocytosed and transported into the interstitial space for removal via the lymph or vascular drainages or reenters the ciliated zone of the airway. Particles deposited in the AL walls are phagocytosed by the AMs and transported from the low surface tension surfactant in the AL lining to the high surface tension bronchial fluid of the ciliate airways for elimination by the mucociliary clearance mechanism (Lauweryns and Baert, 1977). The particle sizes optimal for phagocytosis are 2–3 μm , while particles smaller than 0.26 μm are less effective in activating the macrophages (Holma, 1967). In any case, AMs can phagocytose only a

small fraction of a large number of deposited particles. The nonphagocytosed particles are translocated to the lymphatic system for elimination (Ferin, 1977).

Like the inhaled gases or vapors, soluble and insoluble aerosol particles can directly exert desirable and undesirable local effects at the site of deposition and/or systemic effects after solubilization, absorption, and metabolism (Sackner et al., 1975; Sackner, 1978).

20.6 PHARMACOKINETICS AND PHARMACODYNAMICS OF INHALED AEROSOLS

The inhalation route for administering drugs into the pulmonary system for treatment of respiratory diseases eliminates many bioavailability problems such as plasma binding and “first-pass” metabolism, which are encountered in parenteral or oral administration. Consequently, a small inhalation dose is adequate for achieving the desirable therapeutic response without inducing many undesirable side effects. Furthermore, the direct contact of the drug with the target site ensures rapid action. Nevertheless, the effects from inhaled drug aerosols also depend on the pharmacological properties of the aerosols and the location of their deposition in the respiratory system. For example, the classic experiments on bronchodilation drugs (Dautrebande, 1962a, b) showed that fine aerosol particles of isoproterenol penetrate deeply into the lower respiratory airways (LRA). In this way, a high concentration of the drug aerosol can reach the beta-adrenergic receptors of the bronchial smooth muscles. Stimulation of the receptors causes relaxation of the smooth muscle fibers and results in bronchodilation (Weiner, 1984; McFadden, 1986). Such rapid bronchial responses can be produced in healthy and asthmatic subjects without inducing any cardiac effects. By contrast, the same dose of isoproterenol of large particle sizes deposits mainly along the URT, with a minimal amount reaching the smooth muscles of the LRA. The drug is quickly absorbed into the tracheal and bronchial veins and delivered immediately to the left ventricle of the heart. A high plasma concentration of the drug in the heart causes prominent cardiovascular effects such as tachycardia and hypertension. Other aerosols of beta-adrenergic drugs, such as epinephrine, isoprenaline, terbutaline, and salbutamol, induce bronchodilation effects in animals and humans (Pavia, 1984) via inhalation and stimulate ciliary beat frequency and mucus production at the site of deposition in the trachea (Wanner, 1981). Thus the TB mucociliary clearance mechanism is also stimulated. By contrast, anticholinergic bronchodilators, such as atropine and ipratropium bromide, cause mucus retention in the lungs (Pavia et al., 1983a, b). Therefore, in pharmacological or safety assessments of inhalant beta-adrenergic bronchodilation drugs, aerosols should be of small particle sizes suitable for deposition in the peripheral airways to minimize side

effects. However, anticholinergic agents should be of larger particle sizes suitable for deposition in the large airways (Ingram et al., 1977; Hensley et al., 1978).

Other therapeutic aerosols, such as beclomethasone dipropionate, betamethasone valerate, and budesonide corticosteroid (Williams, 1974); the carbenicillin and gentamicin antibiotics (Hodson et al., 1981); the 2-mercaptoethane sulfonate (Pavia et al., 1983b) and *n*-acetylcysteine (Hollinger, 1985) mucolytics; and even vaccines for the prevention of influenza and tuberculosis (Lourenco and Cotromanes, 1982) are active by inhalation and/or oral administration. When these drugs are administered as aerosols, certain particle sizes may be targeted to a specific region or to multiple regions of the pulmonary system depending on the therapeutic target site(s). In any case, when aerosols are delivered as fine particles, the rate of absorption is increased because of an increase in the distribution area per unit mass of the drug. Thus, an effective aerosol dose of corticosteroid for treatment of asthma and bronchitis is merely a fraction of an oral dose (Williams, 1974). An aerosol of SCG dry powder, a prophylactic for preventing the onset of bronchoconstriction in asthmatic attacks (Cox et al., 1970), is effective mainly by local inhibition of the release of chemical mediators from mast cells in bronchial smooth muscle. Therefore, SCG particle sizes should be approximately 2 μ m in AD for the most effective penetration into the bronchial regions (Godfrey et al., 1974; Curry et al., 1975). Likewise, therapeutic aerosols of local anesthetics and surfactants may require appropriate particle sizes to be targeted to a specific region of the pulmonary system.

Other than undesirable pharmacological effects, toxic concentrations of soluble or insoluble aerosol particles may lead to adverse physiological and/or histopathologic responses. For example, irritating aerosols cause dose-related reflex depression of the respiratory rate (Alarie, 1966, 1981a), while phagocytosed particles cause chemotaxis of AMs and neutrophils to the site of deposition (Brain, 1971). The maximum response usually occurs at 24 h and returns to normal in approximately 3 days postexposure (Kavet et al., 1978). Furthermore, a toxic quantity of phagocytosed particles may interact with the lysosomal membrane within a macrophage, releasing cytotoxic lysosomal enzymes, proteases, and free radicals that in turn damage the adjacent lung tissue (Hocking and Golde, 1979).

In general, a specific category of drug delivered to a specific site of the pulmonary system will exert a specific pharmacological or toxicological action locally or systemically. Therefore, in safety assessments of inhalants, a drug should be delivered to the target sites of the pulmonary system according to the toxicological information required.

Finally, there are many drugs in the categories of amphetamines, anorectics, antihistamines, antipsychotics, tricyclic antidepressants, analgesics and narcotics, and beta-adrenergic blocking agents that are known to accumulate in the lung

(Wilson, 1982; Hollinger, 1985) even though these drugs are not administered via the inhalation route. Therefore, in safety assessments of these drugs, their pulmonary toxicity should also be evaluated.

20.7 METHODS FOR SAFETY ASSESSMENT OF INHALED THERAPEUTICS

Methods for evaluation of inhalation toxicity should be selected according to the pharmacological and/or the toxicological questions asked, and the design of experiments should specify the delivery route of a drug to the target sites in the pulmonary system (Gad, 2015). For example, if an immunologic response of the lungs to a drug is in question, then the lymphoid tissues of the lungs should be the major target of evaluation. The following are some of the physiological, biochemical, and pharmacological tests that are applicable for safety assessment of inhaled medicinal gases, vapors, or aerosols.

URT irritation can occur from inhalation of a medicinal gas, vapor, or aerosol. Because of the necessary fragility of the thin active gas exchange epithelial tissue, severe or repeated irritation can readily cause the development of non-functional scar tissue. For assessing the potential of an inhalant to cause URT irritation, the mouse body plethysmographic technique (Alarie, 1966, 1981a, b) has proven to be extremely useful. This technique operates on the principle that respiratory irritants stimulate the sensory nerve endings located at the surface of the respiratory tract from the nose to the AL region. The nerve endings in turn stimulate a variety of reflex responses (Alarie, 1973; Widdicombe, 1974) that result in characteristic changes in inspiratory and expiratory patterns and, most prominently, depression of respiratory rate. Both the potency of irritation and the concentration of the irritant are positively related to the magnitude of respiratory rate depression. The concentration response can be quantitatively expressed in terms of "RD₅₀," defined as the concentration (in logarithmic scale) of the drug in the air that causes a 50% decrease in respiratory rate. The criteria for positive URT irritation in intact mice exposed to the drug atmosphere are depression in breathing frequency and a qualitative alteration of the expiratory patterns. Numerous experimental results have shown that the responses of mice correlated almost perfectly with those of humans (Alarie et al., 1980; Alarie and Luo, 1986). Thus, this technique is useful for predicting the irritancy of airborne medicinal compounds in humans. From the drug-formulating point of view, an inhalant drug with URT-irritating properties indicates the need for an alternate route of administration. From the industrial hygiene point of view, the recognition of the irritant properties is very important. If a chemical gas, vapor, or aerosol irritates, it has a "warning property." With an adequate warning property, a worker will avoid inhaling

damaging amounts of the airborne toxicant; without a warning property, a worker may unknowingly inhale an injurious amount of the toxicant.

Inhalation of a cardiovascular drug, such as an aerosol of propranolol (a beta-adrenergic receptor agonist), may affect the respiratory cycle of a subject. For evaluating the cardiopulmonary effects of an inhalant, the plethysmographic technique using a mouse or a guinea pig model is useful. The criteria for a positive response in intact mice or guinea pigs are changes in the duration of inspiration and expiration and the interval between breaths (Schaper et al., 1989).

Pulmonary sensitization may occur from inhalation of drug vapors such as enflurane (Schwettmann and Casterline, 1976) and antibiotics such as spiramycin (Davies and Pepys, 1975) and tetracycline (Menon and Das, 1977). To detect pulmonary sensitization from inhalation of drug and chemical aerosols, the body plethysmographic technique using a guinea pig model has been shown to be useful (Patterson and Kelly, 1974; Paterson, 1977; Karol, 1988; Karol and Thorne, 1988; Karol et al., 1989; Thorne and Karol, 1989). The criteria for positive pulmonary sensitization in intact guinea pigs are changes in breathing frequency and their extent and the time of onset of an airway constrictive response after induction and after a challenge dose of the test drug (Karol et al., 1989).

The mucociliary transport system of the airways can be impaired by respiratory irritants, local analgesics and anesthetics, and parasympathetic stimulants (Pavia, 1984). Any one of the earlier agents will retard the beating frequency of the cilia and the secretion of the serous fluid of the mucous membranes. As a result, the propulsion of the inhaled particles, bacteria, or endogenous debris toward the oral pharynx for expectoration or swallowing will be retarded. Conversely, inhalation of adrenergic agonists increases the activity of the mucociliary transport system and facilitates the elimination of noxious material from the pulmonary system. Laboratory evaluation of the adverse drug effects on mucociliary transport in animal models can be achieved by measuring the velocity off the linear flow of mucus in the trachea of surgically prepared animals (Rylander, 1966; Oyarzun and Clements, 1977, 1978). Clinically, the transportation of markers placed on the tracheal epithelium of normal human subjects can also be observed using a fiber-optic bronchoscopic technique (Pavia et al., 1980; Mussatto et al., 1988). The criteria of a positive response are changes in the transport time over a given distance of markers placed on the mucus or changes in the rate of mucus secretion (Davis et al., 1976; Johnson et al., 1983, 1987; Webber and Widdicombe, 1987). More comprehensive discussion on mucociliary clearance can be found in several reviews (Last, 1982; Pavia, 1984).

Cytological studies on the bronchoalveolar lavage fluid (BALF) permit the evaluation of the effects of an inhaled drug on the epithelial lining of the respiratory tract. This fluid can be obtained from intact animals or from excised

lungs (Henderson, 1984, 1988, 1989). Quantitative analyses of fluid constituents such as neutrophils, antibody-forming lymphocytes, and antigen-specific IgG provide information on the cellular and biochemical responses of the lungs to the inhaled agent (Henderson, 1984; Henderson et al., 1985, 1987). For example, BALF parameters were found to be unperturbed by the inhalation of halothane (Henderson and Loery, 1983). The criteria of a positive response are increases in protein content, increase in the number of neutrophils and macrophages for inflammation, increase in the number of lymphocytes and alteration of lymphocyte profiles for immune response, increase in cytoplasmic enzymes (lactate dehydrogenase) for cell lysis (Henderson, 1989), and the presence of antigen-specific antibodies for specific immune responses (Bice, 1985).

Morphological examination of the cellular structure of the pulmonary system is the foundation of most inhalation toxicity studies. Inhalation of airborne drug vapors or aerosols at harmful concentrations results mainly in local histopathologic changes in the epithelial cells of the airways, of which there are two types: nonciliated and ciliated cells. The nonciliated cells are the Clara cells, which contain secretory granules and smooth endoplasmic reticulum (SER); cells that have secretory granules but lack SER; and the brush cells, which have stubby microvilli and numerous cytoplasmic fibers on their free surfaces. If the concentration gradient of the drug in the lung is high enough to reach the alveoli, type I alveoli cells will also be affected (Evans, 1982). Drugs that affect the lungs via the bloodstream, such as bleomycin (Aso et al., 1976), cause changes to the endothelial cells of the vascular system that result in diffuse damage to the alveoli. The criteria of cellular damage are loss of cilia, swelling, and necrosis and sloughing of cell debris into the airway lumina. Tissues recovering from injuries are characterized by increases in the number of dividing progenitor cells followed by increases in intermediate cells that eventually differentiate into normal surface epithelium.

Pulmonary drug disposition studies are essential in research and development of new inhalant drugs. Inhaled drugs are usually absorbed and metabolized to some extent in the lungs because the lungs, like the liver, contain active enzyme systems. A drug may be metabolized to an inactive compound for excretion or to a highly reactive toxic metabolite that causes pulmonary damage. In most pulmonary disposition studies, a gas or vapor is delivered via whole-body exposure (Paustenbach et al., 1983) or head-only exposure (Hafner et al., 1975). For aerosols, over 90% of a dose administered by mouth breathing is deposited in the oropharynx and swallowed. Consequently, the disposition pattern reflects that of ingestion in combination with a small contribution from pulmonary metabolism. For determining the disposition of inhaled drugs by the pulmonary system alone, a dosimetric endotracheal nebulization technique (Leong et al., 1987, 1988) is useful. In this technique, micro-

liter quantities of a radiolabeled drug solution can be nebulized within the trachea using a miniature air-liquid nebulizing nozzle. Alternatively, a small volume of liquid can be dispersed endotracheally using a microsyringe. In either technique, an accurate dose of a labeled drug solution is delivered entirely into the respiratory tract and lungs. Subsequent radioassay of the excreta thus reflects only the pulmonary disposition of the drug without complication from aerosols deposited in the oropharyngeal regions if the drug had been delivered by mouth inhalation. For example, in a study of the antiasthmatic drug Iodoxamide tromethamine, the urinary metabolites produced by beagle dogs after receiving a dose of the radiolabeled drug via endotracheal nebulization showed a high percentage of the intact drug. However, metabolites produced after oral administration were mainly nonactive conjugates. The differences were due to the drug's escape from first-pass metabolism in the liver when it was administered through the pulmonary system. The results thus indicated that the drug had to be administered by inhalation to be effective. This crucial information was extremely important in the selection of the most effective route of administration and formulation of this antiasthmatic drug (Leong et al., 1988).

Cardiotoxicity of inhalant drugs should also be evaluated. For example, adverse cardiac effects may be induced by inhaling vapors of fluorocarbons, which are used extensively as propellants in drug aerosols. Inhalation of vapors of anesthetics also has been shown to cause depression of the heart rate and alteration of the rhythm and blood pressure (Merin, 1981; Leong and Rop, 1989). More important, inhalation of antiasthmatic aerosols of beta-receptor agonists delivered in a fluorocarbon propellant has been shown to cause marked tachycardia, electrocardiogram (ECG) changes, and sensitization of the heart to arrhythmia (Aviado, 1981; Balazs, 1981). Chronic inhalation of drug aerosols can also result in cardiomyopathy (Balazs, 1981). For detection of cardiotoxicity, standard methods of monitoring arterial pressures, heart rate, and ECGs of animals during inhalation of a drug, or at frequent intervals during a prolonged treatment period, should be useful in safety assessments of inhalant drugs.

Since the inhalation route is just a method for administering drugs, other nonpulmonary effects, such as behavioral effects (Ts'o et al., 1975) and renal and liver toxicity, should also be evaluated. In addition, attention should also be given to drugs that are not administered via the inhalation route, but that accumulate in the lungs where they cause pulmonary damage (Wilson, 1982; Hollinger, 1985).

20.8 PARAMETERS OF TOXICITY EVALUATION

Paracelsus stated over 400 years ago that "All substances are poison. The right dose differentiates a poison and a remedy." Thus, in safety assessments of inhaled drugs, the "dose," or

magnitude of inhalation exposure, in relation to the physiological, biochemical, cytological, or morphological response(s), must be determined. Toxicity information is essential to establishing guidelines to prevent the health hazards of acute or chronic overdosage during therapy or of unintentional exposure to the bulk drugs and their formulated products during manufacturing and industrial handling.

20.8.1 The Inhaled "Dose"

Most drugs are designed for oral or parenteral administration in which the dose is calculated in terms of drug weight in milligrams (mg) divided by the body weight in kilograms (kg):

$$\text{Dose} = \frac{\text{drug weight (mg)}}{\text{body weight (kg)}} = \frac{\text{mg}}{\text{kg}}$$

For inhalant drugs, the inhaled "dose" has been expressed in many mathematical models (Dahl, 1990). However, the practical approach is based on exposure concentration and duration rather than on theoretic concepts. Thus, an inhaled "dose" is expressed in terms of the exposure concentration (C) in milligrams per liter (mg L^{-1}) or milligrams per cubic meter (mg m^{-3}) or parts per million (ppm) parts of air, the duration of exposure (t) in minutes, the ventilatory parameters including the respiratory rate (R) in number of breaths per minute and the tidal volume (T_v) in liters per breath, and a retention factor α (alpha), which is related to the reactivity and the solubility of the drug. The product of these parameters divided by the body weight in kilograms gives the dose:

$$\text{Dose} = \frac{CtRT_v\alpha}{\text{body weight}} = \frac{\text{mg}}{\text{kg}}$$

In critical evaluation of the effect of a gas, vapor, or aerosol inhaled into the respiratory tract of an animal, the dosimetric method has been recommended (Oberst, 1961). However, due to the complexity of measuring the various parameters simultaneously, only a few studies on gaseous drugs or chemicals have employed the dosimetric method (Weston and Karel, 1946; Carpenter et al., 1949; Leong and MacFarland, 1965; Landy et al., 1983; Stott and McKenna, 1984; Dallas et al., 1986, 1989). For studies on liquid or powdery aerosols, modified techniques such as intratracheal instillation (Brain et al., 1976) or endotracheal nebulization (Leong et al., 1988) were used to deliver an exact dose of the test material into the lower respiratory tract (LRT) while bypassing the URT and ignoring the ventilatory parameters.

In routine inhalation studies, it is generally accepted that the respiratory parameters are relatively constant when the animals are similar in age, sex, and body weight. This leaves only C and t to be the major variables for dose consideration:

$$\text{"Dose"} = Ct = \text{mg min L}^{-1}$$

The product Ct is not a true dose because its unit is milligrams per minute per liter rather than milligrams per kilogram. Nevertheless, Ct can be manipulated as though it were a dose, an approximated dose (MacFarland, 1976).

The respiratory parameters of an animal will dictate the volume of air inhaled and hence the quantity of test material entering the respiratory system. Commonly used parameters for a number of experimental species and man are given in Table 20.1 to illustrate this point and include the AL surface area because this represents the target tissue for most inhaled materials. It can be seen that by taking the ratios of these parameters and comparing the two extremes, that is, the mouse and man, (i) a mouse inhales approximately 30 times its lung volume in 1 min, whereas a human at rest inhales approximately the same volume as that of his or her lung. This can increase with heavy work up to the same ratio as the mouse, but is not sustained for long periods. This means that the dose per unit lung volume is up to 30 times higher in the mouse than man at the same inhaled atmospheric concentration. (ii) The minute volume of the mouse is in contact with five times less AL surface area than man; hence the dose per unit area is up to five times greater in the mouse. (iii) The lung volume in comparison with the AL surface area in experimental animals is less than in humans, meaning that the extent of contact of inhaled gases with the AL surface is greater in experimental animals.

While it is possible, and common, to refer to standard respiratory parameters for different species in order to calculate inhaled dose and deposited dose with time, it is usually the case that inhaled materials influence the breathing patterns of test animals. The most common examples of this are irritant vapors, which can reduce the respiratory rate by up to 80%. This phenomenon results from a reflexive pause during the breathing cycle due to stimulation by the inhaled material of the trigeminal nerve endings situated in the nasal passages. The duration of the pause and hence the reduction in the respiratory rate are concentration related, permitting concentration–response relationships to be plotted. This has been investigated extensively by Alarie (1981a) and forms the basis of a test screen for comparing quantitatively the irritancy of different materials and has found application in assessing appropriate exposure limits for

human exposure when respiratory irritancy is the predominant cause for concern.

While irritancy resulting from the earlier reflex reaction is one cause of altered respiratory parameters during exposure, there are many others. These include other types of reflex response, such as bronchoconstriction, the narcotic effects of many solvents, the development of toxic signs as exposure progresses, or simply a voluntary reduction in respiratory rate by the test animal due to the unpleasant nature of the inhaled atmosphere. The extent to which these affect breathing patterns and hence inhaled dose can only be assessed by actual measurement.

By simultaneous monitoring of tidal volume and respiratory rate, or minute volume, and the concentration of an inhaled vapor in the bloodstream and the vapor in the exposure atmosphere, pharmacokinetic studies on the Ct relationship have shown that the effective dose was nearly proportional to the exposure concentration for vapors such as 1,1,1-trichloroethane (Dallas et al., 1989), which has a saturable metabolism, and found that the steady-state plasma concentrations were disproportionately greater at higher exposure concentrations.

Acknowledging the possible existence of deviations, this simplified approach of using C and t for dose determination provides the basis for dose–response assessments in practically all inhalation toxicological studies.

20.8.2 The Dose–Response Relationship

The first principle of dose–response determination in inhalation toxicology is based on Haber's rule, which states that responses to an inhaled toxicant will be the same under conditions where C varies in complementary manner to t (Haber, 1924), for example, if Ct elicits a specific magnitude of the same response, that is, $Ct=K$, where K is a constant for the stated magnitude of response.

This rule holds reasonably well when C or t varies within a narrow range for acute exposure to a gaseous compound (Rinehart and Hatch, 1964) and for chronic exposure to an inert particle (Henderson et al., 1991). Excursion of C or t beyond these limits will cause the assumption $Ct=K$ to be incorrect (Adams et al., 1950, 1952; Sidorenko and Pinigin, 1976; Andersen et al., 1979; Uemitsu et al., 1985).

TABLE 20.1 Respiratory Parameters for Common Experimental Species and Man

| Species | Body Weight (kg) | Lung Volume (mL) | Minute Volume (mL min ⁻¹) | Alveolar Surface Area (m ²) | Lung Volume % Surface Area | Minute Volume % Lung Volume | Minute Volume % Surface Area |
|---------|------------------|------------------|---------------------------------------|---|----------------------------|-----------------------------|------------------------------|
| Mouse | 0.023 | 0.74 | 24 | 0.068 | 10.9 | 32.4 | 353 |
| Rat | 0.14 | 6.3 | 84 | 0.39 | 16.2 | 13.3 | 215 |
| Monkey | 3.7 | 184 | 694 | 13 | 14.2 | 3.77 | 53 |
| Dog | 22.8 | 1501 | 2923 | 90 | 16.7 | 1.95 | 33 |
| Human | 75 | 7000 | 6000 | 82 | 85.4 | 0.86 | 73 |

Source: Data from Altman and Ditmar (1974).

For example, an animal may be exposed to 1000 ppm of diethyl ether for 420 min or 1400 ppm for 300 min without incurring any anesthesia. However, exposure to 420 000 ppm for 1 min will surely cause anesthesia or even death of the animal. Furthermore, toxicokinetic study of liver enzymes affected by inhalation of carbon tetrachloride (Uemitsu et al., 1985), which has a saturable metabolism in rats, showed that $Ct=K$ does not correctly reflect the “toxicity value” of this compound. Therefore, the limitations of Haber’s rule must be recognized when it is used in interpolation or extrapolation of inhalation toxicity data.

20.8.3 Exposure Concentration versus Response

In certain medical situations (e.g., a patient’s variable exposure duration to a surgical concentration of an inhalant anesthetic or the repeated exposures of surgeons and nurses to subanesthetic concentrations of an anesthetic in the operating theater), it is necessary to know the duration of safe exposure to a drug. Duration safety can be assessed by determining a drug’s median effective time (ET_{50}) or median lethal time (LT_{50}). These statistically derived quantities represent the duration of exposure required to affect or kill 50% of a group of animals exposed to a specified concentration of an airborne drug or chemical in the atmosphere.

The graph in Figure 20.2 is the probit plot of cumulative percentage response to logarithm of exposure duration. It shows 1000 mg m^{-3} for 10 h or 10 mg m^{-3} for 1000 h, each with a Ct (an approximated dose) of $\sim 10000 \text{ h mg m}^{-3}$. Similar to concentration–response graphs, the slopes indicate the differences in the mechanism of action and the margins of safe exposure of the three drugs. The ratio of ET_{50} over LT_{50} of two drugs indicates their relative toxicity, and the ratio of ET_{50} over LT_{50} of the same drug is the therapeutic ratio.

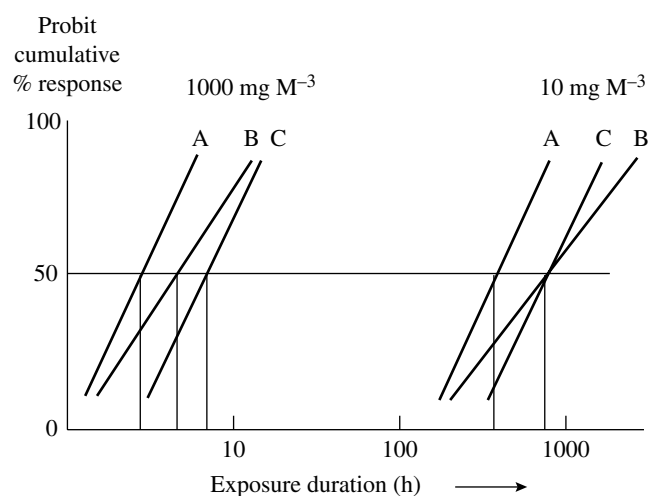


FIGURE 20.2 Dose–response plot in terms of probit of cumulative percentage response to logarithm of exposure duration. A, B, and C indicate three exemplary curves.

20.8.4 Product of Concentration and Duration (Ct) versus Responses

To evaluate inhalation toxicity in situations where workers are exposed to various concentrations and durations of a drug vapor, aerosol, or powder in the work environment during manufacturing or packaging, a more comprehensive determination of $E(Ct)_{50}$ or $L(Ct)_{50}$ values is used. The $E(Ct)_{50}$ or $L(Ct)_{50}$ values are statistically derived values that represent the magnitude of exposure, expressed as a function of the product of C and t , that is expected to affect or kill less than 50% and greater than 50% of the animals. The other curve represents exposures that kill 50% or greater than 50% of each group of animals (Irish and Adams, 1940).

The graph in Figure 20.3 illustrates inhalation exposures to a drug using various combinations of C and t that kill 50% of the animals. For example, a 50% mortality occurs when a group of animals is exposed to drug A at a concentration of 1000 mg m^{-3} for a duration of approximately 2 h or at a concentration of 100 mg m^{-3} for a duration of approximately 20 h. Furthermore, the graph also illustrates that the inhalation toxicity of drug A is more than one order of magnitude higher than that of drug B. For example, an exposure to drug A at the concentration of 100 mg m^{-3} for 100 h kills 100% of the animals, whereas an exposure to drug B at the concentration of 1000 mg m^{-3} for 100 h does not kill any animals.

20.8.5 Units for Exposure Concentration

For gases and vapors, exposure concentrations are traditionally expressed in parts per million (ppm). The calculation for the ppm of a gas or vapor in an air sample is based on

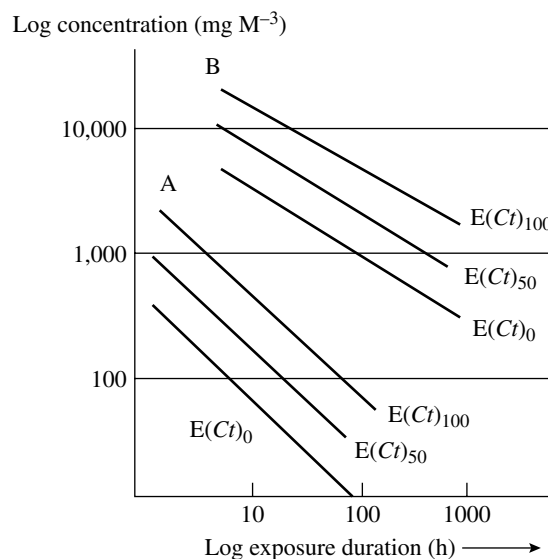


FIGURE 20.3 Dose–response plot in terms of logarithms of exposure concentration and durations. A and B indicate nonspecific example compounds A and B.

Avogadro's law, which states that "Equal volumes contain equal numbers of molecules under the same temperature and pressure." In other words, under standard temperature and pressure (STP), one gram-molecular weight (mole) of any gas under a pressure of 1 atm (equivalent to the height of 760 mm mercury) and a temperature of 273 K has the same number of molecules and occupies the same volume of 22.4 L. However, under ambient conditions, the volume of 22.4 L has to be corrected to a larger volume based on Charles' law, which states that at constant pressure the volume of gas varies directly with the absolute temperature. Thus, at a room temperature of 25°C, 1 mol of a gas occupies a volume of 24.5 L:

$$22.4 \text{ L} \times \frac{298 \text{ K}}{273 \text{ K}} = 24.5 \text{ L}$$

Further correction of volume for an atmospheric pressure deviation from one atmosphere may be done by applying Boyle's law, which states that the volume of a gas without change of temperature varies inversely with the pressure applied to it:

$$24.5 \text{ L} \times \frac{758 \text{ mm Hg}}{760 \text{ mm Hg}} = 24.4 \text{ L}$$

In practice, atmospheric pressure in most animal experimental environments usually varies only a few mm Hg, so little or no correction is required.

Using the aforementioned principles, the volume of a vapor generated from a given weight of a liquid can be calculated. For example, 1 mol of water weighs 18 g, while 1 mol of ethanol weighs 46 g. When 1 mol of each liquid is totally vaporized, each will occupy the same volume of 24.5 L at room temperature (25°C) and pressure (760 mm Hg). In an inhalation experiment, if the volume of test liquid and the rate of airflow being mixed in the animal exposure chamber are known, the vapor concentration in the chamber atmosphere can be calculated in parts per million or milligrams per liter. A conversion table published by the US Bureau of Mines enables quick conversion between parts per million and milligrams per liter for compounds with molecular weights up to 300 g (Fieldner et al., 1921; Patty, 1958).

For aerosols of nonvolatile liquid and powdery compounds, the concentration of the mist or dust atmosphere must be expressed in terms of milligrams per liter or milligrams per cubic meter (mg m^{-3}) of air. With advances in biotechnology, many pharmacological testing techniques are based on specific receptor bindings, in which the ratio of the number of molecules to those of the receptors is considered, in which case the exposure concentration may be more appropriately expressed in micromoles per unit volume of air ($\mu\text{mol m}^{-3}$).

20.9 INHALATION EXPOSURE TECHNIQUES

Many inhalation exposure techniques—such as the whole-body, nose-only, mouth-only, or head-only technique (Drew and Laskin, 1973; MacFarland, 1976; Leong et al., 1981; Smith et al., 1981; Phalen, 2009), the intranasal exposure technique (Elliott and DeYoung, 1970), the endotracheal nebulization technique (Leong et al., 1985, 1988; Schreck et al., 1986), and the body plethysmographic technique (Alarie, 1966; Thorne and Karol, 1989)—have been developed for inhalation toxicity studies. Table 20.2 provides a summary of the advantages and disadvantages of each of the major inhalation exposure methodologies.

The main criteria for the design and operation of any dynamic (as opposed to static) inhalation exposure system are the following:

- The concentration of the test atmosphere must be reasonably uniform throughout the chamber and should increase and decrease at a rate close to theoretical at the start or end of the exposure. Silver (1946) showed that the time taken for a chamber to reach a point of equilibrium was proportional to the flow rate of atmosphere passing through the chamber and the chamber volume. From this, the concentration–time relationship during the “run-up” and “run-down” phase could be expressed by the equation

$$t_x = k \frac{V}{F}$$

where t_x = time required to reach $x\%$ of the equilibrium concentration, k = a constant of value determined by the value of x , V = chamber volume, and F = chamber flow rate. The t_{99} value is frequently quoted for exposure chambers, representing the time required to reach 99% of the equilibrium concentration and providing an estimate of chamber efficiency. Thus, at maximum efficiency, the theoretical value of k at t_{99} is 4.605, and the closer to this that the results of evaluation of actual chamber performance fall, the greater the efficiency and the better the design of the chamber.

- Flow rates must be controlled in such a way that they are not excessive, which might cause streaming effects within the chamber, but must be adequate to maintain normal oxygen levels, temperature, and humidity in relation to the number of animals being exposed. A minimum of 10 air changes per hour is frequently advocated and is appropriate in most cases. However, the chamber design and housing density also need to be taken into account, and some designs, such as that of Doe and Tinston (1981), function effectively at lower air change rates.
- The chamber or exposure manifold materials should not affect the chemical or physical nature of the test atmosphere.

TABLE 20.2 Advantages, Disadvantages, and Considerations Associated with Patterns of Inhalation Exposure

| Mode of Exposure | Advantages | Disadvantages | Design Considerations |
|--|--|--|--|
| Whole body | Variety and number of animals Chronic studies possible Minimum restraint Large historical database Controllable environment Minimum stress Minimum labor | Messy Multiple routes of exposure: skin, eyes, oral Variability of “dose” Cannot pulse exposure easily Poor contact between animals and investigators Capital intensive Inefficient compound usage Difficult to monitor animals during exposure | Cleaning effluent air Inert materials Losses of test material Even distribution in space Sampling Animal care Observation Noise, vibration, humidity Air temperature Safe exhaust Loading Reliability |
| Head only | Good for repeated exposure Limited routes of entry into animal More efficient dose delivery | Stress to animal Losses can be large Seal around neck Labor in loading/unloading | Even distribution Pressure fluctuations Sampling and losses Air temperature, humidity Animal comfort Animal restraint Pressure fluctuations |
| Nose/mouth only | Exposure limited to mouth and respiratory tract Uses less material (efficient) Containment of material Can pulse the exposure | Stress to animal Seal about face Effort to expose a large number of animals | Body temperature Sampling Air locking Animals' comfort Losses in plumbing/masks Air humidity/temperature Stress to the animal Physiologic support |
| Lung only (tracheal administration) | Precision of dose One route of exposure Uses less material (efficient) Can pulse the exposure | Technically difficult Anesthesia or tracheostomy Limited to small numbers Bypasses nose Artifacts in deposition and response Technically more difficult | Stress to animal Physiologic support |
| Partial lung | Precision of total dose Localization of dose Can achieve very high local doses Unexposed control tissue from same animal | Anesthesia Placement of dose Difficulty in interpretation of results Technically difficult Possible redistribution of material within lung | |

Source: Data from Gad and Chengelis (1998).

The whole-body exposure technique is useful for acute and chronic toxicity studies of gases and vapors. For acute whole-body exposure, a few animals are exposed for 1–4 h to a gas, vapor, or aerosol of a drug or chemical in a simple glass jar. The gaseous drug is metered with a precision flow meter into the stream of filtered room air being drawn through the glass jar or chamber. For vapor generation from a volatile liquid, a stream of clean air is bubbled at a constant rate onto the walls of a temperature-regulated flask, which vaporizes the liquid droplets rapidly and continuously. In either method, the vapor emerging from the vaporizer is

directed into the filtered airstream being drawn through the glass jar or chamber. For the generation of drug aerosols from liquids or powders, various types of atomizers or nebulizers and dust generators are available (Drew and Laskin, 1973; Drew and Lippmann, 1978; Leong et al., 1981; Phalen, 2009). For more critical and precision studies, an adequate number of animals per group are calculated by an appropriate statistical method (Gad, 2000), and the exposure is carried out in an elaborate dynamic airflow chamber with precision control of the chamber airflow, temperature, and humidity.

Regardless of the exposure apparatus used, the most important aspect of an exposure study is that the generation of a constant concentration of the airborne drug vapor or aerosol in the chamber atmosphere has to be sampled (Drew and Lippmann, 1978) and analyzed using an appropriate analytical instrument, such as an infrared spectrophotometer for halogenated propellants or a gas chromatograph for other gases and vapors. The concentration of the drug as detected by the analyzer is the "analytical concentration." For characterizing the aerosol atmosphere, particle sizing, in addition to concentration analyses, is essential. Because the breathing patterns of the experimental animals cannot be regulated, it is extremely important to generate aerosols of the appropriate size for bioavailability.

For critical laboratory studies on inhaled drugs, a mono-disperse aerosol of a specified range of MMAD should be used to increase the probability that the aerosol reaches the specified target area of the lungs. The Dautrebande aerosol generators (Dautrebande, 1962c) and the DeVilbiss nebulizer (Drew and Lippmann, 1978) are the classic single-reservoir generators for short-duration inhalation studies. For long-duration inhalation studies, the multiple-reservoir nebulizer (Miller et al., 1981) and the continuous syringe metering and elutriating atomizer (Leong et al., 1981) are frequently used. The nebulizers generate a polydisperse droplet aerosol either by the shearing force of a jet of air over a fine stream of liquid or by ultrasonic disintegration of the surface liquid in a reservoir (Drew and Lippmann, 1978). The aerosols emerging from a jet nebulizer generally have MMADs ranging between 1.2 and 6.9 μm with GSDs of 1.7–2.2, and aerosols from an ultrasonic nebulizer have MMADs ranging between 3.7 and 18.5 μm with GSDs of 1.4–2.0 (Mercer, 1981).

For testing therapeutic formulations, the liquid aerosols are usually generated by the pressurized metered-dose inhaler (MDI) (Newman, 1984; Gad and Chengelis, 1998; Newton, 2000; Gad, 2006). The pressurized MDI generates a bolus of aerosols by atomizing a well-defined quantity of a drug that is solubilized in a fluorocarbon propellant. The aerosols, thus, consist of the drug particles with a coating of the propellant. As the aerosols emerge from the orifice, the mean particle size may be as large as 30 μm (Moren, 1981). After traveling through a tubular or cone-shaped spacer, the propellant may evaporate, reducing the MMADs to a range of 2.8–5.5 μm with GSDs of 1.5–2.2 (Hiller et al., 1978; Sackner et al., 1981; Newman, 1984) and making the aerosols more stable for inhalation studies. In a prolonged animal exposure study, multiple MDIs have to be actuated sequentially with an electromechanical gadget (Ulrich et al., 1984) to maintain a slightly pulsatile but relatively consistent chamber concentration.

For generating an aerosol from dry powders, various dust generators, such as the Wright dust feed, air elutriator or fluidized-bed dust generator, and air impact pulverizer, have been developed for acute and chronic animal inhalation

studies and described in many articles (Hinds, 1980; Leong et al., 1981; Gad and Chengelis, 1998; Hext, 2000; Phalen, 2009; Brown et al., 2014). For generating powdery therapeutic agents, a metered-dose dry powder inhaler, Spinhaler or Rotahaler, is used (Newman, 1984). The particle size of the drug powder is micronized to a specific size range during manufacture, and the Spinhaler or the Rotahaler only disperses the powders.

More recently, another approach for administering dry powders to both humans and test animals has arisen. Dry powders, while less frequently used in nasal drug delivery, are becoming more popular. Powders can be administered from several devices, the most common being the insufflator. Many insufflators work with predosed powder in gelatin capsules. To improve patient compliance a multidose powder inhaler has been developed which has been used to deliver budesonide. These devices can also be used for administration to test animal delivery, both in terms of amounts and aerodynamic size of the particles. While early dry powder inhalers such as the Rotahaler® used individual capsules of micronized drug which were difficult to handle, modern devices use blister packs (e.g., Diskus®) or reservoirs (e.g., Turbuhaler®). The dry powder inhalers rely on inspiration to withdraw drug from the inhaler to the lung, and hence the effect of inhalation flow rate through various devices has been extensively studied. The major problem to be overcome with these devices is to ensure that the finely micronized drug is thoroughly dispersed in the airstream. It has been recommended that patients inhale as rapidly as possible from these devices in order to provide the maximum force to disperse the powder. The quantity of drug and deposition pattern varies enormously depending on the device, for example, the Turbuhaler produces significantly greater lung delivery of salbutamol than the Diskus. Vidgren and coworkers (1987) demonstrated by gamma scintigraphy that a typical dry powder formulation of SCG suffers losses of 44% in the mouth and 40% in the actuator nozzle itself.

It must also be emphasized that the major mass of a heterodispersed aerosol may be contained in a few relatively large particles, since the mass of a particle is proportional to the cube of its diameter. Therefore, the particle-size distribution and the concentration of the drug particles in the exposure atmosphere should be sampled using a cascade impactor or membrane filter sampling technique, monitored using an optical or laser particle-size analyzer, and analyzed using optical or electron microscopy techniques.

In summary, many techniques have been developed for generating gas, vapor, and aerosol atmospheres for inhalation toxicology studies. By proper regulation of the operating conditions of the nebulizers and the formulation of MDIs, together with the use of spacer or reservoir attachments to MDIs, more particles within the respirable range can be generated for inhalation. An accurately controlled exposure concentration is essential to an accurate determination of the

dose–response relationship in a safety assessment of an inhalant drug.

Finally, comparisons of various techniques for animal exposures indicate that the whole-body exposure technique is the most suitable for safety assessment of gases and vapors and permits simultaneous exposure of a large number of animals to the same concentration of a drug; however, this technique is not suitable for aerosol and powder exposures because the exposure condition represents the resultant effects from inhalation, ingestion, and dermal absorption of the drug (Gad and Chengelis, 1998; Phalen, 2009).

20.10 THE UTILITY OF TOXICITY DATA

Regardless of the type of test and the parameters to be monitored, the ultimate goal is to interpolate or extrapolate from the dose–response data to find a no-observable-adverse-effect level (NOAEL) or a no-observable-effect level (NOEL). By applying a safety factor of 1–10 to the NOAEL, a safe single-exposure dose for a phase I clinical trial may be obtained. By applying a more stringent safety factor, a multiple-exposure dose for a clinical trial may also be obtained. After the drug candidate has successfully passed all the drug safety evaluations and entered in the production stage, more toxicity tests may be needed for the establishment of a threshold limit value–time-weighted average (TLV-TWA). A TLV-TWA is defined as “the time weighted average concentration for a normal 8-h workday and a 40-h workweek, to which nearly all workers may be repeatedly exposed, day after day, without adverse effect” (ACGIH, 1991). Using TLVs as guides, long-term safe occupational exposures during production and industrial handling of a drug may be achieved. Appropriate safety assessments of pharmaceutical chemicals and drugs will ensure the creation and production of a safe drug for the benefit of humans and animals. Further, inhalation toxicity data are needed for compliance with many regulatory requirements of the FDA, the Occupational Safety and Health Administration, and the Environmental Protection Agency (Gad and Chengelis, 1998).

20.11 FORMULATION AND POTENTIAL MUCOSAL DAMAGE

- Improved absorption involves interactions with the mucosal membrane.
- Proposed enhancement mechanisms are:
 - Extraction of membrane components
 - Penetration and fluidization of membrane
 - Loosening of tight junctions
 - Perturbation of nasal mucociliary clearance system
 - Simultaneous transport of environmental toxins

- Adverse effects have to be of short duration, mild, and rapidly reversible.
- Kinetics of lipid and protein extraction from the membrane are measures of the extent of damage evaluated by the measuring activity of membrane marker enzymes:
 - Lactate dehydrogenase: cytosolic enzyme, related to intracellular damage
 - 5'-Nucleotidase: membrane-bound enzyme, indicator of membrane perturbations
 - Alkaline phosphatase: membrane-bound enzyme, related to membrane damage
- Ideal characteristics of absorption enhancers include the following:
 - Pharmacological inertness
 - Nonirritant, nontoxic, and nonallergenic
 - Effect on nasal mucosa that should be transient and completely reversible
 - Potent in low concentrations
 - Compatible with other adjuvants
 - Has no offensive odor or taste
 - Inexpensive and readily available
- The factors influencing mucosal damage include:
 - Drug administration
 - Dose
 - Frequency
 - Interspecies difference
 - Sensitivity toward absorption enhancers

20.11.1 Methods to Assess Irritancy and Damage

- Erythrocytes
 - Used to study the membrane activity of absorption enhancers
- Histology
 - Histological studies of nasal membranes
- Intracellular protein release
 - Index of cellular damage due to exposure to absorption enhancers
- Tolerability
 - These are subjective (double-masked) studies in which individuals report any effects due to the use of enhancers in the formulation.
- Cilia A function
 - Cilia beat frequency is obtained from tissue samples at sacrifice using video capture systems.
 - Tissues used for ciliary function studies include chicken embryo trachea, cryopreserved human mucosa taken from sphenoidal sinus, rat nasal mucosa, and recently human nasal epithelial cells.

More comprehensive descriptions and discussions on inhalation toxicology and technology may be found in several recent monographs, reviews, and textbooks (Willeke, 1980; Leong et al., 1981; Witschi and Nettekheim, 1982; Clarke and Pavia, 1984; Witschi and Brain, 1985; Barrow, 1986; McFadden, 1986; Leikauf, 2013; Salem, 1986; Gardner, 2006; McClellan and Henderson, 1989; Gad and Chengelis, 1998; Hext, 2000; Brown et al., 2015; Gardner, 2006; Phalen, 2009; Salem and Katy, 2015).

US FDA regulation of new inhalation drug products is by the Division of Pulmonary, Allergy, and Rheumatology Products (DPAAP). Like its counterpart for dermal route drugs (Dermal and Dental Products (DDP)), actual general/systemic toxicity testing must be performed using the drug in the intended clinical formulation to actually be given to man (FDA, 2002, 2008).

20.12 THERAPEUTIC DRUG DELIVERY BY THE DERMAL ROUTE

The dermal route has both some significant attractions and some significant drawbacks for use in drug administration. The ability to utilize the former and avoid (at least the relevance of) the latter is the key to successful development of new drugs administered by this route.

On the advantage side, dermal administration:

- Avoids first-pass metabolism
- Avoids acidic environment of the stomach
- Can be designed to deliver controlled amounts of drug over a prolonged period outside of a clinical setting
- Can be designed to achieve desired local tissue exposure to the therapeutic area while avoiding (or minimizing) systemic exposure

On the disadvantage side, dermal administration historically:

- Cannot be used for systemic delivery of large molecules (usually with molecular weights in excess of 1000)
- Achieves a lower systemic bioavailability than oral routes
- Uses formulation components which may irritate or damage the application site (Wilhelm et al., 2012) and are perceived as “messy” or aesthetically undesirable

New approaches to formulation or delivery systems (see Toutou and Barry, 2007) have overcome (or minimized) a number of the disadvantages and continue to evolve. For FDA, the division reviewing proposed new dermal products (DDP) requires that even systemic (general toxicology)

studies be performed using intensely clinical formulation product, which presents a challenge to toxicologists. In addition to the traditional gels, lotions, and creams (for which new chemical permeation enhancers and vesicular carriers such as liposomes have and are being added), new technology such as iontophoresis, electrophoresis, and ultrasound is also evolving. Additionally, in some cases, the barrier layer in the skin (stratum corneum) can be bypassed (using such approaches as microneedles) or removed by tape stripping, laser, adhesion, and ballistic methods (such as those used for mass vaccine administration in the military).

Bioavailability is typically defined as the rate and extent at which a drug reaches the general circulation from an administered dosage form. Dermatological drug products include preparations which are designed to exert a local effect in diseased skin following topical application on the skin surface. The objective is to maximize drug concentration at the site of action within the skin with, ideally, a minimal systemic uptake. Thus, systemic availability may not properly reflect local cutaneous bioavailability (as it does for transdermal products which are designed to deliver drug into the systemic circulation). Moreover, topical doses tend to be so small (typically 2–5 mg of product per square centimeter) that serum and/or urine concentrations are often undetectable using conventional assay techniques. Further complicating this is the lack of knowledge of the drug concentration needed at the skin target site (with the exception of antifungal and antibacterial agents whose target site is the SC surface). Topical bioavailability has been more properly defined as the temporal pattern of free drug, but this approach remains largely theoretical due to the difficulty of quantifying drug within the skin. deHeer et al. (1999) present clear guidance on estimating dermal absorption of products.

Available options include estimating a drug's permeability coefficient through human skin from the molecular weight and octanol–water partition coefficient. This information is not really sufficient to estimate topical bioavailability. The algorithms available are only able to approximate values for drug bioavailability (deHeer et al., 1999).

Alternatives include collecting samples by tape stripping or biopsy and then measuring the actual drug (and metabolites) present in the tissue samples. There are numerous complications to this and approaches to solving them (see Herkenne et al., 2008).

There are no direct guidance documents focused on nonclinical safety assessment for topical route drugs or on what must be done before taking such a drug into clinical trials. Rather, the FDA regulatory expectations come from ICH/M4, the 2008 CDER reformulation document (CDER, 2008), and current practice (Jacobs and Loan, 2012). The results are summarized in Table 20.3. Regulatory preference as to a model for dermal systemic toxicology is solidly the minipig (McAnulty et al., 2012), though the rabbit and guinea pig still see use.

TABLE 20.3 Test Requirement Matrix for Topical Agents

| Test Requirement | Species |
|---|-----------------|
| <i>Initial clinical trial/IND requirements</i> | |
| 1. Seven-day DRF toxicity two phase (single-dose MTD and non-GLP pilot in rodents) | R/M |
| 2. Seven-day DRF toxicity two phase (single-dose MTD and non-GLP pilot in nonrodents) | D/S/P |
| 3. Genotoxicity: bacterial mutagenicity (Ames) ^a | <i>In vitro</i> |
| 4. Genotoxicity: <i>in vitro</i> clastogenicity (CHO chromosome aberration) (note that the ICH S2 guidance option 2 allows the use of alternative studies) ^a | <i>In vitro</i> |
| 5. Genotoxicity: <i>in vivo</i> (mouse or rat micronucleus) ^a | R/M |
| 6. Safety pharmacology: CV-hERG (recommended but not required) ^b | <i>In vitro</i> |
| 7. Safety pharmacology: CV <i>in vivo</i> ^a | D/P/S |
| 8. Safety pharmacology: FOB/Irwin ^a | R/M |
| 9. Safety pharmacology: respiratory—rodent ^a | R |
| 10. Pivotal/repeat dose in rodents (14–28-day intended route ^c) | R/M |
| 11. Pivotal/repeat dose in nonrodents (14–28-day intended route ^c) | D/P/S |
| 12. CYP induction/inhibition ^b | <i>In vitro</i> |
| 13. Five-species microsome metabolic panel ^b | <i>In vitro</i> |
| 14. Develop bioanalytical for three species (man/rodent/nonrodent) | NA |
| 15. Local irritation ^c (clinical formulation) | R |
| 16. Of dermal—sensitization | G |
| <i>To support continued clinical development</i> | |
| 17. Developmental tox (seg II)—rat and rabbit pilots and rat and rabbit studies | |
| 18. Immunotoxicity ^c (if immune modulatory claim or there are findings in 14/28 dog studies) | |
| 19. Pivotal/repeat dose in rodents (3/6-month oral) | |
| 20. Pivotal/repeat dose in nonrodents (3/9–12-month oral) | |
| <i>To support marketing approval</i> | |
| 21. Reproductive toxicity—seg I | |
| 22. Reproductive toxicity—seg III | |
| 23. Tumorigenicity/carcinogenicity—rat | |
| 24. Tumorigenicity/carcinogenicity—mouse | |

All studies described in the preceding text must be performed GLP.

Note: Species: R=rat, M=mouse, D=dog, S=pig, P=primate, B=rabbit, G=guinea pig, TBD=to be determined.

^a May be required.

^b Recommended.

^c Dermal/ophthalmic/vaginal/rectal.

The CDER (2008) document presents that route-specific expectations for dermal include the following:

- Delayed hypersensitivity of any previously not evaluated topical drug or new formulation. In the United States, this means performing the LLNA, GPMT, or Buehler sensitization tests. As the Buehler is not accepted in Europe, use of one of the other two (see the chapter on immunotoxicology for details on these and their limitations) should be used.
- Photoirritation should be evaluated if the new formulation absorbs ultraviolet or visible radiation (290–700nm) and if the product is applied to sun-exposed skin. If the new formulation is a patch, then photoirritation should be considered if the patch is permeable to light and is applied to sun-exposed skin (see details in local tissue tolerance). A screen using 3T3 cells is acceptable in place of the traditional rabbit study for evaluation prior to IND.
- If the new formulation contains an active ingredient that has not been used by the dermal route, the repeat-dose local toxicity study mentioned earlier should be conducted in a nonrodent species (preferably the minipig). This study should be of at least the same duration as clinical use (up to 9 months) and include both local and systemic evaluation. For NCEs, repeat-dose studies in both rodents and nonrodents (preferably the pig) are required. In one of these (typically the rodent), it should be ensured that sufficient systemic exposure of the NCE is achieved usually by conducting such a study by the oral or a parenteral route.
- The administered skin dose from topically applied drug products can be orders of magnitude larger than the skin dose after systemic administration. Therefore, a dermal carcinogenicity study might be recommended for drugs with a chronic indication even if systemic carcinogenicity studies are available.

- The photocarcinogenic potential should be evaluated if the new formulation is used chronically on sun-exposed skin. Evaluation of photocarcinogenicity generally is not recommended for path products (see guidance for industry *photosafety testing*).
- Nonclinical dermal studies generally should be conducted with untreated control, vehicle control, and complete formulation groups.

A key point is if significant systemic exposure is or is not achieved. If it is not, then genotoxicity and safety pharmacology testing is not required. Such testing is required if it cannot be shown that systemic absorption is minimal or not detectable.

For the repeat-dose systemic toxicity studies, added to the usual details of study design is the need to evaluate local tissue effects of application sites and to collect tissue and histopathologically evaluate such sites.

As a last point, it should be added that topical administration is not limited to the dermal route. It also includes the body surfaces (vaginal and rectal) for which evaluation of irritation and hypersensitivity is required and for which systemic toxicology repeat-dose studies are performed by administration to these sites (but there is no phototoxicity concern) and ocular for which it is expected that:

- The dermal irritation and delayed contact hypersensitivity potential of the new formulation should be evaluated (because dermal exposure will occur with such drugs).
- The local tissue (eye) irritation must be evaluated.

REFERENCES

- ACGIH (1991) *Threshold limit values and biological exposure indices*. American Conference of Governmental Industrial Hygienists (ACGIH), Cincinnati.
- Adams EM, Spencer HC, Rowe VK, Irish DD. (1950) Vapor toxicity of 1,1,1-trichloroethane (methylchloroform) determined by experiments on laboratory animals. *Arch. Ind. Hyg. Occup. Med.* 1:225–236.
- Adams EM, Spencer HC, Rowe VK, McCollister DD, Irish DD. (1952) Vapor toxicity of carbon tetrachloride determined by experiments on laboratory animals. *Arch. Ind. Hyg. Occup. Med.* 6:50–66.
- Agnew JE. (1984) Physical properties and mechanisms of deposition of aerosols. In: *Aerosols and the Lung, Clinical Aspects* (Clarke SW and Pavia D, Eds.). Butterworth, London, pp. 49–68.
- Alarie Y. (1966) Irritating properties of airborne material to the upper respiratory tract. *Arch. Environ. Health.* 13:433–449.
- Alarie Y. (1973) Sensory irritation by airborne chemicals. *CRC Crit. Rev. Toxicol.* 2:299–363.
- Alarie Y. (1981a) Toxicological evaluation of airborne chemical irritants and allergens using respiratory reflex reactions. In: *Inhalation Toxicology and Technology* (Leong BKJ, Ed.). Ann Arbor Science, Ann Arbor, MI, pp. 207–231.
- Alarie Y. (1981b) Bioassay for evaluating the potency of airborne sensory irritants and predicting acceptable levels of exposure in man. *Food Cosmet. Toxicol.* 19:623–626.
- Alarie Y, Luo JE. (1986) Sensory irritation by airborne chemicals: a basis to establish acceptable levels of exposure. In: *Toxicology of the Nasal Passages* (Barrow CS, Ed.). Hemisphere, New York, pp. 91–100.
- Alarie Y, Kane L, Barrow C. (1980) Sensory irritation: the use of an animal model to establish acceptable exposure to airborne chemical irritants. In: *Toxicology: Principles and Practice 1* (Reeves AL, Ed.). John Wiley & Sons, New York.
- Altman PL, Dittmer DS. (1974) *Biological Data Book*, Vol. III. Federation of American Societies for Experimental Biology, Bethesda, MD.
- Andersen ME, French JE, Gargas ML, Jones RA, Jenkins LJ Jr. (1979) Saturable metabolism and the acute toxicity of 1, 1-dichloroethylene. *Toxicol. Appl. Pharmacol.* 47:385–393.
- Aso Y, Yoneda K, Kikkawa Y. (1976) Morphologic and biochemical study of pulmonary changes induced by bleomycin in mice. *Lab. Invest.* 35:558–568.
- Aviado DM. (1981) Comparative cardiotoxicity of fluorocarbons. In: *Cardiac Toxicology* (Balazs T, Ed.), Vol. II. CRC Press, Boca Raton, FL, pp. 213–222.
- Aviado DM, Micozzi MS. (1981) Fluorine-containing organic compounds. In: *Patty's Industrial Hygiene and Toxicology* (Clayton GD and Clayton FE, Eds.), Vol. 2B. John Wiley & Sons, Inc., New York, pp. 3071–3115.
- Balazs T. (1981) Cardiotoxicity of adrenergic bronchodilator and vasodilating antihypertensive drugs. In: *Cardiac Toxicology* (Balazs T, Ed.), Vol. II. CRC Press, Boca Raton, FL, pp. 61–73.
- Barrow CS. (1986) *Toxicology of the Nasal Passages*. Hemisphere, New York.
- Bell KA. (1978) Local particle deposition in respiratory airway models. In: *Recent Developments in Aerosol Science* (Shaw DT, Ed.). John Wiley & Sons, Inc., New York, pp. 97–134.
- Bice DE. (1985) Methods and approaches to assessing Immunotoxicology of the lower respiratory tract. In: *Immunotoxicology and Immunopharmacology* (Dean JH, Luster MI, Munson AE and Amos HA, Eds.). Raven Press, New York, pp. 145–157.
- Brain JD. (1971) *The effects of increased particles on the number of alveolar macrophages*. Inhaled Particles III, Proceedings of an international symposium organized by the British Occupational Hygiene Society in London, September 14–23, 1970, Vol. 1, pp. 200–233, BOHS Symposium. Unwin Brothers Ltd, London.
- Brain JD, Knudson DE, Sorokin SP, Davis MA. (1976) Pulmonary distribution of particles given by intratracheal instillation or by aerosol inhalation. *Environ. Res.* 11:13–33.
- Brown JD, Kreyling WG, Godleski JJ. (2014) Inhalation toxicology. In: *Principles and Methods of Toxicology* (Hayes W, Ed.), 6th Ed. Raven Press, New York, pp. 1385–1444.
- Budavari S. (Ed.) (1989) *Merck Index*, 11th Ed. Merck & Co., Rahway, NJ, p. 82.

- Carpenter CP, Smyth HF Jr, Pozzani UC. (1949) The assay of acute vapor toxicity, and grading and interpretation of results on 96 chemical compounds. *J. Ind. Hyg. Toxicol.* 31:343–346.
- CDER. (2008) *Guidance for Industry and Review Staff: Nonclinical Safety Evaluation of Reformulated Drug Products and Products Intended for Administration by an Alternate Route*. U.S. Department of Health and Human Services, Rockville, MD.
- Chan TL, Lippmann M. (1980) Experimental measurements and empirical modelling of the regional deposition of inhaled particles in humans. *Am. Ind. Hyg. Assoc. J.* 41:399–409.
- Chenoweth MB, Leong BKJ, Sparschu GL, Torkelson TR. (1972) Toxicities of methoxyflurane, halothane and diethyl ether in laboratory animals on repeated inhalation at subanesthetic concentrations. In: *Cellular Biology and Toxicity of Anesthetics* (Fink BR, Ed.). Williams & Wilkins, Baltimore, MD, pp. 275–284.
- Clarke SW, Pavia D. (Eds.) (1984) *Aerosol and the Lung: Clinical and Experimental Aspects*. Butterworth, London.
- Cox JSG, Beach JE, Blair AMJN, Clarke AJ, King J, Lee TB, Loveday DEE, Moss GF, Orr TSC, Ritchie JT, Sheard P. (1970) Disodium cromoglycate. *Adv. Drug Res.* 5:115–196.
- Curry SH, Taylor AJ, Evans S, Godfrey S, Zeidifard E. (1975) Disposition of disodium cromoglycate administered in three particle sizes. *Br. J. Clin. Pharmacol.* 2:267–270.
- Dahl AR. (1990) Dose concepts for inhaled vapors and gases. *Toxicol. Appl. Pharmacol.* 103:185–197.
- Dallas CE, Bruckner JV, Maedgen JL, Weir FW. (1986) A method for direct measurement of systemic uptake and elimination of volatile organics in small animals. *J. Pharmacol. Methods.* 16:239–250.
- Dallas CE, Ramanathan R, Muralidhara S, Gallo JM, Bruckner JV. (1989) The uptake and elimination of 1,1,1-trichloroethane during and following inhalation exposure in rats. *Toxicol. Appl. Pharmacol.* 98:385–397.
- Dautrebande L. (1962a) Importance of particle size for therapeutic aerosol efficiency. In: *Microaerosols* (Dautrebande L, Ed.). Academic Press, New York, pp. 37–57.
- Dautrebande L. (1962b) Practical recommendation for administering pharmacological aerosols. In: *Microaerosols* (Dautrebande L, Ed.). Academic Press, New York, pp. 86–92.
- Dautrebande L. (1962c) Production of liquid and solid micromellar aerosols. In: *Microaerosols* (Dautrebande L, Ed.). Academic Press, New York, pp. 1–22.
- Davies RJ, Pepys J. (1975) Asthma due to inhaled chemical agents—the macrolide antibiotic spiramycin. *Clin. Allergy.* 5:99–107.
- Davis CN. (1961) A formalized anatomy of the human respiratory tract. In: *Inhaled Particles and Vapours* (Davis CN, Ed.). Pergamon Press, London, pp. 82–91.
- Davis CN, Heyder J, Subba Ramv MC. (1972) Breathing of half micron aerosols. I. Experimental. *J. Appl. Physiol.* 32:591–600.
- Davis B, Marin MG, Fischer S, Graf P, Widdicombe JG, Nadel JA. (1976) New method for study of canine mucus gland secretion *in vivo*: cholinergic regulation. *Am. Rev. Respir. Dis.* 113:257 (abstract).
- Dennis WL. (1961) The discussion of a paper by C.N. Davis: a formalized anatomy of the human respiratory tract. In: *Inhaled Particles and Vapours* (Davis CN, Ed.). Pergamon Press, London, p. 88.
- Doe JE, Tinston DJ. (1981) Novel chamber for long-term inhalation studies. In: *Inhalation Toxicology and Technology* (Leong KJ, Ed.). Ann Arbor Science, Ann Arbor, MI.
- Drew RT, Laskin S. (1973) Environmental inhalation chambers. In: *Methods of Animal Experimentation* (Gay WI, Ed.), Vol. IV. Academic Press, New York, pp. 1–41.
- Drew RT, Lippmann M. (1978) Calibration of air sampling instruments. In: *ACGIH Air Sampling Instruments for Evaluation of Atmospheric Contaminants*, 5th Ed. American Conference of Governmental Industrial Hygienists, Cincinnati, OH, Section I, pp. 1–32.
- Elliot GA, DeYoung EN. (1970) Intranasal toxicity testing of antiviral agents. *Ann. N. Y. Acad. Sci.* 173:169–175.
- Evans MJ. (1982) Cell death and cell renewal in small airways and alveoli. In: *Mechanisms in Respiratory Toxicology* (Witschi H and Nettesheim P, Eds.), Vol. 1. CRC, Boca Raton, FL, pp. 189–218.
- FDA. (2002) *Guidance for Industry: Nasal Spray and Inhalation Solution, Suspension, and Spray Drug Products—Chemistry, Manufacturing and Controls Documentation*. U.S. Department of Health and Human Services, Food and Drug Administration, Center for Drug Evaluation and Research, Rockville, MD.
- FDA. (2008) *Guidance for Industry and Review Staff: Nonclinical Safety Evaluation of Reformulated Drug Products and Products Intended for Administration by an Alternate Route*. CDER, Rockville, MD.
- Ferin J. (1977) Effect of particle content of lung on clearance pathways. In: *Pulmonary Macrophage and Epithelial Cells* (Sanders CL, Schneider RP, Dagle GE and Ragan HA, Eds.). Technical Information Center, Energy Research & Development Administration, Springfield, VA, pp. 414–423.
- Fieldner AC, Kazt SH, Kinney SP. (1921) *Gas Masks for Gases Met in Fighting Fires*. Technical Paper No. 248. U.S. Bureau of Mines, Washington, DC.
- Gad S. (2000) *Statistics and Experimental Design for Toxicologists*, 3rd Ed. CRC Press, Boca Raton, FL.
- Gad SC. (2006) Safety assessment of therapeutic agents administration by the respiratory route. In: *Toxicology of the Lung* (Garlner DE, Ed.), 4th Ed. Taylor & Francis, New York.
- Gad SE. (2015) Safety assessment of therapeutic agents administered by the respiratory route. In: *Inhalation Toxicology* (Salem H and Katz SA, Eds.), 3rd Ed. CRC Press, Boca Raton, FL, pp. 537–580.
- Gad S, Chengelis CP. (1998) *Acute Toxicology Testing: Perspectives and Horizons*, 2nd Ed. Academic Press, San Diego, CA, pp. 404–466.
- Gamsu G, Singer MM, Vincent HH, Berry S, Nadel JA. (1976) Postoperative impairment of mucous transport in the lung. *Am. Rev. Respir. Dis.* 114:673–679.
- Gardner DE. (Ed.) (2006) *Toxicology of the Lung*, 4th Ed. Taylor & Francis, New York.
- Godfrey S, Zeidifard E, Brown K, Bell JH. (1974) The possible site of action of sodium cromoglycate assessed by exercise challenge. *Clin. Sci. Mol. Med.* 46:265–272.

- Haber FR. (1924) *Fünf Vorträge aus den Jahren 1920–23, No. 3*, Die Chemie im Kriege. Julius Springer, Berlin.
- Hafner RE Jr, Watanabe PG, Gehring PJ. (1975) Preliminary studies on the fate of inhaled vinyl chloride monomer in rats. *Ann. N. Y. Acad. Sci.* 246:135–148.
- Hatch TF, Gross P. (1964) *Pulmonary Deposition and Retention of Inhaled Aerosols*. Academic press, New York, pp. 16–17 51–52, 147–168.
- deHeer C, Wilschut A, Stevenson H, Halpert BC. (1999) *Guidance Document on the Estimation of Dermal Absorption According to a Tiered Approach: An Update*. TNO Nutrition and Food Research Institute, Zeist.
- Henderson RF. (1984) Use of bronchoalveolar lavage to detect lung damage. *Environ. Health Perspect.* 56:115–129.
- Henderson R. (1988) Use of bronchoalveolar lavage to detect lung damage. In: *Toxicology of the Lung* (Gardner DE, Crapo JD and Massaro EJ, Eds.). Raven Press, New York, pp. 239–268.
- Henderson R. (1989) Bronchoalveolar lavage: a tool for assessing the health status of the lung. In: *Concepts in Inhalation Toxicology* (McClellan RO and Henderson RF, Eds.). Hemisphere, Washington, DC, pp. 414–442.
- Henderson RF, Loery JS. (1983) Effect of anesthetic agents on lavage fluid parameters used as indicators of pulmonary injury. *Lab. Anim. Sci.* 33:60–62.
- Henderson RF, Benson JM, Hahn FF, Hobbs CH, Jones RK, Mauderly JL, McClellan RO, Pickrell JA. (1985) New approaches for the evaluation of pulmonary toxicity: bronchoalveolar lavage fluid analysis. *Fundam. Appl. Toxicol.* 5:451–458.
- Henderson RF, Mauderly JL, Pickrell JA, Hahn FF, Muhle H, Rebar AH. (1987) Comparative study of bronchoalveolar lavage fluid: effect of species, age, method of lavage. *Exp. Lung Res.* 1:329–342.
- Henderson RF, Barr EB, Hotchkiss JA. (1991) Effect of exposure rate on response of the lung to inhaled particles (abstr.) *Toxicologists*. 11:234.
- Hensley MJ, O'Cain CF, McFadden ER Jr, Ingram RH Jr. (1978) Distribution of bronchodilation in normal subjects: beta agonist versus atropine. *J. Appl. Physiol.* 45:778–782.
- Herkenne C, Alberti I, Naik A, Kalia YN, Mathy F-X, Préat V, Guy RH. (2008) *In vivo* methods for the assessment of topical drug bioavailability. *Pharm. Res.* 25(1):87–103.
- Hext PM. (2000) Inhalation toxicology. In: *General and Applied Toxicology* (Ballantyne B, Marrs T and Syversen T, Eds.). Macmillan, London, pp. 587–601.
- Heyder J, Gebhart J, Stahlhofen W. (1980) Inhalation of aerosols: particle deposition and retention. In: *Generation of Aerosols and Facilities for Exposure Experiments* (Willeke K, Ed.). Ann Arbor Science, Ann Arbor, MI, pp. 80–99.
- Hiller FC, Mazunder MK, Wilson JD, Bone RC. (1978) Aerodynamic size distribution of metered dose bronchodilator aerosols. *Am. Rev. Respir. Dis.* 118:311–317.
- Hinds WC. (1980) Dry-dispersion aerosol generators. In: *Generation of Aerosols and Facilities for Exposure Experiments* (Willeke K, Ed.). Ann Arbor Science, Ann Arbor, MI, pp. 171–187.
- Hocking WG, Golde DW. (1979) The pulmonary alveolar macrophage. *N. Engl. J. Med.* 310:580–587 639–645.
- Hodson ME, Penketh AR, Batten JC. (1981) Aerosol carbenicillin and gentamicin treatment of *Pseudomonas aeruginosa* infection in patients with cystic fibrosis. *Lancet*. 2:1137–1139.
- Hollinger MA. (1985) Mucokinetic agents, radionuclides, and anti-coagulant-thrombolytic agents. In: *Respiratory Pharmacology and Toxicology*. WB Saunders, Philadelphia, pp. 133–134.
- Holma B. (1967) Lung clearance of mono- and di-disperse aerosols determined by profile scanning and whole body counting: a study on normal and SO₂ exposed rabbits. *Acta Med. Scand. Suppl.* 473:1–102.
- Horsfield K, Cuning G. (1968) Morphology of the bronchial tree in man. *J. Appl. Physiol.* 24:373–383.
- Ingram RH, Wellman JJ, McFadden ER Jr, Mead J. (1977) Relative contributions of large and small airways to flow limitation in normal subjects before and after atropine and isoproterenol. *J. Clin. Invest.* 59:696–703.
- Irish DD, Adams EM. (1940) Apparatus and methods for testing the toxicity of vapors. *Ind. Med. Surg.* 1:1–4.
- Jacob A, Loan JW. (2012) Regulatory aspects: pharmaceuticals, medical devices, food additives: minipigs and global safety regulations. In: *The Minipig in Biomedical Research* (McAnulty PA, Dayan AD, Ganderup N-C and Hastings KL, Eds.). CRC Press, Boca Raton, FL, pp. 565–572.
- Johnson HG, McNee ML, Johnson MA, Miller MD. (1983) Leukotriene C₄ and dimethylphenylpiperazinium-induced responses in canine airway tracheal muscle contraction and fluid secretion. *Int. Arch. Allergy Appl. Immunol.* 71: 214–218.
- Johnson HG, McNee ML, Braughler JM. (1987) Inhibitors of metal catalyzed lipid peroxidation reactions inhibit mucus secretion and 15 HETE levels in canine trachea. *Prostaglandins Leukot. Med.* 30:123–132.
- Jones JG. (1984) Clearance of inhaled particles from alveoli. In: *Aerosols and the Lung: Clinical and Experimental Aspects* (Clarke SW and Pavia D, Eds.). Butterworth, London, pp. 172–196.
- Karol MH. (1988) Immunologic responses of the lung to inhaled toxicants. In: *Concepts in Inhalation Toxicology* (McClellan RO and Henderson R, Eds.). Hemisphere Publishing, Washington, DC, pp. 403–413.
- Karol MH, Thorne PS. (1988) Hypersensitivity and hyperreactivity. In: *Toxicology of the Lung* (Gardner DE, Crapo JD and Massaro EJ, Eds.). Raven Press, New York, pp. 427–488.
- Karol MH, Hillebrand JA, Thorne PS. (1989) Characteristics of weekly pulmonary hypersensitivity responses elicited in the guinea pig by inhalation of ovalbumin aerosols. In: *Toxicology of the Lung* (Gardner DE, Crapo JD and Massaro EJ, Eds.). Raven Press, New York, pp. 427–448.
- Kavet RI, Brain JD, Levens DJ. (1978) Characteristics of weekly pulmonary macrophages lavaged from hamsters exposed to iron oxide aerosols. *Lab. Invest.* 38:312–319.
- Landy TD, Ramsey JC, McKenna MJ. (1983) Pulmonary physiology and inhalation dosimetry in rats: development of a method and two examples. *Toxicol. Appl. Pharmacol.* 71:72–83.

- Last JA. (1982) Mucus production and ciliary escalator. In: *Mechanisms in Respiratory Toxicology* (Witschi H and Nettekheim P, Eds.), Vol. 1. CRC Press, Boca Raton, FL, pp. 247–268.
- Lauweryns JM, Baert JH. (1977) Alveolar clearance and the role of the pulmonary lymphatics. *Am. Rev. Respir. Dis.* 115:625–683.
- Lee W-C, Wang C-S. (1977) Particle deposition in systems of repeated bifurcating tubes. In: *Inhaled Particles IV* (Walton WH, Ed.). Pergamon Press, Oxford, pp. 49–60.
- Leikauf GD. (2013) Toxic responses of the respiratory system. In: *Casarett and Doull's Toxicology* (Klaassen CD, Casarett LJ and Doull J, Eds.), 8th Ed. McGraw Hill, New York, pp. 691–731.
- Leong BKJ, MacFarland HN. (1965) Pulmonary dynamics and retention of toxic gases. *Arch. Environ. Health.* 11:555–563.
- Leong BKJ, Rop DA. (1989) *The combined effects of an inhalation anesthetic and an analgesic on the electrocardiograms of beagle dogs*. Abstract-475. International Congress of Toxicology, Brighton, July 16–21, 1989. Taylor & Francis, London, p. 159.
- Leong BKJ, Powell DJ, Pochyla GL. (1981) A new dust generator for inhalation toxicological studies. In: *Inhalation Toxicology and Technology* (Leong BKJ, Ed.). Ann Arbor Science, Ann Arbor, MI, pp. 157–168.
- Leong BKJ, Coombs JK, Petzold EN, Hanchar AH, McNee ML. (1985) Endotracheal nebulization of drugs into the lungs of anesthetized animals (abstr.) *Toxicologist.* 5:31.
- Leong BKJ, Lund JE, Groehn JA, Coombs JK, Sabaitis CP, Weaver RJ, Griffin L. (1987) Retinopathy from inhaling 4,4'-methylenedianiline aerosols. *Fundam. Appl. Toxicol.* 9:645–658.
- Leong BKJ, Coombs JK, Petzold EN, Hanchar AJ. (1988) A dosimetric endotracheal nebulization technique for pulmonary metabolic disposition studies in laboratory animals. *Inhal. Toxicol.* 1(Premier issue):37–51.
- Lichtiger M, Landa JF, Hirsch JA. (1975) Velocity of tracheal mucus in anesthetized women undergoing gynaecologic surgery. *Anesthesiology.* 42:753–756.
- Lippmann M. (1970) "Respirable" dust sampling. *Am. Ind. Hyg. Assoc. J.* 31:138–159.
- Lippman M. (1977) Regional deposition of particles in the human respiratory tract. In: *Handbook of Physiology, Section 9: Reactions of Environmental Agents* (Lee DHK, Falk LK and Murphy SD, Eds.). American Physiology Society, Bethesda, MD, pp. 213–232.
- Lippman M, Yeates DB, Albert RE. (1980) Deposition, retention and clearance of inhaled particles. *Br. J. Ind. Med.* 37:337–362.
- Lourenco RV, Cotromanes E. (1982) Clinical aerosols. II. Therapeutic aerosols. *Arch. Intern. Med.* 142:2299–2308.
- MacFarland HN. (1976) Respiratory toxicology. In: *Essays in Toxicology* (Hayes WJ, Ed.), Vol. 7. Academic Press, New York, pp. 121–154.
- Marple VA, Rubow KL. (1980) Aerosol generation concepts and parameters. In: *Generation of Aerosols and Facilities for Exposure Experiments* (Willeke K, Ed.). Ann Arbor Science, Ann Arbor, MI, p. 6.
- McAnulty PA, Dayan AD, Ganderup N-C, Hastings KL. (Eds.) (2012) *The Minipig in Biomedical Research*. CRC Press, Boca Raton, FL.
- McClellan RO, Henderson RF. (1989) *Concepts in Inhalation Toxicology*. Hemisphere, Washington, DC.
- McFadden ER Jr. (1986) *Inhaled Aerosol Bronchodilators*. Williams & Wilkins, Baltimore, MD, pp. 40–41.
- Menon MPS, Das AK. (1977) Tetracycline asthma—a case report. *Clin. Allergy.* 7:285–290.
- Mercer TT. (1981) Production of therapeutic aerosols; principles and techniques. *Chest.* 80(Suppl. 6):813–818.
- Merin RG. (1981) Cardiac toxicity of inhalation anesthetics. In: *Cardiac Toxicology* (Balazs T, Ed.), Vol. II. CRC Press, Boca Raton, FL, pp. 4–10.
- Miller FJ, Gardner DE, Graham JA, Lee RE Jr, Wilson WE, Bachmann JD. (1979) Size considerations for establishing a standard for inhalable particles. *J. Air Pollut. Control Assoc.* 29:610–615.
- Miller JL, Stuart BO, Deford HS, Moss OR. (1981) Liquid aerosol generation for inhalation toxicology studies. In: *Inhalation Toxicology and Technology* (Leong BKJ, Ed.). Ann Arbor Science, Ann Arbor, MI, pp. 121–207.
- Moren F. (1981) Pressurized aerosols for oral inhalation. *Int. J. Pharm.* 8:1–10.
- Mussatto DJ, Garrad CS, Lourenco RV. (1988) The effect of inhaled histamine on human tracheal mucus velocity and bronchial mucociliary clearance. *Am. Rev. Respir. Dis.* 138:775–779.
- National Academy of Sciences (NAS). (1958) *Handbook of Respiration*. NAS National Research Council. W.B. Saunders, Philadelphia, PA, p. 41.
- Newhouse M, Sanchis J, Bienenstock J. (1976) Lung defense mechanisms. *N. Engl. J. Med.* 295:1045–1052.
- Newman SP. (1984) Therapeutic aerosols. In: *Aerosols and the Lung, Clinical and Experimental Aspects* (Clarke SW and Pavia D, Eds.). Butterworth & Company, London, pp. 197–224.
- Newman SP, Pavia D, Clarke SW. (1981a) How should a pressurized beta-adrenergic bronchodilator be inhaled? *Eur. J. Respir. Dis.* 62:3–21.
- Newman SP, Moren F, Pavia D, Corrado O, Clarke SW. (1981b) The effects of changes in metered volume and propellant vapour pressure on the deposition of pressurized inhalation aerosols. *Int. J. Pharm.* 11:337–344.
- Newton PE. (2000) Techniques for evaluating hazards of inhaled products. In: *Product Safety Evaluation Handbook*, 2nd Ed. Marcel Dekker, Inc., New York, pp. 243–298.
- Oberst FW. (1961) Factors affecting inhalation and retention of toxic vapours. In: *Inhaled Particles and Vapours* (Davis CN, Ed.). Pergamon Press, New York, pp. 249–266.
- Oyarzun MJ, Clements JA. (1977) Ventilatory and cholinergic control of pulmonary surfactant in rabbit. *J. Appl. Physiol.* 43:39–45.
- Oyarzun MJ, Clements JA. (1978) Control of lung surfactant by ventilation, adrenergic mediators and prostaglandins in the rabbits. *Am. Rev. Respir. Dis.* 117:879–891.
- Parent RA. (2015) *Comparative Biology of the Normal Lung*, 2nd Ed. Elsevier, San Diego, CA.
- Paterson JW. (1977) Bronchodilators. In: *Asthma* (Clark TJH and Godfrey S, Eds.). Chapman and Hall, London, pp. 251–271.
- Paterson JW, Woocock AJ, Shenfield GM. (1979) Bronchodilator drugs. *Am. Rev. Respir. Dis.* 120:1149–1188.

- Patterson R, Kelly JF. (1974) Animal models of the asthmatic state. *Annu. Rev. Med.* 25:53–68.
- Patty FA. (1958) *Industrial Hygiene and Toxicology*, 2nd Ed. Interscience, New York.
- Paustenbach DJ, Carlson GP, Christian JE, Born GS, Rausch JE. (1983) A dynamic closed-loop recirculating inhalation chamber for conducting pharmacokinetic and short-term toxicity studies. *Fundam. Appl. Toxicol.* 3:528–532.
- Pavia D. (1984) Lung mucociliary clearance. In: *Aerosols and the Lung, Clinical and Experimental Aspects* (Clarke SW and Pavia D, Eds.). Butterworth & Co, London, pp. 127–155.
- Pavia D, Bateman JRM, Sheahan NF, Agnew JE, Newman SP, Clarke SW. (1980) Techniques for measuring lung mucociliary clearance. *Eur. J. Respir. Dis.* 67(Suppl. 110):157–177.
- Pavia D, Sutton PP, Agnew JE, Lopez-Vidriero MT, Newman SP, Clarke SW. (1983a) Measurement of bronchial mucociliary clearance. *Eur. J. Respir. Dis.* 64(Suppl. 127):41–56.
- Pavia D, Sutton PP, Lopez-Vidriero MT, Agnew JE, Clarke SW. (1983b) Drug effects on mucociliary function. *Eur. J. Respir. Dis.* 64(Suppl. 128):304–317.
- Phalen RF. (2009) *Inhalation Studies: Foundations and Techniques*, 2nd Ed. Ed. Informa Press, Boca Raton, FL, pp. 35–46 51–57.
- Raabe OG. (1982) Deposition and clearance of inhaled aerosols. In: *Mechanisms in Respiratory Toxicology* (Witschi H and Nettekheim P, Eds.), Vol. 1. CRC Press, Boca Raton, FL.
- Rinehart WE, Hatch T. (1964) Concentration–time product (Ct) as an expression of dose in sublethal exposures to phosgene. *Am. Ind. Hyg. Assoc. J.* 25:545–553.
- Rylander R. (1966) Current techniques to measure alterations in the ciliary activity of intact respiratory epithelium. *Am. Rev. Respir. Dis.* 93(Suppl.):67–85.
- Sackner MA. (1978) Effect of respiratory drugs on mucociliary clearance. *Chest.* 73(Suppl. 6):958–966.
- Sackner MA, Landa J, Hirsch J, Zapata A. (1975) Pulmonary effects of oxygen breathing. A 6-hour study in normal men. *Ann. Intern. Med.* 82:40–43.
- Sackner MA, Brown LK, Kim CS. (1981) Basis of an improved metered aerosol delivery system. *Chest.* 80(Suppl. 6):915–918.
- Salem H. (1986) Principles of inhalation of inhalation toxicology. In: *Inhalation Toxicology: Research Methods, Applications, and Evaluation*. Marcel Dekker, New York, pp. 1–33.
- Salem H, Katz SA. (Eds.) (2015) *Inhalation Toxicology*, 3rd Ed. CRC Press, Boca Raton, FL.
- Schaper M, Detwiler K, Alarie Y. (1989) Alteration of respiratory cycle timing by propranolol. *Toxicol. Appl. Pharmacol.* 97:538–547.
- Schreck RW, Sekuterski JJ, Gross KB. (1986) *Synchronized intra-tracheal aerosol generation for rodent studies*. Aerosols, Formation and Reactivity: Second International Aerosol Conference, September 22–26, 1986, West Berlin. Pergamon Press, New York, pp. 37–51.
- Schwettmann RS, Casterline CL. (1976) Delayed asthmatic response following occupational exposure to enflurane. *Anesthesiology.* 44:166–169.
- Sidorenko GI, Pinigin MA. (1976) Concentration-time relationship for various regimens of inhalation of organic compounds. *Environ. Health Perspect.* 13:17–21.
- Silver SD. (1946) Constant flow gassing chambers: principles influencing design and operation. *J. Lab. Clin. Med.* 31:1153–1161.
- Smith DM, Ortiz LW, Archuleta RF, Spalding JF, Tillery MI, Ettinger HJ, Thomas RG. (1981) A method for chronic nose-only exposure of laboratory animals to inhaled fibrous aerosols. In: *Inhalation Toxicology and Technology* (Leong BKJ, Ed.). Ann Arbor Science, Ann Arbor, MI, pp. 89–105.
- Stahlhofen W, Gebhart J, Heyder J. (1980) Experimental determination of the regional deposition of aerosol particles in the human respiratory tract. *Am. Ind. Hyg. Assoc. J.* 41:385–398.
- Stahlhofen W, Gebhart J, Heyder J. (1981) Biological variability of regional deposition of aerosol particles in the human respiratory tract. *Am. Ind. Hyg. Assoc. J.* 42:348–352.
- Stott WT, McKenna MJ. (1984) The comparative absorption and excretion of chemical vapors by the upper, lower and intact respiratory tract of rats. *Fundam. Appl. Toxicol.* 4:594–602.
- Swift DB, Proctor DF. (1982) Human respiratory deposition of particles during oronasal breathing. *Atmos. Environ.* 16:2279–2282.
- Task Group on Lung Dynamics. (1966) Deposition and retention models for internal dosimetry of the human respiratory tract. *Health Phys.* 12:173–207.
- Taulbee DB, Yu CP, Heyder J. (1978) Aerosol transport in the human lung from analysis of single breaths. *J. Appl. Physiol.* 44:803–812.
- Thorne PS, Karol MH. (1989) Association of fever with late-onset pulmonary hypersensitivity responses in the guinea pig. *Toxicol. Appl. Pharmacol.* 100:247–258.
- Toutou E, Barry BW. (Eds.) (2007) *Enhancement of Drug Delivery*. CRC Press, Boca Raton, FL.
- Ts'o TOT, Leong BKJ, Chenoweth MB. (1975) Utilities and limitations of behavioral techniques in industrial toxicology. In: *Behavioral Toxicology* (Weiss B and Laties VG, Eds.). Plenum Press, New York, pp. 265–291.
- Uemitsu N, Minobe Y, Nakayoshi H. (1985) Concentration-time-response relationship under conditions of single inhalation of carbon tetrachloride. *Toxicol. Appl. Pharmacol.* 77:260–266.
- Ulrich CE, Klonne DR, Church SV. (1984) Automated exposure system for metered-dose aerosol pharmaceuticals. *Toxicologist.* 4:48.
- Vidgren MT, Karkkainen A, Paronen TP, Karjalainen P. (1987) Respiratory tract deposition of ⁹⁹Tc-labeled drug particles administered via a dry powder inhaler. *Int. J. Pharm.* 39:101–105.
- Wanner A. (1979) The role of mucociliary dysfunction in bronchial asthma. *Am. J. Med.* 67:477–485.
- Wanner A. (1981) Alteration of tracheal mucociliary transport in airway disease. Effect of pharmacologic agents. *Chest.* 80(Suppl. 6):867–870.
- Webber SE, Widdicombe JG. (1987) The actions of methacholine, phenylephrine, salbutamol and histamine on mucus secretion from the ferret in-vitro trachea. *Agents Actions.* 22:82–85.

- Weibel ER. (1963) *Morphometry of the Human Lung*. Springer-Verlag, Heidelberg.
- Weibel ER. (1983) How does lung structure affect gas exchange? *Chest*. 83:657–665.
- Weiner N. (1984) Norepinephrine, epinephrine, and the sympathomimetic amines. In: *Goodman and Gilman's, The Pharmacological Basis of Therapeutics* (Gilman AG, Goodman LS, Rall TW and Murad F, Eds.), 7th Ed. Macmillan, New York, pp. 145–180.
- Weston R, Karel L. (1946) An application of the dosimetric method for biologically assaying inhaled substances. *J. Pharmacol. Exp. Ther.* 88:195–207.
- Widdicombe JG. (1974) Reflex control of breathing. In: *Respiratory Physiology* (Widdicombe JG, Ed.), Vol. 2. University Park Press, Baltimore, MD.
- Wilhelm KP. (2012). In: *Dermatotoxicology* (Zhou H and Maibach HI, Eds.), 8th Ed. CRC Press/Informa Healthcare, Chichester.
- Willeke K. (1980) *Generation of Aerosols and Facilities for Exposure Experiments*. Ann Arbor Science, Ann Arbor, MI.
- Williams MH. (1974) Steroids and antibiotic aerosols. *Am. Rev. Respir. Dis.* 110:122–127.
- Wilson AGE. (1982) Toxicokinetics of uptake, accumulation, and metabolism of chemicals by the lung. In: *Mechanisms in Respiratory Toxicology* (Witschi H and Nettesheim P, Eds.), Vol. 1. CRC Press, Boca Raton, FL, pp. 162–178.
- Witschi HP, Brain JD. (1985) *Toxicology of Inhaled Materials*. Springer-Verlag, New York.
- Witschi HP, Nettesheim P. (1982) *Mechanisms in Respiratory Toxicology*, Vol. 1. CRC Press, Boca Raton, FL.

SPECIAL CASE PRODUCTS: IMAGING AGENTS

21.1 INTRODUCTION

Special cases in the assessment of nonclinical safety arise from a variety of situations or for a number of reasons. We have already explored the major cases that arise as a result of different routes of administration (inhalation and topical routes with their variations). These are the most common variations from the general cases of oral and parenteral administration of drugs.

Likewise, the major case associated with different physical forms (small molecules, being the general case, the major variation being biotechnology-sourced materials) has been discussed.

In this chapter, the first of two special cases which arise from the nature of the therapeutic use of these therapeutics is considered. These are imaging agents (*in vivo* diagnostic agents), with the second (being oncology drugs) being addressed in Chapter 22.

21.2 IMAGING AGENTS

Medical imaging agents used for the diagnosis or monitoring serve a variety of different modes of activity such as radiography, computed tomography (CT), ultrasonography, magnetic resonance imaging (MRI), and radionuclide imaging (Weissleder et al., 2010; Szabo, 2013; Wang et al., 2015). These imaging agents can be classified into at least two general categories contrast agents and diagnostic radiopharmaceuticals (Van Brocklin, 2008). Treated as drugs and regulated by either CDER (if small molecules) or CDER (if biological immunochemistry-based entities) as the mechanisms of therapeutic use (action) do not involve chemically or metabolically altering or interacting with the body, these

would seem to be devices rather than drugs, but regulation by the drug centers has developed for historical reasons.

Historically, imaging agents (and the subset of radiopharmaceuticals) start as a real exception to the general case in that they do not required an IND before an initial clinical evaluation. Rather, under the Radioactive Drug Research Committee (RDRC—a special type of IRB) process as specified in CFR 21:36.1, single-dose studies may be conducted in one of the few academic medical centers with only RDRC review and with use limited to in a small number of individuals (up to 30) (CDER, 2000). A number of cardiovascular event-related deaths associated with imaging agent use (Health Imaging News, 2008) has led to greater scrutiny of the safety of the agents.

The RDRC is an institutional body that reviews research protocols for scientific and technical merit. An overview of the RDRC program from 1975 to 2004 has been recently reported (Van Brocklin, 2008). As of 2003, there were 84 active RDRCs in the United States.

In order to be generally regarded as safe (GRAS), a radiopharmaceutical is limited in terms of pharmacological and radiation dose. The mass associated with the radiopharmaceutical must “be known not to cause any clinically detectable pharmacological effect in human beings.” As a result of this limitation, FIH studies must be limited to “microdosing” under an RDRC protocol. Typically, RDRCs require published human studies involving the tracer to be evaluated before approving a protocol at near therapeutic dose. The dose limitation requires that the smallest radiation dose needed to obtain meaningful data from the study be administered to the study subject. The maximum allowable single dose to the whole body, blood-forming organs, lens of the eye, and gonads is 30 mSv (3 rem) with a maximum annual or total dose to all other organs of 50 mSv (5 rem).

The maximum single dose and total annual dose to all other organs are 50 mSv (5 rem) and 150 mSv (15 rem), respectively. There is also a significant radiation dose limit on studies involving research subjects that are less than 18 years of age. The dose may not exceed 10% of the adult doses reported previously. Additionally, all radiation doses associated with the study must be included in the total study dose. This means that the CT dose from a positron emission tomography (PET)/CT study must be included in the total, and this total may not exceed the maximum limits set forth in the regulations.

The types of studies that may be conducted under an RDRC-approved protocol are also regulated. The research must be basic in nature and may include the evaluation of the radiopharmaceutical PK, metabolism, and excretion. The distribution of a radiopharmaceutical to evaluate human physiology, pathophysiology, or biochemistry is permitted as long as the studies are not for diagnostic or therapeutic benefit. Safety and efficacy studies are not permitted under these regulations. An example of a study that is permissible under RDRC would be the brain distribution of [^{18}F]fluorodopa relative to subject age or neurodegenerative disease.

The RDRC may not approve protocols that require more than 30 subjects. If more than 30 subjects need to be studied and may be justified by the researcher, then a special summary form is submitted to the FDA for review and a formal IND must be opened. A pediatric consultant to the RDRC must review studies involving minors under 18 years old and a special summary must be submitted to the FDA. In addition, all adverse reactions “attributable to the use of the radioactive drug” must be reported immediately to the FDA. It is interesting to note, however, that over 30 years since the inception of the RDRC regulations with an estimated 60 000 subjects enrolled in the studies not one adverse event has been reported.

The regulations also stipulate the constitution of the RDRC with appropriate expertise to review the protocol applications. The committee must have at least five members. Three of those members must be a nuclear medicine physician, a qualified individual with radiopharmaceutical preparation experience, and a radiation dosimetry/radiation safety expert. The remaining members must have experience and qualifications in disciplines related to nuclear medicine.

21.2.1 Contrast Agents

As used in the guidance, a contrast agent is a medical imaging agent used to improve the visualization of tissues, organs, and physiologic processes by increasing the relative difference of imaging signal intensities in adjacent regions of the body. Types of contrast agents include, but are not limited to, (i) iodinated compounds used in radiography and CT; (ii) paramagnetic metallic ions (such as ions of gadolinium, iron, and manganese) linked to a variety of

molecules and microparticles (such as superparamagnetic iron oxide) used in MRI; and (iii) microbubbles, microaerosomes, and related microparticles used in diagnostic ultrasonography.

21.2.2 Diagnostic Radiopharmaceuticals

As used in the guidance, a diagnostic radiopharmaceutical is (i) an article that is intended for use in the diagnosis or monitoring of a disease or a manifestation of a disease in humans and that exhibits spontaneous disintegration of unstable nuclei with the emission of nuclear particles or photons or (ii) any nonradioactive reagent kit or nuclide generator that is intended to be used in the preparation of such an article. As stated in the preamble to FDA’s proposed rules on regulations for *in vivo* radiopharmaceuticals used for diagnosis and monitoring, the agency interprets this definition to include articles that exhibit spontaneous disintegration leading to the reconstruction of unstable nuclei and subsequent emission of nuclear particles or photons (63 FR 28301 at 28303; May 22, 1998).

Diagnostic radiopharmaceuticals are generally radioactive drug or biological products that contain a radionuclide that typically is linked to a ligand or carrier. These products are used in nuclear medicine procedures, including planar imaging, single-photon emission computed tomography (SPECT), PET, or in combination with other radiation detection probes.

Diagnostic radiopharmaceuticals used for imaging typically have two distinct components:

1. A radionuclide that can be detect *in vivo* (e.g., technetium-99m, iodine-123, indium-111).
 - The radionuclide typically is a radioactive atom with a relatively short physical half-life that emits radioactive decay photons having sufficient energy to penetrate the tissue mass of the patient. The photons can then be detected with imaging devices or other detectors.
2. A nonradioactive component to which the radionuclide is bound that delivers that radionuclide to specific areas within the body.
 - This nonradionuclidic portion of the diagnostic radiopharmaceutical often is an organic molecule such as a carbohydrate, lipid, nucleic acid, peptide, small protein, or antibody.

As technology advances, new products may emerge that do not fit into these traditional categories (e.g., agents for optical imaging, magnetic resonance spectroscopy, combined contrast, and functional imaging). It is anticipated that the general principles discussed here could apply to these new diagnostic products. Developers of these products should contact the appropriate reviewing FDA division for advice on product development.

21.2.3 Medical Imaging Agent Characteristics Relevant to Safety

The following sections discuss the special characteristics of a medical imaging agent that can lead to a more focused safety evaluation. Characteristics include its radiation-absorbed dose, mass dose, route of administration, frequency of use, biodistribution, and biological, physical, and effective half-lives in the serum, the whole body, and critical organs.

21.2.3.1 Mass Dose Some medical imaging agents can be administered at low mass doses. For example, the mass dose of a single administration of a diagnostic radiopharmaceutical can be small because device technologies can typically detect relatively small amounts of radionuclide (e.g., radiopharmaceuticals for myocardial perfusion imaging). When a medical imaging agent is administered at a mass dose that is at the low end of the dose response curve, safety concerns are minimal.

21.2.3.2 Route of Administration Some medical imaging agents are administered by routes that decrease the likelihood of systemic adverse events. For example, medical imaging agents that are administered as contrast media for radiographic examination of the gastrointestinal tract (e.g., barium sulfate) can be administered orally, through an oral tube, or rectally. In patients with normal gastrointestinal tracts, many of these products are not absorbed, so systemic adverse events are less likely to occur. In general, nonradiolabeled contrast agents pose safety issues similar to therapeutic drugs generally should be treated with therapeutic agents for the purpose of conducting clinical safety assessments.

21.2.3.3 Frequency of Use Many medical imaging agents, including both contrast and diagnostic radiopharmaceuticals, are administered infrequently or as single doses. Accordingly, adverse events that are related to long-term use or to accumulation are less likely to occur with these agents than with agents that are administered repeatedly to the same patient. Therefore, the nonclinical development programs for such single-use products usually can omit long-term (i.e., 3 months duration or longer), repeat-dose safety studies. In clinical settings where it is possible that the medical imaging agent will be administered to a single patient repeatedly (e.g., to monitor disease progression), we recommend that repeat-dose studies (of 14–28 days duration) be performed to assess safety.

Biological medical imaging agents are frequently immunogenic, and the development of antibodies after intermittent, repeated administration can alter the pharmacokinetics, biodistribution, safety, and/or imaging properties of such agents and, potentially, of immunologically related agents. We recommend that studies in which repeat dosing of a biological imaging agent is planned to incorporate pharma-

cokinetic data, human antimouse antibody (HAMA), human antihumanized antibody (HABA), or human antichimeric antibody (HACA) levels as well as whole body biodistribution imaging to assess for alterations in the biodistribution of the imaging agent following repeat dosing. Studies of immunogenicity in animal models are generally of little value. Therefore, we recommend that human clinical data assessing the repeat use of a biological imaging agent be obtained prior to application to licensure of such an agent.

21.2.3.4 Biological, Physical, and Effective Half-Lives Diagnostic radiopharmaceuticals often use radionuclides with short physical half-lives or that are excreted rapidly. The biological, physical, and effective half-lives of diagnostic radiopharmaceuticals are incorporated into radiation dosimetry evaluations that require an understanding of the kinetics of distribution and excretion of the radionuclide and its mode of decay. We recommend that biological, physical, and effective half-lives be considered in planning appropriate safety and dosimetry evaluations of diagnostic radiopharmaceuticals.

21.2.4 Performance of Nonclinical Safety Assessments

FDA recommends that the nonclinical development strategy for an agent be based on sound scientific principles, the agent's unique chemistry (including those of its components, metabolites, and impurities), and the agent's intended use. Because each product is unique, we encourage sponsors to consult with us before submitting an IND application and during product development. The number and types of nonclinical studies recommended would depend in part on the phase of development, what is known about the agent or its pharmacologic class, its proposed use, and the indicated patient population. If it is determined that nonclinical pharmacology or toxicology studies are not needed, FDA is prepared to grant a waiver under 21 CFR 312.0 if adequate justification is provided.

In the discussion that follows, a distinction is made between drug products and biological products. Existing specific guidance for biological products is referenced but not repeated here (see Section III.B.2 of CDER, 2004).

21.2.4.1 Nonclinical Safety Assessments for Nonbiological Drug Products (CDER, 2004)

- Timing of Nonclinical Studies Submitted to an IND Application
 - FDA recommends that nonclinical studies be timed so that they help facilitate the timely conduct of clinical trials (including appropriate safety monitoring based on findings in nonclinical studies) and reduce the unnecessary use of animals and other resources. The recommended timing of nonclinical studies for medical imaging drugs is summarized in

TABLE 21.1 Timing of Nonclinical Studies for Nonbiological Imaging Agents Submitted to an IND

| Study Type | Before Phase I | Before Phase II | Before Phase III | Before NDA |
|--|--|-----------------------------------|---|------------|
| Safety pharmacology | Major organs, ^a and organ systems the drug is intended to visualize | | | |
| Toxicokinetic/ pharmacokinetic | As part of repeat-dose GLP study in P2 | | | |
| Expanded single-dose toxicity | Expanded acute single-dose ^b two species | | | |
| Short-term (2–4 weeks) multiple-dose toxicity | | Repeat-dose toxicity ^c | | |
| Special toxicology | Conduct as necessary based on route irritancy, blood compatibility, protein flocculation, misadministration, and extravasation | | | |
| Radiation dosimetry | If applicable | | | |
| Genotoxicity | <i>In vitro</i> ^d | | | |
| Immunotoxicity | | | May be needed based on molecular structure, biodistribution pattern, class concern, or clinical or nonclinical signal | |
| Reproductive and developmental toxicity | | | Needed or waiver obtained ^d | |
| Drug interaction | | | | As needed |
| Other based on data results | | | | As needed |

^aSee the guidances *S7A Safety Pharmacology Studies for Human Pharmaceutical* and *S7B Safety Pharmacology Studies for Assessing the Potential for Delayed Ventricular Repolarization (QT Interval Prolongation) by Human Pharmaceuticals* (note that S7B allows for phase evaluation of the required studies).

^bSee the guidance *Single-Dose Acute Testing for Pharmaceuticals*.

^cWhen repeat-dose toxicity studies have been performed, but single-dose toxicology studies have not, dose selection for initial human studies will likely be based on the results of the no-observed-adverse-effect level (NOAEL) obtained in the repeat-dose study. The likely result will be a mass dose selection for initial human administration that is lower than if the dose selection had been based on the results of acute, single-dose toxicity studies.

^dSee radiopharmaceutical discussion in CDER (2000).

Table 21.1. The allograft for converting nonclinical animal doses to human dose equivalents is presented in Table 21.2.

- Contrast Agents
 - Because of the characteristics of contrast drug products (e.g., variable biologic half-life) and the way they are used, FDA recommends that nonclinical safety evaluations of such drug products be made more efficient with the following modifications:
 - Long-term (i.e., >3 months), repeat-dose toxicity studies in animals usually can be omitted. (Exceptions are products with long residence time, e.g., >90 days.)
 - Long-term rodent carcinogenicity studies usually can be omitted.
 - Reproductive toxicology studies required under Section 312.23(a)(8)(ii)(a) often can be limited to an evaluation of embryonic and fetal toxicities in rats and rabbits and to evaluations of reproductive organs in other short-term toxicity studies. If you determine that such reproductive studies are not needed, we are prepared to grant a waiver under Section 312.10 if you provide adequate justification.

TABLE 21.2 Factors Used for Milligram per Kilogram to Milligram per Square Meter Conversions

| Species | Conversion Factor |
|---------------|-------------------|
| Mouse | 3 |
| Rat | 6 |
| Guinea pig | 7.7 |
| Hamster | 4.1 |
| Rabbit | 11.8 |
| Dog | 20 |
| Monkey | 12 |
| Human (60 kg) | 37 |

FDA recommends that studies be conducted to address the effects of large mass dose and volume (especially for iodinated contrast materials administered intravenously); osmolality effects; potential transmetalation of complexes of gadolinium, manganese, or iron (generally MRI drugs); potential effects of tissue or cellular accumulation on organ function (particularly if the drug is intended to image a diseased organ system); and the chemical, physiological, and physical effects of ultrasound microbubble drugs (e.g., coalescence, aggregation, margination, and cavitation).

21.2.4.2 Diagnostic Radiopharmaceuticals (Nonbiological Products) Because of the characteristics of diagnostic radiopharmaceuticals and the way they are used, FDA recommends that nonclinical safety evaluations of these drugs be made more efficient by the following modifications:

- Long-term, repeat-dose toxicity studies in animals typically can be omitted.
- Long-term rodent carcinogenicity studies typically can be omitted.
- Reproductive toxicology studies can be waived when adequate scientific justification is provided.
- Genotoxicity studies should be conducted on the non-radioactive component because the genotoxicity of the nonradioactive component should be identified separately from that of the radionuclide. Genotoxicity studies can be waived if adequate scientific justification is provided.

FDA recommends that special safety considerations for diagnostic radiopharmaceuticals include verification of the mass dose of the radiolabeled and unlabeled moiety; assessment of the mass, toxic potency, and receptor interactions for any unlabeled moiety; assessment of potential pharmacologic or physiologic effects due to molecules that bind with receptors or enzymes; and evaluation of all components in the final formulation for toxicity (e.g., excipients, reducing drugs, stabilizers, antioxidants, chelators, impurities, and residual solvents). It is recommended that the special safety considerations include an analysis of particle size (for products containing particles) and an assessment of instability manifested by aggregation or precipitation. It is also

recommended that an individual component be tested if specific toxicological concerns are identified or if toxicological data for that component are lacking. However, if toxicological studies are performed on the combined components of a radiopharmaceutical and no significant toxicity is found, toxicological studies of individual components are seldom required.

REFERENCES

- CDER (2000) *Biological Products: Developing Medical Imaging Drugs and Biological Products, Guidance for Industry*. Center for Drug Evaluation and Research, Silver Spring, MD.
- CDER (2004) *Developing Medical Imaging Drug and Biological Products Part I: Conducting Safety Assessments, Guidance for Industry*. Center for Drug Evaluation and Research, Silver Spring, MD.
- Health Imaging News (2008) FDA Warns that Deaths Continue with Heart Imaging Agents, June 23, 2008.
- Szabo TL. (2013) *Diagnostic Ultrasound Imaging: Inside Out*. Academic Press, San Diego, CA.
- Van Brocklin HF. (2008) Radiopharmaceuticals for drug development: United States regulatory perspective. *Curr. Radiopharm.* 1:2–6.
- Wang ZJ, Chang T-TA, Slauter R, Porter SR, Srivatsan A, Chen X. (2015) Molecular imaging techniques in laboratory animals. In: *Animal Models in Toxicology* (Gad SC, Ed.), 3rd Ed. CRC Press, Boca Raton, FL.
- Weissleder R, Ross BD, Rehemtulla A, Gambbin SS. (2010) *Molecular Imaging: Principles and Practice*. Peoples Medical Publishing House, Shelton, CT.

SPECIAL CASE PRODUCTS: DRUGS FOR TREATMENT OF CANCER

22.1 INTRODUCTION

While cancer is not the most frequent disease cause of death in the United States and developed world (it is second behind cardiovascular disease, with only ~ 500,000 US deaths a year attributed to it in 2015), it is easily the most feared. Treatment options have progressed from surgery to radiation to chemotherapy to use of targeted biological therapies and combinations of these (Mukherjee, 2010). The best approaches are (as with antibiotics for multiresistance bacterial infections) combinations of these, usually in a staged manner.

The most common NME drugs entering development—fully one-third of all INDs—are for oncology. Many reach approval (see Table 22.1), but oncology drugs also have the highest failure rates in all stages of development (Hay et al., 2014; DiMasi et al., 2015).

Traditional oncology drugs now have a specific FDA and ICH guidance for their nonclinical safety assessment requirements (ICH, 2010), reflecting what have been the accepted (and expected) practices and approaches for the assessment of this largest portion of new drugs undergoing development (fully one-third of all new drugs).

These drugs are intended for use in individuals with a serious and life-threatening disease. Initial clinical trials are conducted in patients who have already failed other forms of therapy, not volunteers. Requirements for entry into initial clinical trials for new cancer agents are on a sliding scale, depending on frequency of drug administration to patients (as shown in Table 22.2). For an oncology drug with only one administration a month, a single-dose study in each of two species would be sufficient. Note, however, that data (not wishful thinking) must be provided to support the projected frequency of administration (especially if said administration is projected to be infrequent). Because of

this, as summarized in Tables 22.3 and 22.4, the requirements are not as strenuous as the general case. Indeed, genotoxicity and safety pharmacology are usually not required (Baguley and Kerr, 2002) (see Table 22.5).

The other unusual feature in such trials, whether the drug is a small or large (protein) molecule, is that clinical dosing regimens are not daily but rather in accordance with a schedule set as much by tradition and clinical operations as by drug pharmacokinetics. While drug pharmacodynamics are initially established in mouse models (though a new trend of evaluation of efficacy in companion animal dogs with cancer is developing), only later in clinical development is an understanding of these sufficient to guide clinical practice.

Administrations are usually once or twice a week on a 3- or 4-week sequence of repeat dosing followed by a period without dosing. It is a normal practice for an initial clinical tolerance trial to be conducted with the dose level being escalated in each 3- or 4-week dosing series (Baguley and Kerr, 2002; Teicher and Andrews, 2004; Hidalgo et al., 2010). The 28-day two species nonclinical safety assessment studies are taken to serve for the entire multiseries set of administrations and allow for the most flexible coverage of changed needs in clinical development.

Dose scaling as we now practice it across the range of pharmaceuticals arose from the practice of oncology. Clinicians in this therapeutic area think of dosing in terms of milligrams per square meter, which leads to the need to perform conversions of the expression of dose in safety studies (milligrams per kilogram) to the dose in the per body surface area (BSA) form. In general, the highest dose administered in nonclinical studies does not set an upper limit on the dose escalation in clinical studies—such doses can be advanced to yet higher doses in the face of a lack of toxic-limited dose (TLD) findings in patient populations.

TABLE 22.1 New Drug and Biologic Agents Approved to Treat Cancer, 2004–2010

| Agents | Year Approved | Original Indication(s) | Orphan Drugs | | Clinical Testing Duration (IND to NDA), years |
|--------------|------------------|---|---|---|---|
| | | | Drug Class (Other Similar Drugs) ^a | Alternative Therapies Available at Time of Approval ^a | |
| Pemetrexed | 2004 | Malignant pleural mesothelioma | Antifolate (methotrexate) | None approved for this indication | 11.2 |
| Azacitidine | 2004 | Myelodysplastic syndrome of certain FAB subtypes | DNA methylation inhibitor (none) | Supportive care, HSCT | 3.5 |
| Clofarabine | 2004 | Pediatric patients with ALL after at least two prior treatments | Purine antagonist (fludarabine, cladribine) | Numerous | 5.3 |
| Nelarabine | 2005 | Adult patients with T-cell lymphoblastic lymphoma and ALL after at least two prior treatments; pediatric patients with same conditions | Purine antagonist (fludarabine, cladribine) | None approved for this indication | 10.9 |
| Sorafenib | 2005 | Advanced renal cell carcinoma | Multiple tyrosine kinase inhibitor (none) | Interleukin 2 | 5.1 |
| Lenalidomide | 2005 | Transfusion-dependent anemia due to low- or intermediate-1- risk myelodysplastic syndromes associated with a 5q deletion cytogenetic abnormality | Immunomodulator (thalidomide) | Azacitidine | 4.9 |
| Decitabine | 2006 | Myelodysplastic syndrome previously treated and untreated, de novo and secondary, of all FAB subtypes | DNA methylation inhibitor (azacitidine) | Azacitidine | 7.3 |
| Dasatinib | 2006 | Chronic-phase, accelerated-phase or myeloid or lymphoid blast-phase chronic myeloid leukemia with resistance or intolerance to prior therapy including imatinib; Philadelphia chromosome-positive ALL with resistance or intolerance to prior therapy | BCR-ABL tyrosine kinase inhibitor (imatinib) | None approved for this indication | 2.7 |
| Vorinostat | 2006 | Cutaneous T-cell lymphoma on or following two systemic therapies | Histone deacetylase inhibitor (none) | None approved for this indication | 6.2 |
| Temsirolimus | 2007 | Advanced renal cell carcinoma | mTOR kinase inhibitor (sirolimus) | Sorafenib, sunitinib, interleukin 2 | 4.9 |
| Nilotinib | 2007 | Philadelphia chromosome-positive chronic myelogenous leukemia in chronic phase and accelerated phase in adults resistant or intolerant to prior therapy that included imatinib | BCR-ABL tyrosine kinase inhibitor (imatinib, dasatinib) | Dasatinib | 2.4 |
| Bendamustine | 2008 | Chronic lymphocytic leukemia | Alkylating agent (carmustine, others) | Chlorambucil, rituximab, fludarabine | 4.2 |
| Pralatrexate | 2009 | Replasted or refractory peripheral T-cell lymphoma | Antifolate (methotrexate) | None approved for this indication | 12.1 |
| Ofatumumab | 2009 | Chronic lymphocytic leukemia refractory to fludarabine and alemtuzumab | CD20 monoclonal antibody (rituximab) | Chlorambucil, bendamustine, cyclophosphamide | 4.7 |
| Romidepsin | 2009 | Cutaneous T-cell lymphoma after ≥1 prior systemic therapy | Histone deacetylase inhibitor (vorinostat) | Bexarotene, vorinostat | 6.6 |

| Nonorphan Drugs | | | | | |
|-----------------|------|---|---|---|------|
| Cetuximab | 2004 | EGFR-expressing, metastatic colorectal carcinoma refractory to irinotecan | Anti-EGFR monoclonal antibody (none) | Oxaliplatin, fluorouracil | 8.7 |
| Bevacizumab | 2004 | First-line treatment of metastatic carcinoma of the colon or rectum | Anti-VEGF monoclonal antibody (none) | Oxaliplatin, fluorouracil, irinotecan, capecitabine | 6.7 |
| Erlotinib | 2004 | Locally advanced or metastatic nonsmall cell lung cancer after failure of ≥ 1 prior chemotherapy regimen | HER1/EGFR tyrosine kinase inhibitor (gefitinib) | Vinorelbine, gemcitabine, paclitaxel, others | 7.0 |
| Panitumumab | 2006 | EGFR-expressing, metastatic colorectal carcinoma with disease progression on or following fluoropyrimidine, oxaliplatin, and irinotecan | Anti-EGFR monoclonal antibody (cetuximab) | Bevacizumab, cetuximab | 6.8 |
| Sunitinib | 2006 | GIST after disease progression on or intolerance to imatinib; advanced renal cell carcinoma | Multiple tyrosine kinase inhibitor (sorafenib) | GIST: none approved for this indication; advanced renal cell cancer: interleukin 2, sorafenib | 4.3 |
| Lapatinib | 2007 | Advanced or metastatic breast cancer whose tumors overexpress HER2 and prior to therapy including an anthracycline, a taxane, and trastuzumab | HER2/EGFR tyrosine kinase inhibitor (none) | Palliative chemotherapy | 5.7 |
| Ixabepilone | 2007 | Metastatic or locally advanced breast cancer as combination therapy with capecitabine after failure of an anthracycline and a taxane or as individual therapy after failure of an anthracycline, a taxane, and capecitabine | Epothilone microtubule inhibitor (none) | Combination therapy: gemcitabine; monotherapy: none approved for this indication | 7.7 |
| Degarelix | 2008 | Advanced prostate cancer | GnRH receptor inhibitor (abarelix) | GnRH agonists | 6.6 |
| Pazopanib | 2009 | Advanced renal cell carcinoma | Multiple tyrosine kinase inhibitor (sorafenib, sunitinib) | Temsirolimus, bevacizumab, interleukin 2 | 6.3 |
| Everolimus | 2009 | Advanced renal cell carcinoma after failure of sunitinib or sorafenib | mTOR kinase inhibitor (sirolimus, temsirolimus) | Temsirolimus, bevacizumab, interleukin 2 | 11.5 |
| Eribulin | 2010 | Metastatic breast cancer after ≥ 2 chemotherapeutic regimens | Halichondrin microtubule inhibitor (none) | Lapatinib, capecitabine, ixabepilone | 7.0 |
| Cabazitaxel | 2010 | Hormone-refractory metastatic prostate cancer previously treated with docetaxel | Taxane (paclitaxel, docetaxel) | None approved for this indication (mitoxantrone and prednisone indicated for palliation) | 11.0 |

Source: Kesselheim et al. (2011).

ALL, acute lymphoblastic leukemia; EGFR, epidermal growth factor receptor; GIST, gastrointestinal stromal tumor; GnRH, gonadotropin-releasing hormone; HER, human epidermal growth factor; HSCT, hematopoietic stem cell transplantation; IND, investigational new drug application; mTOR, mammalian target of rapamycin; NDA, new drug application; VEGF, vascular endothelial growth factor. ^aDrug class and available alternative therapies based primarily on summary in US Food and Drug Administration clinical review document.

TABLE 22.2 Examples of Treatment Schedules for Anticancer Pharmaceuticals to Support Initial Clinical Trials

| Clinical Schedule | Examples of Nonclinical Treatment Schedule ^{a-e} |
|---------------------------------------|---|
| Once every 3–4 weeks | Single dose |
| Daily for 5 days every 3 weeks | Daily for 5 days |
| Daily for 5–7 days, alternating weeks | Daily for 5–7 days, alternating weeks (2-dose cycles) |
| Once a week for 3 weeks, 1 week off | Once a week for 3 weeks |
| Two or three times a week | Two or three times a week for 4 weeks |
| Daily | Daily for 4 weeks |
| Weekly | Once a week for 4–5 doses |

^aTable 22.1 describes the dosing phase. The timing of the toxicity assessment(s) in the nonclinical studies should be scientifically justified based on the anticipated toxicity profile and the clinical schedule. For example, both a sacrifice shortly after the dosing phase to examine early toxicity and a later sacrifice to examine late onset of toxicity should be considered.

^bThe treatment schedules described in the table do not specify recovery periods.

^cThe treatment schedules described in this table should be modified as appropriate for molecules with extended pharmacodynamics effects, long half-lives, or potential for anaphylactic reactions. In addition, the potential effects of immunogenicity should be considered (see ICH, 2011).

^dFor nonrodent studies, dose groups usually consist of at least three animals per sex per group, with an additional two per sex per group for recovery, if appropriate. Both sexes should generally be used, or justification should be given for specific omissions.

^eA common approach for many small molecules is to set a start dose at 1/10 the severely toxic dose in 10% of the animals (STD₁₀) in rodents. If the nonrodent is the most appropriate species, then 1/6 the highest nonseverely toxic dose (HNSTD) is considered an appropriate starting dose. The HNSTD is defined as the highest dose level that does not produce evidence of lethality, life-threatening toxicities, or irreversible findings.

The toxicity of the traditional cytotoxic or antimetabolite anticancer drugs correlates more closely with BSA than with body weight (BW). Thus, it is a standard practice to administer and compare such drugs on the basis of BSA. To interconvert between units of milligrams per kilogram and milligrams per square meter, it is necessary to know either BSA or BW. Accurate values for BW are easy to obtain. In contrast, BSA is difficult to measure and therefore is usually calculated from a known BW. The relationship between BW and BSA is described by the formula: $BSA(m^2) = b \times [BW/(kg)]^{2/3}$, where b is a species-specific factor derived from empirical data. In practice, the values for b vary between laboratories (e.g., values for b for the mouse have been found to vary from 0.090 to 0.096). Confusion and miscommunication can occur as a result of the use of different species-specific factors. For example, difficulties have arisen during the dose selection process for a novel cytotoxic anticancer agent with a sleep dose–response curve. In this case, if different investigators use b values as varied as 0.099, 0.101, and 0.111, then animal BSA values for a 10 kg dog will likewise vary, depending on the choice of b

TABLE 22.3 Special Case: Oncology Agents (Cytotoxic)

| Test Requirement | Species |
|--|-----------------|
| <i>Initial Clinical Trial/IND Requirements</i> | |
| 1. Acute toxicity in rodents (IV) | R/M |
| 2. Acute toxicity in nonrodents (IV) | D/S/P |
| 3. 7-day DRF toxicity in rodents (IV) | R/M |
| 4. 7-day DRF toxicity in nonrodents (IV) | D/S/P |
| 5. Safety pharmacology: CV-hERG ^a | <i>In vitro</i> |
| 6. Pivotal/repeat dose in rodents (14–28 day IV) | R/M |
| 7. Pivotal/repeat dose in nonrodents (14–28 day IV) | D/P/S |
| 8. CYP induction/inhibition ^a | <i>In vitro</i> |
| 9. Five species microsome metabolic panel ^a | <i>In vitro</i> |
| 10. Develop bioanalytical for three species (man/rodent/nonrodent) | NA |
| <i>To Support Continued Clinical Development</i> | |
| 11. Pivotal/repeat dose in rodents (3/6 months oral) | R/M |
| 12. Pivotal/repeat dose in nonrodents (3/9–12 months oral) | D/P/S |
| <i>Marketing Approval</i> | |
| 13. Unless issues arise, other requirements are generally varied | |

Note: Species: R=rat, M=mouse, D=dog, S=pig, P=primate, B=rabbit; TBD=to be determined. All studies described previously must be performed GLP.

^aRecommended.

TABLE 22.4 Special Case: Oncology Agents (Protein-Targeted Molecules)

| Test Requirement | Species |
|--|-----------------|
| <i>Initial Clinical Trial/IND Requirements</i> | |
| 1. Acute toxicity in rodents (IV) | R/M |
| 2. Acute toxicity in nonrodents (IV) | D/S/P |
| 3. 7-day DRF toxicity in rodents (IV) | R/M |
| 4. 7-day DRF toxicity in nonrodents (IV) | D/S/P |
| 5. Safety pharmacology: CV <i>in vivo</i> | D/P/S |
| 6. Safety pharmacology: FOB/Irwin | R/M |
| 7. Safety pharmacology: respiratory—rodent | R |
| 8. Pivotal/repeat dose in rodents (14–28 day IV) | R/M |
| 9. Pivotal/repeat dose in nonrodents (14–28 day IV) | D/P/S |
| 10. CYP induction/inhibition ^a | <i>In vitro</i> |
| 11. Five species microsome metabolic panel ^a | <i>In vitro</i> |
| 12. Develop bioanalytical for three species (man/rodent/nonrodent) | NA |
| <i>To Support Continued Clinical Development</i> | |
| 13. Developmental toxicity (segment II)—rat and rabbit pilots and rat and rabbit studies | R/B |
| 14. Immunotoxicity | TBD |
| 15. Pivotal/repeat dose in nonrodents (3/6 months oral) | D/P/S |
| <i>To Support Marketing Approval</i> | |
| 16. Reproductive toxicity—segment I | R |
| 17. Reproductive toxicity—segment III | R |

Note: Species: R=rat, M=mouse, D=dog, S=pig, P=primate, B=rabbit; TBD=to be determined. All studies described previously must be performed GLP.

^aRecommended.

TABLE 22.5 Pre-IND Nonclinical Safety

| | |
|--------------------------------|---|
| Systemic toxicity | 2 species, TLD In case of episodic (1/3–4 week) usage, a single dose may be sufficient to evaluate |
| Genotoxicity | Not required |
| Safety pharmacology | Generally only |
| Reproductive/ developmental | Generally not required |
| Carcinogenicity | Not required |
| Scaling | By body surface area allometry |
| Dosing coverage | Day for day within cohort |

(0.926, 0.945, and 1.038 mg, respectively, for a 2 mg m^{-2} dose). The use of a single set of species-specific factors by all investigators is proposed as it would be beneficial in removing confusion from discussions of study designs and results. Such a proposal has an even more immediate benefit in the case of compounds with very steep dose–response curves, where even small variations in dosing can profoundly influence the results obtained and their interpretation.

22.1.1 Dose Conversions: Perspective

Dosages of pharmaceuticals are typically calculated on the basis of BW interspecies comparisons of toxicological effects of drugs, and chemicals are commonly based on exposures normalized with respect to BW or systemic exposure (i.e., AUC and C_{max} values).

22.1.2 The Use of the mg/m^2 Dose Unit

22.1.2.1 Calculations of Drug Dosages for Treatment

Cytotoxic anticancer drugs are typically administered in dose units of milligrams per square meter. However, BSA is difficult to measure. In contrast, values for height and BW are easy to obtain. Two solutions present themselves:

1. In clinic, estimates of BSA can be accomplished from height and weight data using established nomograms.
2. In the toxicology laboratory, BSA can be calculated from animal BW data *using the appropriate formula*.

22.1.2.2 Conversion of mg/kg BW Doses to Units of mg/m^2

Conversions from milligrams per kilogram BW to milligrams per square meter occur often in the design of toxicity studies, in the interpretation of data from a number of different studies, including old studies, and in comparisons of animal

and human data where dose units are not similar. Two solutions are commonly used.

Another unusual feature in oncology is that the maximum utilized clinical dose is not based on an NOAEL but rather the TLD. Patients are typically titrated to overt toxicity, and actual clinical treatment practice doses patients up to a frankly toxic dose.

Somewhat in parallel, what constitutes a dose-limiting adverse effect in a nonclinical study (a TLD) is usually taken to be a frank toxicity and not a simple intolerance (such as sporadic emesis).

Additionally, it is frequently the case that even initial clinical trials will be done with the investigative agent being administered along with another agent and possibly also radiation. In these cases, such initial clinical trials will also require that at least an additional short study, usually 14 days in a rodent (and not necessarily GLP), be done to establish the safety of the actual combined intended clinical treatment.

REFERENCES

- Baguley BC, Kerr DJ. (2002) *Anticancer Drug Development*. Academic Press, San Diego, CA.
- DiMasi JA, Hermann JC, Twyman K, Kondru RK, Stergiopoulos S, Getz KA, Rackoff W. (2015) A tool for predicting regulatory approval after phase II testing of new oncology compounds. *Clin. Pharmacol. Ther.* 98(5):506–513.
- Hay M, Thomas DW, Craighead JL, Economides C, Rosenthal J. (2014) Clinical development success rates for investigational drugs. *Nat. Biotechnol.* 32(1):40–51.
- Hidalgo M, Garrett-Mayer F, Clendeninn FJ, Eckhardt GS. (2010) *Principles of Anticancer Drug Development*. Humana Press, Totawa, NJ.
- ICH (2010) *ICH S9: Nonclinical Evaluation for Anticancer Pharmaceuticals*. International Conference on Harmonisation, Geneva, Switzerland.
- ICH (2011) *S6(R1): Preclinical Safety Evaluation of Biotechnology-Derived Pharmaceuticals*. International Conference on Harmonisation, Geneva, Switzerland.
- Kesselheim AS, Myers JA, Avorn J. (2011) Characteristics of clinical trials to support approval of orphan vs nonorphan drugs for cancer. *JAMA*. 305(22):2320–2326.
- Mukherjee S. (2010) *The Emperor of All Maladies: A Biography of Cancer*. Scribner, New York.
- Teicher BA, Andrews PA. (2004) *Anticancer Drug Development Guide: Preclinical Screening Clinical Trials and Approval*. Humana Press, Totawa, NJ.

PEDIATRIC PRODUCT SAFETY ASSESSMENT (2006 GUIDANCE, INCLUDING JUVENILE TOXICOLOGY)

23.1 INTRODUCTION

Two of the “special” or especially susceptible populations at the top of all lists of concern as to drug toxicity are the young and the old. There remain no guidelines or specific requirements for assessing the special risks of drugs to the elderly (geriatrics), but there are now requirements and guidelines for the nonclinical assessment of risks of drugs to pediatric and juvenile populations (FDA, 2006; EMEA, 2008; JMHLW, 2012; ICH, 2014). These nonclinical studies are the front end of the Pediatric Investigation Program (PIP) required of most new drugs. While existing studies required to evaluate reproductive effects evaluate pre-natal and early post-natal effects, the FDA remains concerned that effects on juveniles, such as growth retardation (FDA, 1998), may not be adequately screened for.

Designing and conducting relevant evaluation programs present a number of difficulties. Start with modeling the continuing process of anatomical, physiological, and metabolic development as neonates progress from birth through childhood and adolescence in humans. These changes are accompanied with continuous major changes on the absorption, distribution, metabolism, and elimination/clearance of both endogenous and exogenous molecules (Table 23.1). The changes cannot be simply extrapolated from adult situations, dosing, pharmacodynamics, or pharmacokinetics (as was attempted in the past), and no single species of animal can serve as the most appropriate model for developing humans for all drugs.

Studies must be designed to support specific classes of drugs and clinical trials, mimicking what is to be done in such trials. Included in considerations must be the human age at the time of the pediatric trial, the duration of dosing relative to the phases of development and growth in pediatric and juvenile animals, what is known about potential target

organs (one must remember that except in very rare cases, juvenile/pediatric trials are conducted after primary adult toxicity studies, providing some broad knowledge as to potential target organs), and the potential differences in toxicodynamic and pharmacodynamic profiles in developing and mature organisms.

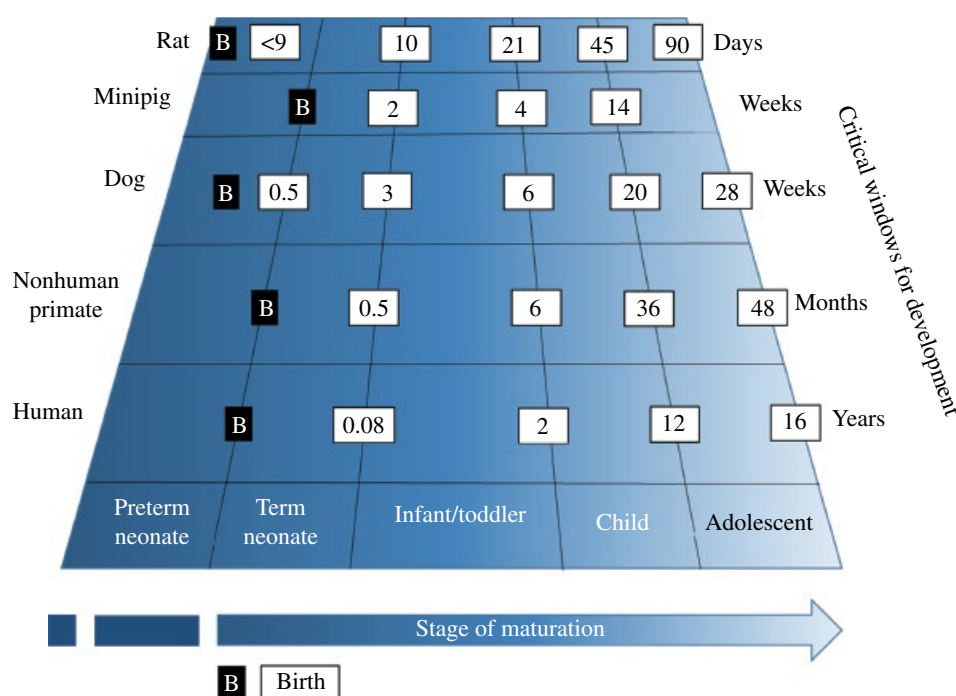
The stage of development in animals being studied must be comparable to those in humans, a comparison summarized in Figure 23.1. The range of utilized model species is more restricted than for general toxicity studies, mirroring available knowledge on developmental specifics.

It is thought that organ systems at highest risk for drug toxicity are those that undergo significant postnatal development. Thus, evaluation of postnatal developmental toxicity is a primary concern. The structural and functional characteristics of many organ systems differ significantly between juveniles and adults as a result of the growth and development that takes place during postnatal maturation, and the same is the case for model species. Examples include the following:

- Brain, where neural development continues through adolescence (Rice and Barone, 2000).
- Kidneys, where adult levels of function are first reached at approximately 1 year of age (Radde, 1985).
- Lungs, where most alveolar maturation occurs in the first 2 years of life (Burri, 1997).
- Immune system, where adult levels of IgG and IgA antibody responses are not achieved until about 5 and 12 years of age, respectively (Miyawaki et al., 1981).
- Reproductive system, where maturation is not completed until adolescence (Zoetis and Walls, 2003).
- Skeletal system, where maturation continues well into adulthood for 25–30 years (Zoetis and Walls, 2003).

TABLE 23.1 Drugs that Exhibit Differences in Toxicity between Adult and Pediatric Patients

| | |
|-------------------------|---|
| Acetaminophen | Acute acetaminophen toxicity is a classic example of how maturation can affect the toxicity profile of a drug. Young children are far less susceptible to acute acetaminophen toxicity than adults because children possess a higher rate of glutathione turnover and more active sulfation. Thus, they have a greater capacity to metabolize and detoxify an overdose of acetaminophen when compared to adults |
| Valproic acid | In contrast to acetaminophen, young children treated with valproic acid appear disproportionately vulnerable to fatal hepatotoxicity |
| Chloramphenicol | Chloramphenicol is associated with mortality in newborns because exposure is increased due to a longer half-life ($t_{1/2} = 26$ h) compared to adults ($t_{1/2} = 4$ h) |
| Inhaled corticosteroids | Inhaled corticosteroids have been found to decrease growth velocity in children, an irrelevant end point in adults (FDA, 1998) |
| Aspirin | Aspirin should not be used to treat children with influenza or varicella infections because of their increased risk of developing Reye's syndrome, a complication not seen in adults |
| Lamotrigine | Children are at a greater risk for developing hypersensitivity-type reactions, including Stevens–Johnson syndrome, when treated with lamotrigine |

**FIGURE 23.1** Comparative stages of development for neonatal and juvenile toxicity studies. *Source:* Beulke-Sam (2003).

- Gastrointestinal systems, which may have direct consequences on bioavailability, clearance, and biotransformation of drugs, are functionally mature by about 1 year of age (Walthall et al., 2005).

Studies in juvenile animals may be useful in the prediction of age-related toxicity in children, as shown in the following examples:

- The effects of phenobarbital on cognitive performance in children were predicted by experimental studies examining the effects of this drug on the developing rodent nervous system (Fonseca et al., 1976; Diaz et al., 1977; Farwell et al., 1990).
- The vulnerability of human neonates to hexachlorophene neurotoxicity was modeled in developing rats and monkeys (Towfighi, 1980).

- The increased susceptibility of infants to verapamil-induced cardiovascular complications would be expected based on animal studies demonstrating a greater sensitivity of the immature heart to calcium channel blockade (Boucek et al., 1984; Skovranek et al., 1986).
- An increased risk of convulsions in young children treated with theophylline was predicted by studies of the preconvulsant effects of this agent in developing rodents (Mares et al., 1994; Yokoyama et al., 1997).

Examples of drug-induced postnatal developmental toxicity demonstrated in animals include the following:

- Neurobehavioral impairment in adult rats following early postnatal exposure to methamphetamine (Vorhees et al., 1994)

- The effects of methylphenidate on growth and endocrine function in young rats (Greeley and Kizer, 1980; Pizzi et al., 1987), particularly given the wide use in children with ADD and frequent polypharmacy in such cases with other drugs such as clonidine.
- Apoptotic neurodegeneration in neonatal rats treated with NMDA receptor antagonists (Ikonomidou et al., 1999)
- Decreased myelination and axonal damage induced in preweaning rats by vigabatrin (Sidhu et al., 1997)
- Long-term changes in serotonergic innervation in rats exposed to fluoxetine during early juvenile life (Wegerer et al., 1999)
- Chondrotoxicity in immature animals treated with fluoroquinolones (Stahlmann et al., 1997)

23.1.1 Scope of Nonclinical Safety Evaluation

The nonclinical safety evaluation of pediatric therapeutics should primarily focus on their potential effects on growth and development that have not been studied or identified in previous nonclinical and clinical studies. Juvenile animal testing may be useful in assessing potential developmental age-specific toxicities and differences in sensitivity between adult and immature animals. Although the toxicological assessment should focus primarily on the active moiety, testing the inactive ingredients in the clinical formulation can also be important, particularly when a drug's pharmacodynamics or distribution is altered by the inactive ingredients or when uncharacterized excipients are present. The toxicological assessment should include local and systemic analyses of effects on postnatal growth and development in the anticipated pediatric population. The known pharmacological and toxicological properties of the drug relative to the proposed patient population (Yaffe and Aranda, 2005) should be considered. Any concerns for postnatal developmental toxicity can be addressed either in juvenile animal studies or by modified study designs (e.g., modification of Segment III reproductive toxicity studies to include animals of similar developmental status as the pediatric population of concern). Juvenile animal studies are especially relevant when a known target organ toxicity occurs in adults in tissues that undergo significant postnatal development. The extent and timing of nonclinical safety studies will depend on the available new pediatric indication for an approved product used in adults that may be quite different from the information needed to support pediatric use of a new molecular entity because of the postnatal developmental safety concerns in the later population. These concerns will be considered for their particular clinical indications on a case-by-case basis within the drug review divisions.

23.1.2 Timing of Juvenile Animal Studies in Relation to Clinical Testing

Specific recommendations regarding the timing of nonclinical toxicology studies are available in the ICH guidance for

industry *M3 Nonclinical Safety Studies for the Conduct of Human Clinical Trials for Pharmaceuticals* (ICH M3 safety studies guidance). The recommendations presented here for juvenile animal studies may assist in identifying postnatal developmental toxicities that are not adequately assessed in general toxicity studies with mature animals and that may not be adequately and safely tested in pediatric clinical trials.

23.1.2.1 Long-Term Exposure in Pediatric Subjects

Most clinical studies in pediatric subjects do not involve long-term exposure to a therapy because they are generally of short-term duration (<6 months). This is especially true when the trials are intended to determine pharmacokinetics rather than efficacy. As a result, long-term exposure during postnatal developmental periods is not usually addressed in pediatric clinical trials. If the drug is indicated for chronic use, then some assessment of the long-term developmental effects of the drug in animals should be made before marketing. However, in those cases when pediatric clinical studies do involve long-term exposure, we recommend conducting juvenile animal studies *before* initiation of the long-term clinical studies. When designing juvenile animal studies, the age of the pediatric population for which the drug is intended is important. Neonates, infants, and older children are at very different developmental stages, and appropriate nonclinical data should support the drug's use in the intended pediatric population.

23.1.2.2 Short-Term Exposure in Pediatric Subjects

Depending on the indication and use of the drug, safety concerns, and the number of subjects exposed, there may be a need for juvenile animal studies in conjunction with clinical studies even if the trials are designed for short-term exposure. Because juvenile animal studies may identify potential hazards and these hazards may have relevance to human safety, it may be more useful to complete juvenile animal studies before conducting clinical studies so that appropriate monitoring can be incorporated into the clinical trial design to limit human risk.

23.1.2.3 Insufficient Clinical Data to Support Initiation of Pediatric Studies

Typically, pediatric subjects are included in clinical trials after there has been considerable experience in the adult population. When there is insufficient clinical data or experience because of minimal prior adult and pediatric experience, completed juvenile animal studies are needed before initiation of pediatric clinical trials regardless of whether the clinical trials involve long-term exposures. Similarly, when there have been reports of adverse effects with off-label use in pediatric patients and there are inadequate data to evaluate the relationship between the drug and the adverse effects, completed juvenile animal studies are needed before initiation of pediatric clinical studies. The timing of juvenile animal studies relative to clinical testing of therapeutics indicated for serious or life-threatening pediatric conditions will be considered on a case-by-case basis by the review division.

23.2 ISSUES TO CONSIDER REGARDING JUVENILE ANIMAL STUDIES

It is important for juvenile animal toxicology studies to be designed efficiently using the least number of animals to identify potential pediatric safety concerns. Whenever feasible, we recommend designing an initial study to address end points of concern for multiple potential pediatric populations. In all cases, studies using juvenile animals are appropriate when adequate information cannot be generated using standard nonclinical studies or from clinical trials. The following issues are specific to studies in juvenile animals for assessing toxicity.

23.2.1 Developmental Stage of Intended Population

Consideration should be given to the age of the intended population and thus the stage of postnatal development. The condition to be treated may also influence the type, extent, and timing of testing considered appropriate. Selection of appropriate end points in the nonclinical studies to address concerns for the specific pediatric populations is important.

Recommendations regarding specific age ranges of pediatric subpopulations are discussed in the ICH guidance for industry *E11 Clinical Investigation of Medicinal Products in the Pediatric Population*.

23.2.2 Evaluating Data to Determine When Juvenile Animal Studies Should Be Used

Evaluation of the available data is important when considering the need for studies in juvenile animals. Toxicity studies in juvenile animals may be appropriate when available nonclinical or clinical data are insufficient to support reasonable safety of a therapeutic for pediatric patients. Gaps in the age ranges of rodent and nonrodent species used in standard toxicity testing are widely acknowledged. These age gaps can affect assessment of nervous system toxicity end points in particular because of the extended process of maturation. Standard toxicity studies with adult animals cannot assess all of the relevant end points, especially growth present in the immature animal. In other circumstances, however, juvenile animal studies might not be necessary when (i) data from similar therapeutics in a class have identified a particular hazard and additional data are unlikely to change this perspective, (ii) there are adequate clinical data and adverse events of concern have not been observed during clinical use, and (iii) target organ toxicity would not be expected to differ in sensitivity between adult and pediatric patients because the target organ of toxicity is functionally mature in the intended pediatric population and younger children with the functionally immature tissue are not expected to receive the drug.

Most drugs that are intended for use in pediatric patients have established efficacy and safety profiles in adult humans. Some data may also be available from pediatric patients aged 12 years or older. For some drugs a preponderance of

clinical data will be obtained from children, as in the case of inhaled corticosteroids (FDA, 1998). For approved drugs that have already undergone extensive clinical testing, substantial nonclinical pharmacology and toxicology data will have already been performed. The toxicology assessment generally includes studies of general toxicity, reproductive toxicity, genetic toxicity, carcinogenicity, and special toxicities, as well as studies in juvenile animals, if available. Target organs of toxicity of the drug both in humans and animals should have been identified in these studies. A thorough evaluation of these data should enable scientists to (i) judge the adequacy of the nonclinical information, (ii) identify some of the potential safety concerns for the intended population, and (iii) identify any gaps in the data that might be addressed by testing in juvenile animals.

23.2.3 Considering Developmental Windows When Determining Duration of Clinical Use

Based on the observation that embryo–fetal development is especially sensitive to perturbation during organogenesis, tissues that undergo significant postnatal development in pediatric patients and juvenile animals may also have greater sensitivity to certain drug-induced toxicities than mature tissues. Organ systems identified as undergoing considerable postnatal growth and development include the nervous, reproductive, pulmonary, renal, skeletal, gastrointestinal, hepatobiliary, and immune systems. Given the variable rate of postnatal development during different periods of childhood, the definition of long-term treatment can vary by pediatric population. Intended treatment of several weeks may not be considered long term in early adolescence but might involve considerable development for the neonate given the duration of some developmental windows.

23.2.4 Timing of Exposure

The timing of the intended use of the drug as it relates to periods of rapid postnatal growth and development is important. If the drug is intended for use in children undergoing phases of rapid overall growth and development, it is important to evaluate an animal model undergoing a corresponding growth phase. Organ systems mature at specific times in specific species. Human-to-animal comparisons of developmental periods for the nervous, reproductive, skeletal, pulmonary, immune, renal, cardiac, and metabolic systems are presented in Section VII of FDA (2006) at the end of this guidance. These can be used as general guides to appropriate periods of treatment to assess the development of specific systems in various animal models. Immature animals have accelerated chronological development compared to humans, which can facilitate evaluation of long-term effects following acute or chronic exposure using well-defined end points (e.g., assessment of reproductive or nervous function).

23.2.5 Selection of Study Models

In addition to consideration of models and end-point assessments based on the intended pediatric human use, target organs for toxicological and pharmacological activity identified in adults need special consideration. It is important that organ systems identified as specific targets of drug toxicity in adults and that undergo significant postnatal development be studied in juvenile animals for those specific effects, even when the primary postnatal developmental period in humans does not coincide with the intended treatment phase. This is based on the observation that development is generally a continuous event. Additionally, a therapeutic target tissue may be developmentally regulated by other tissues or organ systems. In such cases, it may be advisable to examine the effects of the drug during the stages of development relevant to all of those tissues/organs in a test species.

23.3 GENERAL CONSIDERATIONS IN DESIGNING TOXICITY STUDIES IN JUVENILE ANIMALS

Modified repeat-dose toxicity studies can provide a more general screen for potential hazards in some instances.

However, we recommend that such studies modify the animal age at the study initiation, duration of treatment, and end points assessed to address the specific concerns. Modification of standard ICH studies designed to address developmental stages C–F (ICH, 1994) would include ensuring adequate exposure in juvenile animals during the postnatal period and assessment of developmental end points appropriate for the intended pediatric population. Assessment of developmental end points not usually included in standard repeat-dose toxicity studies also may be appropriate. In addition to ensuring adequate exposure to the drug, histopathologic examinations and effects on specific growth parameters and functionally immature tissues in the juvenile animal would be important. In these modified designs, dosing can be initiated with animals younger than usual and extended until the developmental period for the intended pediatric population has been completed in the animal species in accordance with the age of the pediatric patients who would use the drug. Information from such studies can be compared with the findings from treated adults of the same species to evaluate whether the effects are specific to juvenile animals.

If one has knowledge of the toxico-/pharmacodynamics of the compound of interest, there is a wealth of information available on comparative developmental end points (see FDA, 2006), including on enzymatic metabolism (see Table 23.2).

TABLE 23.2 Enzymatic Metabolism Development (Humans, Rats, and Rabbits)

| Enzyme | Developmental Modulation of Phase I/II Metabolism | | |
|--------------------------------|---|--|----------------------------|
| | Maturation of Enzyme Activity | | |
| | Human (Years) | Rat (Days) | Rabbit (Days) |
| CYP2D6 ^{a,b} | 0–3 | NA* | NA* |
| CYP2E1 ^{b,c,d} | 0–1 | 4–17 ↓ postweaning Male > female | 14–35 2× adult At 35 |
| CYP1A2 ^{a,e,f,g} | 0.5 1 (>adult) | 7–100 low levels | 21–60 |
| CYP2C8 ^{a,b} | <1 | NA* | NA* |
| CYP2C9 ^{a,b} | <0.5 0.5 (>adult) | NA* | NA* |
| CYP3A4 ^b | 0–2 | NA* | NA* |
| Acetylation ^{a,b} | 1 (35% adult) | NA* | NA* |
| Methylation ^{a,b} | <1 (50% adult) | NA* | NA* |
| Glucuronidation ^{a,b} | 0 (>adult) 12 | | NA* |
| Sulfation ^{a,b} | 0 | NA* | NA* |

NA*, Data as to point of appearance not available.

^aKearns and Reed (1989).

^bLeeder and Kearns (1997).

^cWaxman et al. (1989).

^dPeng et al. (1991).

^eDing et al. (1992).

^fImaoka et al. (1991).

^gPineau et al. (1991).

23.4 STUDY DESIGNS AND CONSIDERATIONS

As a starting place, it must be noted that one cannot generally initiate dosing of animals immediately at birth—their chances of survival are very adversely affected. Table 23.3 summarizes currently accepted guidance on earliest dosing start dates for different routes according to model species.

Second there are several approaches to assigning animals in separate litters (for rodents) to different test groups (see Table 23.4).

For nonrodents, the complications of small litter (and acceptable) group sizes in primates (Table 23.5), dogs (Table 23.6), and minipigs (Table 23.7) must be considered.

TABLE 23.3 Earliest Starting Day Based on Dosing Routes

| | Rat | Mice | Rabbit | Dog | Minipig |
|-------------------------|--------|--------|--------|--------------------|--------------------|
| Oral gavage | PND 1 | PND 4 | PND 14 | PND 1 | PND 1 |
| Subcutaneous | PND 1 | PND 1 | PND 6 | PND 1 | PND 1 |
| Intramuscular | PND 1 | PND 1 | PND 6 | PND 1 | PND 1 |
| IV bolus (repeated) | PND 4 | PND 7 | PND 6 | PND 1 | PND 7 ^a |
| IV infusion | PND 21 | PND 21 | PND 28 | PND 56 | PND 7 ^a |
| Inhalation (whole body) | PND 4 | PND 4 | PND 6 | PND 10 | PND 2–7 |
| Inhalation (nose only) | PND 21 | PND 21 | PND 28 | PND 4 ^b | PND 4 ^b |
| Dermal ^c | PND 21 | PND 21 | PND 28 | PND 42 | PND 28 |

Source: Hoberman and Lewis (2012) and Parker (2014).

^aDifficult because there are no easily accessible veins (vascular implant port can be surgically installed after PND 7).

^bMasks can be fit as young as PND 4 but restraint issues arise.

^cNot recommended in preweaning animals.

TABLE 23.4 Litter Composition Study Design Options

| | |
|-----------------------------------|---|
| Within-litter design | Each litter has all treatment groups |
| Split litter design | Each litter has some of the treatment groups |
| Between-litter design | Each litter has the same treatment group |
| Fostering design | Each litter is composed of pups from other litters without using any siblings |
| | All pups within new litter receive the same treatment |
| One pup per sex per litter design | Self explanations |

Note: Each method has its advantages and disadvantages (logistically, cross-contamination possibility, IACUC issues, and statistically).

TABLE 23.5 Juvenile Nonhuman Primate Model

| Advantages | Disadvantages |
|---|--|
| Physical size facilitates collection of multiple biologic specimens | Procurement of appropriately aged animals |
| Postnatal development of many organ systems well characterized | Not practical to test full span of postnatal development |
| Potentially less immunogenic than other species | Limited reference or historical control data for some end points |
| Standardized tests available | More expensive (vs. rodent) |
| • Neurobehavioral testing | Limited appropriate technical expertise/experience |
| | Statistical analysis |

TABLE 23.6 Juvenile Dog Model

| Advantages | Disadvantages |
|---|--|
| Provides model of immaturity | Long time to sexual maturity limits assessment of reproductive development |
| Postnatal development of many organ systems well characterized | Limited reference or historical control data for some end points |
| Physical size facilitates collection of multiple biologic specimens | Learning and memory tests not well developed |
| Samples can be collected with minimum restraint and without anesthesia | Potential immunogenic response |
| Techniques for handling and treating puppies well established | Pharmacological relevance may not have been previously characterized |
| Ability to procure appropriate numbers of animals, even for early-age assessments | Statistical analysis |

TABLE 23.7 Juvenile Minipig Model

| Advantages | Disadvantages |
|---|--|
| High pregnancy rate and synchronization of mating possible | Limited pharmacological relevance |
| Gestation length shorter than NHP | Neurobehavioral tests not well developed |
| Similarity of organ system anatomy, size, and physiology to humans | Potential immunogenic response |
| Availability of reference data for juvenile organ development | Limited labs available with experience |
| Piglets are easily accessible for handling and dosing and allow frequent blood sampling | Sexual maturity at 6.5–7.5 months ^a |
| Sexual maturity at 4–5 months | |

^aTortereau et al. (2013).

Generally, the use of a single species is considered sufficient to evaluate juvenile toxicity subsequent to evaluation of adult systemic toxicity.

It is critical that those designing, conducting, and analyzing juvenile toxicity studies be very knowledgeable about the developmental biology of the model species being employed.

REFERENCES

- Beulke-Sam J. (2003) *Comparative schedules of development in rats and humans: implications for developmental neurotoxicity testing*. Presented at the 42nd Annual Meeting of the Society of Toxicology, March 9–13, 2003, Salt Lake City, UT.
- Boucek RJ Jr, Shelton M, Artman M, Mushlin PS, Starnes VA, Olson RD. (1984) Comparative effects of verapamil, nifedipine, and diltiazem on contractile function in the isolated immature and adult rabbit heart. *Pediatr. Res.* 18:948–952.
- Burri P. (1997) Structural aspects of prenatal and postnatal development and growth of the lung. In: *Lung Growth and Development* (McNoald JA, Ed.). Marcel Dekker, New York, pp. 1–35.
- Diaz J, Schain RJ, Bailey BG. (1977) Phenobarbital-induced brain growth retardation in artificially reared rat pups. *Biol. Neonate.* 32:77–82.
- Ding X, Peng HM, Coon MF. (1992) Cytochromes P450 NMa, NMb (2G1) and LM4 (1A2) are differentially expressed during development in rabbit olfactory mucosa and liver. *Mol. Pharmacol.* 42(6):1027–1032.
- EMA (2008) *Guideline on the Need for Nonclinical Testing in Juvenile Animals of Pharmaceuticals for Pediatric Indications*. European Medicines Agency, London.
- Farwell JR, Lee YJ, Hirtz DG, Sulzbacher SI, Ellenberg JH, Nelson KB. (1990) Phenobarbital for febrile seizures—effects on intelligence and on seizure recurrence. *N. Engl. J. Med.* 322: 364–369.
- FDA (2006) *Nonclinical Safety Evaluation of Pediatric Drug Products*. Center for Drug Evaluation and Research (CDER), Silver Spring, MD.
- Fonseca NM, Sell AB, Carlini EA. (1976) Differential behavioral responses of male and female adult rats treated with five psychotropic drugs in the neonatal stage. *Psychopharmacologia.* 46:253–268.
- Greeley GH, Kizer JS. (1980) The effects of chronic methylphenidate treatment on growth and endocrine function in the developing rat. *J. Pharmacol. Exp. Ther.* 215:545–551.
- Hoberman AM, Lewis EM. (Eds.) (2012) *Pediatric Non-clinical Drug Testing: Principles, Requirements and Practices*. John Wiley & Sons, Inc., Hoboken, NJ.
- ICH (1994) *Guidance for Industry: S5A Detection of Toxicity to Reproduction for Medicinal Products*. International Conference on Harmonisation, Geneva, Switzerland.
- ICH (2014) *E11(R1) Clinical Investigation of Medicinal Products in the Pediatric Population*. International Conference on Harmonisation, Geneva, Switzerland.
- Ikonomidou C, Bosch F, Miksa M, Bittigau P, Vockler J, Dikranian K, Tenkova TI, Stefovskaya V, Turski L, Olney JW. (1999) Blockade of NMDA receptors and apoptotic neurodegeneration in the developing brain. *Science.* 283:70–74.
- Imaoka S, Fujita S, Funae Y. (1991) Age dependent expression of cytochrome P450s in rat liver. *Biochim. Biophys. Acta.* 1097(3):187–192.
- JMHLW (2012) *Guideline on Nonclinical Safety Studies in Juvenile Animals for Pediatric Drugs*. Japanese Ministry of Health, Labour and Welfare, Tokyo.
- Kearns LK, Reed MD. (1989) Clinical pharmacokinetics in infants and children. A reappraisal. *Clin. Pharmacokinet.* 17(Suppl. 1): 29–67.
- Leeder JS, Kearns GL. (1997) Pharmacogenetics in pediatrics: implications for practice, new frontiers in pediatric drug therapy. *Pediatr. Clin. North Am.* 44(1):55–77.
- Mares P, Kubova H, Czuczwar SJ. (1994) Aminophylline exhibits convulsant action in rats during ontogenesis. *Brain Dev.* 16(4):296–300.
- Miyawaki T, Moriya N, Nagaoki T, Taniguchi N. (1981) Maturation of B-cell differentiation ability and T-cell regulatory function in infancy and childhood. *Immunol. Rev.* 57:61–87.
- Parker RM. (2014) Juvenile toxicology. In: *Handbook of Toxicology* (Derelanko M and Auletta C, Eds.), 3rd Ed. CRC Press, Boca Raton, FL, pp. 399–452.
- Peng HM, Porter TD, Ding XX, Coon MJ. (1991) Differences in the developmental expression of rabbit cytochromes P450 2E1 and 2E2. *Mol. Pharmacol.* 40(1):58–62.
- Pineau T, Daujat M, Pichard L, Girard F, Angevain J, Bonfils C, Maurel P. (1991) Developmental expression of rabbit cytochrome P450 CYP1A1, CYP1A2, CYP3A6 genes, effect of weaning and rifampicin. *Eur. J. Biochem.* 197(1):145–153.
- Pizzi WJ, Rode EC, Barnhart JE. (1987) Differential effects of methylphenidate on the growth of neonatal and adolescent rats. *Neurotoxicol. Teratol.* 9:107–111.
- Radde IC. (1985) Mechanism of drug absorption and their development. In: *Textbook of Pediatric Clinical Pharmacology*

- (Macleod SM and Radde IC, Eds.). PSG Publishing Co., Littleton, MA, pp. 17–43.
- Rice D, Barone S Jr. (2000) Critical periods of vulnerability for the developing nervous system: evidence from humans and animal models. *Environ. Health Perspect.* 108(Suppl. 3):511–533.
- Sidhu RS, Del Bigio MR, Tuor UI, Sechia SS. (1997) Low-dose vigabatrin (gamma-vinyl GABA)-induced damage in the immature rat brain. *Exp. Neurol.* 144:400–405.
- Skovranek J, Ostadal B, Pelouch V, Prochazka J. (1986) Ontogenetic differences in cardiac sensitivity to verapamil in rats. *Pediatr. Cardiol.* 7:25–29.
- Stahlmann R, Chahoud I, Theil R, Klug S, Forster C. (1997) The developmental toxicity of three antimicrobial agents observed only in nonroutine animal studies. *Reprod. Toxicol.* 11:1–7.
- Tortereau A, Howroyd P, Lorentsen H. (2013) Onset of puberty and normal histological appearances of the reproductive organs in peripubertal female Göttingen minipigs. *Toxicol. Pathol.* 41(8):1116–1125.
- Towfighi J. (1980) Hexachlorophene. In: *Experimental and Clinical Neurotoxicology* (Spencer S, Ed.). Williams & Wilkins, Baltimore, MD, pp. 440–455.
- United States Food and Drug Administration. *Class Labeling for Intranasal and Orally Inhaled Corticosteroid Containing Drug Products Regarding the Potential for Growth Suppression in Children*. FDA Talk Paper 1998.
- Vorhees CV, Ahrens KG, Acuff-Smith KD, Schilling MA, Fisher JE. (1994) Methamphetamine exposure during early postnatal development in rats: I. Acoustic startle augmentation and spatial learning deficits. *Psychopharmacology.* 114:392–401.
- Walthall K, Cappon GD, Hurtt ME, Zoetis T. (2005) Postnatal development of the gastrointestinal system: a species comparison. *Birth Defects Res.* 74(Part B):132–156.
- Waxman DJ, Morrissey JJ, LeBalnc GA. (1989) Female predominant rat hepatic P450 forms (IIE1) and 3(IIA1) are under hormonal regulatory controls distinct from those of sex specific P450 forms. *Endocrinology.* 270(2):458–471.
- Wegerer V, Moll GH, Bagli M, Rothenberger A, Ruther E, Huether G. (1999) Persistently increased density of serotonin transporters in the frontal cortex of rats treated with fluoxetine during early juvenile life. *J. Child Adolesc. Psychopharmacol.* 9:13–24.
- Yaffe SJ, Aranda JV. (2005) *Neonatal and Pediatric Pharmacology*. Lippincott Williams & Wilkins, Philadelphia, PA.
- Yokoyama H, Onodera K, Yagi T, Iinuma K. (1997) Therapeutic doses of theophylline exert proconvulsant effects in developing mice. *Brain Dev.* 19:403–407.
- Zoetis T, Walls I. (2003) *Principles and Practices for Direct Dosing of Pre-Weaning Mammals in Toxicity Testing and Research*. ILSI Press, Washington, DC, p. 11, 13.

USE OF IMAGING, IMAGING AGENTS, AND RADIOPHARMACEUTICALS IN NONCLINICAL TOXICOLOGY

24.1 INTRODUCTION

Rather than split these into three separate chapters, here the aim is to integrate these interrelated subjects into a single cohesive unit, with the goal of a coverage that better relays the power (and challenges) of the use of available imaging technologies and how contrast agents and radiopharmaceuticals fit into this utilization, followed by a review of current regulatory nonclinical safety requirements for the use of those two classes of drugs.

Imaging and radiopharmaceutical techniques for diagnostic use all share the goal of using a growing toolbox of methodology to evaluate drug targeting and effects. In recent years, these methodologies have become available for use in model species in toxicology, from larger nonrodents (dogs, swine, and primates) to rodents (rats and mice) (Kiessling and Pichler, 2011).

The first section seeks to provide an overview of the methods and their potential use. The next two sections will address consideration of how to evaluate the nonclinical safety of actual agents (regulated as drugs).

The available methods for imaging cover a broad range of possibilities:

- X-rays
- Positron emission tomography (PET)
- Computed tomography (CT)
- Magnetic resonance imaging (MRI)
- Optical imaging
- Ultrasound (US)
- Radiopharmaceuticals
- Nanoparticle-assisted imaging

Others are likely to join the list; terahertz imaging (based on wavelengths between infrared and microwaves) is now being pursued for cancer detection (Jacoby, 2015).

The development of therapeutic agents and modalities spans a period of 5–15 years leading to current approaches (and is still ongoing). The most recent technique available for real-time cellular, tissue, and molecular imaging (MI) of organisms and the disposition of drug in living body provides a powerful means to identify whether a drug candidate hits the target and effects tissues in laboratory animals in early stages of drug development, improving the quality of selection of molecules entering development.

Unlike the more purely energy-driven imaging technologies (also discussed here), MI technology provides not only visualization and characterization but also measurement of biological processes at the molecular and cellular levels in humans and other living systems. MI technologies are increasingly recognized as important preclinical and clinical research tools to speed up the long timelines of drug development process (Hargreaves, 2008; Rogge and Taft, 2010; Weissleder et al., 2010; Kiessling and Pichler, 2011), with a focus on utilizing noninvasive techniques to visualize the anatomic structures and physiologic activity in nonclinical and clinical research and in clinical practice. Modern imaging assessments provide information about tissues by penetrating the living body via physical phenomena. These imaging modalities have been reengineered for use with laboratory animals by pushing the resolution and sensitivity of each modality to the physical limit. The improved high resolution and sensitivity make the imaging technology highly translational in the preclinical drug development process.

These technologies are able to provide evidence of *in vivo* biodistribution of imaging probes and agents, confirm

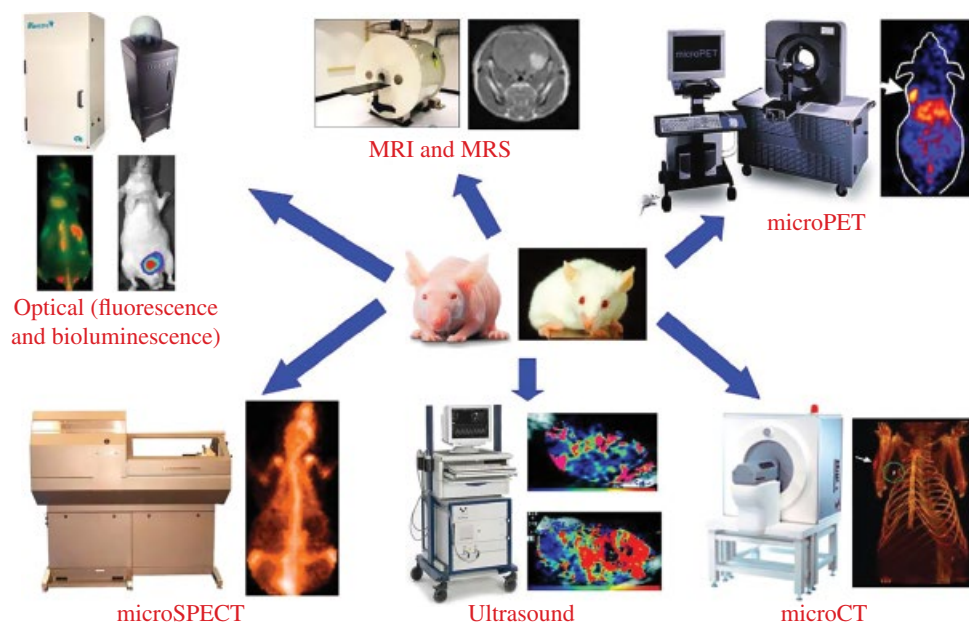


FIGURE 24.1 Multimodality imaging instrumentations. The modern molecular imaging equipment includes magnetic field/radiofrequency (magnetic resonance imaging (MRI)), X-ray (computed tomography (CT)), high-frequency sound waves (ultrasound (US)), optical (bioluminescence/fluorescence), gamma rays (single-photon emission computed tomography (SPECT)), and annihilation twin photons from beta emission (positron emission tomography (PET)). *Source:* Images are courtesy of LOMIN of NIBIB at National Institutes of Health.

on-target biological activity, discern disease mechanism, and potentially provide proof of concept for the therapeutic in question. Several excellent review articles (Bruggen and Roberts, 2003; Willmann et al., 2008) and one book (Weissleder et al., 2010) have been published on the general role of MI on the process of both preclinical and clinical drug development.

24.1.1 Multimodality Imaging Techniques

MRI, SPECT, PET, CT, US, and optical imaging all generate sectional images that can be reconstructed into two-dimensional (2-D) or three-dimensional (3-D) images. For non-clinical evaluations, the imaging technique must have spatial resolution that is adequate for small animals and sensitivity to detect biochemical events. Each technique has certain advantages and limitations, with platform selection being dependent upon which techniques best serve the need of answering a particular nonclinical question. Combining imaging techniques (multimodality) in the same animal, such as using PET/CT, SPECT/CT, and PET/MRI coregistration, can provide temporal pathophysiological information such as development of structural and functional changes, progression and resolution during disease treatment, and the drug-related on-target- and even off-target-related toxicities (Cali et al., 2012).

24.1.1.1 Agents, Probes, and Biomarkers MI probes are agents used to visualize, characterize, and measure biological processes in living systems. MI instrumentation, including

MRI, CT, US, optical imaging, SPECT, and PET as shown in Figure 24.1, are tools that enable the visualization and quantification in space and over time of signals from MI probes. In nonclinical research, diagnostic probes or imaging agents are administered in trace amounts and typically do not induce any physiological response or significant pharmaceutical effect in living systems (Schwaiger et al., 2013).

Imaging probes include radiolabeled contrast agents, biologics, nanoparticles, and different cell types. A radiotracer can be defined as a specific radiolabeled molecule that monitors the *in vivo* behavior of a functional molecule and can be used to provide biological information in a living system. The synthesis of a radiotracer generally includes radiolabeling of the molecule of interest, purification of final product, resting specific activity, and analysis of *in vivo* stability. The commercially available imaging probes and customized radiotracers for PET or SPECT imaging can be used to detect such functional changes as cerebral blood flow, myocardial perfusion, glucose metabolism, infection, bone lesion, kidney function, apoptosis, angiogenesis, thyroid dysfunction, tumor detection, gene expression, neuroreceptor binding, and cell trafficking process (Cai and Chen, 2010; Morone and Betterhauser, 2015).

24.1.2 Dynamic Molecular Imaging Techniques

The radiotracer approach used for PET and SPECT functional imaging is especially valuable at early stages of drug discovery, where researchers can directly label a

therapeutic candidate to see where it localizes in laboratory animals. If the drug candidate does not distribute to its intended physiologic target, as desired, it can be excluded from further development at an early stage. In addition, PET and SPECT imaging experiments can reveal molecular changes in targeted and nontargeted tissues brought about by the drug and its side effect caused within the laboratory animals. This can be extremely important in view of the fact that changes on the molecular level occur long before becoming visible in any anatomic imaging. Such approaches allow the rank ordering of development candidates based on relevant physiologic processes very early in drug development.

In vivo imaging techniques such as PET and SPECT provide a means to perform pharmacokinetic (PK) and pharmacodynamic (PD) studies in animal models without terminating the animal, allowing for serial time samples. With the introduction of dynamic and static imaging approaches, it is now possible to obtain high-resolution PET and SPECT images of “motion-frozen” time points in PK and PD studies and be able to quantify the amount of drug in the target organs of interest in the same laboratory animal over a period of time during the drug evaluation process.

24.2 X-RAY

This is the oldest and most generally familiar technology for visualizing structures *in vivo* and noninvasively. It generally does not require use of an imaging agent, though in some cases.

Instruments (crude ones are easy to build—the author and his brother built one for a science fair project in high school) with electrical energy and generally any radiation in a wavelength ranging from 0.01 to 10 nm by electrons cause a metal anode to release photons. These wavelengths are shorter than UV or visible light and carry more energy. These waves pass through most objects—certainly through living tissues—with varying degrees of absorption. Capturing the “scanned” region onto a suitable receptor (originally film, now electronic detectors) provides an image (“radiograph”) of internal structures.

The difficulty is that X-ray photons carry (and can transmit/transfer) enough energy to ionize atoms and harm the tissues they transit.

Contrast (imaging) agents are used to increase the sharpness and clarity of images in a specific region of interest in the body.

Fluoroscopy is a form of X-ray imaging used to obtain real-time mobile images of structures within the body. It may also be used by surgeons to allow proper placement of devices such as catheters.

24.2.1 Angiography

Angiography, utilizing X-rays to image coronary vasculature and blood flow in discrete regions during a study, has become a standard tool for evaluating effects on vascular

function (including blockage, revascularization, and angiogenesis) when such are concerns or desired outcomes of treatment. The performance of such investigations is usually restricted to dogs and swine, where it is a valuable tool for evaluating vascular/flow functional effects.

24.3 POSITRON EMISSION TOMOGRAPHY (PET)

In PET imaging, a compound (small-molecule drug or biological molecule) labeled with a positron-emitting radioisotope is injected into the subject in a nonpharmacological trace quantity. A positron ejected by a radionuclide combines with an electron in adjacent tissue to emit a pair of photons (511 keV) resulting from annihilation of a positron–electron pair. The PET scanner uses the annihilation coincidence detection (ACD) method to obtain projection images of the localization and quantification of the radiolabeled compound in a living subject (Wang and Maurer, 2005). PET imaging technology can be used for drug distribution, organ perfusion, cell trafficking, tumor targeting, tumor metabolism/proliferation, tumor angiogenesis, tumor hypoxia, tumor apoptosis, tumor volume, anticancer therapeutic response, bone growth/healing, and mechanism studies in animal models of central nervous system (CNS) and autoimmune diseases. Advantage of this technique includes high molecular sensitivity (nanomolar) with almost unlimited depth penetration. Disadvantages of PET include radiation safety and relatively high cost. PET functional imaging is often combined with either micro-CT or animal MRI for better identification of functional change in deep tissue against anatomical structure in an *in vivo* animal model.

24.4 SINGLE-PHOTON EMISSION COMPUTED TOMOGRAPHY (SPECT)

SPECT camera images individual high-energy photons—gamma rays, resulting from radionuclide decay. The technique requires injection of a gamma-emitting radioisotope into laboratory animal for research or patient in clinical practice. Radioisotopes can be coordinated to a test article (ligand) to allow for the evaluation of the binding properties of the test article *in vivo*. The combination of radioisotope and the ligand will bind to a place of interest in the body of the living subject to be seen by the gamma ray detectors, thus producing an image by CT. For example, an imaging agent with affinity for areas of growing bone can be utilized to evaluate potential impact of a test article on the process of ossification.

The single-photon emission radiates from the source in all directions with equal probability. Because of this, an aperture composed of highly attenuation material is required to identify the path of origin of each emission prior to being

able to create an image with a gamma camera. In preclinical SPECT imaging, the most recently developed multipinhole collimator technology achieves submillimeter, high resolution in small laboratory animals.

SPECT can be used to complement any gamma imaging studies. Biodistribution, tumor imaging, infection, bone imaging, and cardiac gated imaging can also be performed with SPECT. SPECT (Wernick and Aarsvold, 2004) offers radiolabeled image resolution and sensitivity that can translate from mouse to human. It also can provide both qualitative and quantitative measurement of physiologic processes. The isotopes used for SPECT imaging (i.e., ^{123}I , ^{125}I , $^{99\text{m}}\text{Tc}$, ^{111}In) are readily available and relatively inexpensive, with longer half-lives than those commonly used in PET imaging. The extended half-life of these isotopes translates to longer scan times and/or scanning periods without the need for secondary isotope administration. These characteristics of SPECT imaging results in a reduction in the isotope quantity needed, as well as the total number of animals utilized in preclinical studies—more data using fewer animals over a longer investigational period. SPECT imaging/isotopes also provide the ability to simultaneously image multiple radiopharmaceuticals with different energies (i.e., $^{99\text{m}}\text{Tc}$ vs. ^{125}I). Micro-SPECT functional imaging is able to coregister with a high-quality CT imaging framework for anatomical imaging comparison.

One characteristic of SPECT imaging is the infrastructure and instrumentation cost, as well as the need for radioactive material handling and dosimetry. This platform requires the investigators' multidisciplinary knowledge and experience to ensure the study design provides an adequate data set to meet their objectives. Although the molecular sensitivity of SPECT is almost one order of magnitude lower in sensitivity than PET, SPECT imaging continues to be widely used in both clinical practice and preclinical research due to its advantages of a relatively lower operating cost, dual-labeled compounds, and a valuable long half-life radionuclides, which allow for lengthier monitoring periods of *in vivo* biological process compared to PET imaging.

24.5 COMPUTED TOMOGRAPHY (CT)

CT is an application of X-ray imaging that provides a 3-D anatomical image. In CT scan, X-rays are emitted from an X-ray source rotating around the subject placed in the center of the CT scanner. A detector opposite the X-ray source senses X-rays that are not absorbed by the tissue; this absorption is inversely related to the density of the tissue structures. The X-ray absorption profile is then used to reconstruct high-resolution (roughly 6–50 μm with no depth limit) tomographic anatomical images. Micro-CT can be used for bone studies, developmental and reproductive toxicology (DART) studies, fetal skeletal evaluations

vascular studies, and lung studies. The use of contrast media (see Table 24.1) enables soft tissue segmentation and some functional imaging, for example, in kidney function studies. The disadvantages of micro-CT include low soft tissue contrast, use of radiation, and limited molecular-level applications. Micro-CT is able to provide a high-quality anatomical framework for functional imaging techniques, particularly in PET and SPECT imaging data analysis.

24.6 MAGNETIC RESONANCE IMAGING (MRI)

MRI relies mainly on the detection of hydrogen nuclei in water and fat to construct high-resolution images. The contrast in these images results from different T_1 and T_2 relaxation times for hydrogen nuclei in different tissue environments. MRI is based on the same principles as liquid-state nuclear magnetic resonance (NMR)—that is, the behavior of nuclei in a magnetic field under the influence of radio-frequency (RF) pulses—but the hardware, pulse sequences, and data processing are somewhat different. Improvements in electronics and computers since the mid-1990s have given MRI resolution capabilities in intact organisms down to approximately 3 mm and therefore tremendous potential as a tool for studying the mechanisms of toxicology. MRI techniques provide detailed information on the response of specific organs to toxicants and can also be used to monitor xenobiotic metabolism *in vivo*. In addition, they could reduce the number of animals required for toxicology studies, as a single animal can be followed over an extended period of time to monitor internal changes.

MRI utilizes NMR and the signal is derived primarily from the hydrogen nuclei (protons) of water molecules. The technique uses a powerful magnetic field to align the magnetization of atoms in the organism and a pulse of RF to alter the alignment of this magnetization. The scanner then detects the rotating magnetic field to produce an image of the scanned area. Unlike radiography or CT, no ionizing radiation is used. Intravenous contrast agents are used to enhance the signal and/or help delineate vessels or tumors. MRI is most useful for imaging soft tissues, especially those with little density contrast, such as the liver or brain, and is most frequently used to provide anatomical images and delineate lesions such as tumors or areas of necrosis in living animal models. Magnetic resonance microscopy (MRM) is MRI with resolutions of better than 100 μm^3 . An advantage of this technique is high resolution (roughly 10–100 μm with no limit of depth) and high soft tissue contrast. MRI contrast agents include chemical exchange saturation transfer (CEST) agents, paramagnetic metal chelates, iron oxide particles, ^{19}F -containing agents, and hyperpolarized molecules. The disadvantage includes limited molecular applications and long scanning times. Functional information can be gathered in a related technique known as magnetic resonance

TABLE 24.1 Classification of X-Ray Contrast Media for Small-Animal CT Imaging

| Iodinated Extracellular Fluid CM (Urographic Water-Soluble CM) | Vascular Space CM (Blood Pool Agent) | Hepatocellular CM (Tissue-Specific CM) | Gastrointestinal CM |
|---|--|--|--------------------------------|
| Monomeric nonionic CM | Macromolecular CM | Iodine-containing lipid | Iodinated CM |
| Iopromide | Dysprosium–DTPA–dextran CM | Iopromide liposomes | Diatrizoate (ionic CM) |
| Iohexol | | Iodixanol liposomes | Iopromide-370 (nonionic CM) |
| Iopamidol | Oil emulsions of ethiodol and iodine | Cholesteryl ioponate | |
| Ioversol | | Polyiodinated triglyceride emulsion (ITG-LE) | Barium sulfate |
| | | Fenestra DC | |
| Iopentol | Iodine-containing micelles | | |
| Iomeprol | MPEG–iodolysine micelles | | |
| Iobitridol | | Dysprosium EOB DTPA | |
| Ioxilan | Nanoparticles | | |
| | Bismuth–sulfide polymer-coating nanoparticles (BPNP) | | |
| | Iron nanoparticles (FeNP) | | |
| | Gold nanoparticles | | |
| Dimeric nonionic CM | Liposomal encapsulation of iodinated CM | | |
| Iotrolan | Polyethylene glycol (PEG)- coated iopromide liposomes | | |
| Iodixanol | PEG iohexol liposomes | | |
| | Lipid mixture with iodixanol | | |
| | Polyiodinated triglyceride emulsion (ITG-LE) | | |
| | Fenestra VC | | |

Source: Adapted from Kiessling and Pichler (2011).

spectroscopy (MRS), which provides information either on the concentration and distribution of magnetic nuclear isotope-labeled drugs in tissues of laboratory animal or on particular endogenous biochemical (metabolites) since a specific pattern of metabolites can be associated with certain diseases and tumors.

24.7 OPTICAL IMAGING

A variety of different kinds of optical imaging techniques have been developed for biomedical applications. They include various microscopy methods such as confocal microscopy, two-photon microscopy, and coherent anti-Stokes Raman scattering (CARS) microscopy for *in vitro* and *ex vivo* applications. These techniques are combined with a number of methods for *in vivo* applications such as various types of microscopic imaging, bioluminescence imaging (BLI), fluorescence imaging, diffused optical tomography, and optical coherence tomography. Over the last decade, the various modes of optical imaging have

grown in popularity and sophistication. Today optical imaging has a wide variety of applications in genomics, proteomics, cell biology, and drug discovery and development. Optical imaging has emerged as a real-time, sensitive, and noninvasive modality for visualization, localization, and measurement of bioactive molecules and molecular processes *in vivo*.

Even though optical imaging cannot compete with PET, SPECT, MRI, and CT in clinical applications today, the advantages of optical imaging (i.e., convenience, sensitivity, cost effectiveness, and nonradioactive material safety) have made it very popular among the traditional imaging modalities for MI in preclinical studies. Optical imaging of receptors, enzymes, gene expression, live cells, and tumors *in vitro* and *in vivo* has deepened our understanding of disease progression and therapeutic response at the molecular, cell, tissue, and whole-animal levels.

In vivo optical imaging includes fluorescence and bioluminescence imaging. Both techniques are highly sensitive (picomolar) at limited depths of a few millimeters, are quick and easy to perform (with a high-throughput capability),

and in general do not require costly instrumentation. This character makes it particularly suited to the drug development and validation process.

Fluorescence¹ imaging uses high-intensity illumination² of a certain wavelength to excite fluorescent molecules in a sample. When a molecule absorbs photons at the appropriate wavelength, electrons are excited to a higher energy level. As these excited electrons “relax” back to the ground state, vibrational energy is lost and, as a result, the emission spectrum is shifted to longer wavelengths. Fluorescence imaging causes excitation of certain fluorophores in a living system by using external light and then detects fluorescence emission with a sensitive CCD camera. The fluorophores can be endogenous molecules (such as collagen or hemoglobin), exogenous fluorescent molecules such as green fluorescent protein (GFP), or small synthetic optical contrast agents.

A limitation of *in vivo* fluorescence imaging is from light attenuation and scattering by adjacent living tissues. Light in the near-infrared (NIR) range (650–900 nm) instead of visible light can improve the light penetration within laboratory animal. It can also minimize the autofluorescence of some endogenous molecules such as hemoglobin, water, lipids, and other biomacromolecules. In a whole mouse illumination experiment, photon counts in the NIR range (670 nm) are about four orders of magnitude higher compared to those in the green light range (530 nm) under similar conditions. NIR fluorescence imaging has provided an effective solution for improving the imaging depth along with sensitivity and specificity. Therefore, fluorophores with emission maxima in the NIR region are important for successful *in vivo* optical imaging and future clinical applications. Commercially available fluorophores,³ with well-defined excitation and emission spectra, can be used to “stain” specific structures or molecules in a specimen. Proper selection of fluorophores allows the identification of multiple targets as long as emission spectra can be cleanly separated and distinguished from autofluorescence.

24.8 ULTRASOUND

The principles of US imaging are founded in the basic physical interaction of sound waves with living tissues of various densities (Szabo, 2004). It is extensively used in

small-animal (dogs and cats) veterinary care (Mattoon and Nyland, 2015). Ultimately the results of these interactions are displayed as an image. Doppler-based modes can also be utilized to evaluate fluid dynamics, primarily velocity, of the circulatory system or within specific organs. Ultrasonography can be used to evaluate both structural and functional end points of multiple target organs within various species, including small and large animals. Contrast agents, such as microbubbles, can be functionalized with specific molecules, such as monoclonal antibody, thus improving image quality (Lindner, 2004). Micro-US is a useful tool for measurement of internal volumes.

Ultrasonography is practiced in various formats such as brightness mode (B-mode), motion mode (M-mode), Doppler (including both color flow and spectral Doppler), and contrast-enhanced imaging. A distinct advantage to the utilization of ultrasonography is simply the many applications that can be practiced in preclinical development. Subjects can be imaged at multiple time points within a given study, and multiple targets can be evaluated. A disadvantage is the lack of high-resolution capabilities, as well as the need for an experienced imaging expert to obtain the acceptable images for analysis. Poor image capture will ultimately lead to unsatisfactory data sets. Ultrasonography represents a noninvasive imaging solution that can be utilized to provide quantitative structural and functional data sets. Investigators have many options available through the application of ultrasonography to match the imaging format to the specific needs of their research projects.

24.8.1 Echocardiography

Echocardiography is the use of US technology to assess various aspects of cardiac function and morphology (Kaddoura, 2002). By using different echo windows (i.e., standard placements of the echo probe) and types of US propagation, qualitative information can be obtained regarding indices of cardiac size, systolic function, diastolic function, and hemodynamics. Furthermore, certain parameters are a reflection of an integrated input of cardiac functions and can be used to provide a global impression of cardiac function (Table 24.2).

In the present studies, the parameters measured varied due to differences in heart size, anatomy, species, movement artifacts, and the equipment used. However, certain parameters are constant across all studies, and these are used for comparative purposes. Two such parameters (or surrogates) are also those used in the Common Terminology Criteria for Adverse Events (CTCAE), v3, an adverse event rating which is used for assessing adverse events during clinical trials. These parameters include an assessment of the left ventricular ejection fraction and fractional shortening of the left ventricle.

¹Fluorescence is the phenomenon where absorption of light of a given wavelength by a molecule is followed by the emission of light at longer (visible) wavelengths.

²Microscope illumination is the radiation incident to a specimen.

³A molecule or part of a molecule that emits fluorescence following excitation with light.

TABLE 24.2 Glossary of Echocardiographic Terms and Changes Associated with a Decreased Cardiac Function

| Parameter | Effect | Change in Cardiac Function |
|--------------------|--|--|
| Heart size | A decrease is associated with a decrease in cardiac function, as there is a reduced volume | LVIDd: left ventricular internal diameter at end diastole (mm) LAD: left atrial diameter (mm) |
| Systolic function | FxS: fractional change in left ventricular diameter (%) | A decrease is associated with decreased cardiac function, reflecting reduced systolic contraction |
| | FxArea: fractional change in left ventricular area (%) | A decrease is associated with decreased cardiac function, reflecting a reduced area of systolic contraction |
| | LVIDs: left ventricular internal diameter at end systole (mm) | An increase is associated with decreased cardiac function, reflecting reduced systolic contraction |
| | PEP/LVET: ratio of the preejection period to left ventricular ejection time | An increase is associated with deficits in contractility or high afterload, a relatively shortened period of systolic contraction |
| | VCFm: mean velocity of circumferential fiber shortening (circ s^{-1}) | A decrease is associated with decreased cardiac function, reflecting a reduced rate of systolic contraction or reduced contractility |
| Diastolic function | IVRT: left ventricular isovolumic relaxation time (ms) | An increase is often associated with impaired myocardial relaxation |
| Hemodynamics | Heart rate (bpm) VTI: velocity time integral of aortic flow profile (mm) | Related to stroke volume decreased with a decrease in cardiac function |
| | VTI \times HR: proportional to cardiac output ($\text{mm}\times\text{min}$) | Decreased with a decrease in cardiac function |
| | AoVel: peak aortic velocity (mm s^{-1}) | Decreased with a decrease in cardiac function |
| Global function | EPSS: separation of the mitral leaflets from the septal wall during the early wave of mitral flow (mm) | Increased with a decrease in cardiac function, may be associated with mitral regurgitation |

A glossary of the terms used, how these reflect indices of cardiac function, and how changes in these parameters reflect decreases in cardiac function are listed in Table 24.2.

24.9 NANOPARTICLE CONTRAST AGENTS

Nanoparticles have been found to be useful in a wide range of therapeutic applications. Used as devices, they can be used to target and treat (by selectively absorbing energy or delivering an anticancer agent) cancer. They can also be used, particularly in small animals, as contrast agents in micro-CT for small animals (Ashton et al., 2015).

24.10 RADIOPHARMACEUTICALS

Radioactive pharmaceutical agents (such as technetium) can be used as diagnostics or to enhance imaging (allowing evaluation of the functioning of tissues such as the brain, heart, lungs, thyroids, liver, kidney, gallbladder, and tumors) or as therapeutics (Kowalsky and Falen, 2011).

Another common use is in studying the metabolism of drugs in development. A small batch of a drug can be synthesized containing one or more radioactive atoms at selected locations in the structure. This allows qualitative and quantitative evaluation of the distribution, metabolism, and excretion of the drug molecules and their metabolites.

24.11 APPLICATIONS OF PRECLINICAL IMAGING IN LABORATORY ANIMALS

24.11.1 Molecular Imaging as an ADME Platform in Drug Screen

Traditional nonclinical studies on absorption, distribution, metabolism, and excretion (ADME) are currently conducted either by “cold” chemistry or by introducing a radiolabeled drug candidate into laboratory animals, usually rodents, and then sacrificing a certain number of the animals at times following dosing that represent each step in the ADME process. The animals are then either necropsied or sectioned and examined and radioactivity is quantitated with scintillation counters or using quantitative whole-body autoradiography (QWBA), respectively, to demonstrate where the drug is distributed in high resolution and how it is cleared from the body. This traditional approach to ADME is a labor-intensive process that requires several groups of animals, with three to five animals per group (if rodents) and five to ten groups, to capture temporal and dose-related changes in the data.

PET, SPECT, QWBA, fluoroscopy, high-resolution US, optical imaging, and other imaging techniques offer numerous benefits in ADME studies. The coregistration of PET/CT and SPECT/CT imaging enables the measurement of rapid kinetic processes in real time and can therefore generate a more accurate picture of the ADME profile. Many of these imaging technologies allow the use of mice as animal models, thereby reducing costs.

In vivo imaging also provides biodistribution information in real time, as compared to the extensive postmortem evaluation procedures required for more traditional approaches (Wang et al., 2015).

In nonclinical imaging studies, animals can be imaged with the same radioisotopes as used in the clinic for humans but at dose levels adapted to the smaller subjects. As a result, successful preclinical studies in animals quickly translate to clinical studies on humans. Provided that small-animal imagers are indeed able to image small subjects with the same utility as humans in the clinic, the ability to use the same tracer in clinical trials as those used in nonclinical studies can result in significant efficiency improvements and time and cost savings when translated into humans in clinical practice.

As a result of these new imaging modalities being available for preclinical development, PET and SPECT imaging have important roles to play in both the nonclinical and clinical stages leading up to submission of a prospective therapeutic to the FDA for marketing approval. *In vivo* MI permits earlier determination of whether a given group of drug candidates will work or not in animal models. By using imaging to look at responses earlier in animal, safety can be further ensured, and resources can be saved by abandoning drugs that do not translate to the human model in favor of those that do by early “go/no go” decisions.

Microscopic imaging methods are also increasingly being recognized as a valuable tool in ADME testing, which can be employed in both *in vitro* and *in vivo* nonclinical studies. Microscope images can reveal multiple pieces of information on the cellular response to drug compounds in one experiment. This might include the simultaneous acquisition of data on drug–receptor binding and any morphological effects of drug treatment. The ability to use multiple protein-specific fluorescent probes in a single experiment is a key enabling technique in high-content imaging. Fluorescence microscopic imaging allows molecules beyond the resolution limit of the light microscope to be visualized (Giepmans et al., 2006). Fluorescence microscopy is a key technique in clinical diagnostic as well as research settings. Confocal⁴ fluorescence microscopy, in particular, has become an essential tool central to the study of structural and molecular dynamics in living cells. Time-lapse imaging can be used to monitor downstream drug effects and, ultimately, excretion from the cell. Once images have been digitally captured, it is also possible to examine the data retrospectively in response to new questions about the drug compound. Microscopy images have the advantage that they are easily machine readable with appropriate image analysis software. This makes microscopy amenable to medium- to high-throughput analysis. Motorized and computer-controlled microscopes are

essential for automated image capture and for incorporation into medium- to high-throughput environments.

Bioluminescent probes have also been used for ADME evaluation. There have been several reports in the literature published describing the synthesis and utility of bioluminescent probes that are substrates of CYP family of enzymes. CYPs are a critical focus of drug development efforts because they are the dominant drug-metabolizing enzymes. CYPs predominantly facilitate drug clearance and in certain cases activation of prodrugs into active compounds. The identification of CYP oxidation sites on newly designed drug molecules might guide redesign efforts so as to minimize breakdown by the CYP family enzymes while retaining its activity. Klaubert and colleagues presented a series of CYP-specific luminescent probes that could help uncover specific CYP functions against a background of multiple CYPs in samples such as liver microsomes and cultured hepatocytes.

It is ideal to have an imaging agent that selectively targets the target tissue, organ, or pathophysiologic lesions for best imaging contrast and diagnostic accuracy *in vivo*. Nevertheless, most of the dyes themselves are not target specific. The *in vivo* performance of an imaging agent or probe can be complicated by its interactions with many biomolecules, membranes, and related cellular permeability or tissue penetration as well as PK processes including ADME. Therefore, it has been challenging to discover and develop an optimal imaging agent for *in vivo*, namely, targeted imaging.

24.11.2 Preclinical Imaging in Oncology

Cancer drug development has the lowest to market success rate of any therapeutic class and is protracted due to the nature of current efficacy markers. Many imaging techniques have been routinely used in the drug discovery process to directly monitor the therapeutic in blood and tumor tissues to evaluate the effects of the drug treatment in the context of tumor. Imaging recently emerges with increasing popularity as it can be used to monitor the changes at the molecular level *in vivo*, and it can help in evaluating treatment efficacy much earlier than traditional clinical end point. For example, FDG PET imaging is one of the most powerful MI techniques available for clinical use to detecting, staging, monitoring, and evaluating the prognosis of cancer (Reske and Kotzerke, 2001). In clinical practice, ¹⁸F-FDG could offer adequate contrast and tumor detection in several cancers, such as lung, colon, and breast cancers and lymphoma. However, the clinical result is not satisfying in other cancers, such as renal, head and neck, prostate, and pancreatic cancers. Therefore, new imaging probes are needed for early tumor detection in both preclinical and clinical research. Recently, one tumor imaging study aims at verifying the capability of ⁶⁴Cu-radiolabeled (a positron emitter; half-life, 12.7h)

⁴Confocal microscopy is an imaging technique that eliminates out-of-focus light in specimens and enables 3-D imaging of thick specimens.

bevacizumab, an anticancer therapeutic antibody targeting tumor angiogenesis, in detecting different types of tumors in early stage of tumor model and comparing with the gold standard ^{18}F -FDG. The project used different strategy from ^{18}F -FDG in order to accurately detect tumors in earlier stage and effectively decrease the nontumor-related hot spots in the background of the ^{18}F -FDG PET imaging. This preliminary imaging result illustrates that ^{64}Cu -bevacizumab would have superior properties as a new generation of tumor detection probe compare to ^{18}F -FDG probe in preclinical imaging study.

The imaging techniques can also be used to evaluate efficacy and therapeutic response of anticancer drug candidate in the same tumor model so as to minimize the subject-to-subject variability and reduce the number of animals required by more traditional methods. One efficacy imaging study represents a powerful way for tracking adoptively transferred T cells and studying their *in vivo* distribution and therapeutic effect in DUC18/CMS5 tumor model by micro-PET imaging. In the efficacy study, the anti-Thy1.2 antibodies conjugated to 1,4,7,10-tetraazacyclododecane-1,4,7,10-tetraacetic acid (DOTA) and radiolabeled with ^{64}Cu were administered to three groups of BALB-Thy1.1 mice on days 4, 7, or 14 post-DUC18 T-cell transfer. The imaging probe successfully detected the transferred cells in tumor tissue and imaged the size and shape of the tumors in the same living subject for therapeutic evaluation. Information obtained from the *in vivo* T-cell trafficking studies could aid in designing protocols that would enhance the efficacy of tumor-specific T cells. The cutting-edge imaging technology helps us further understand in the area of trafficking behavior of adoptively transferred tumor-specific T cells and improve efficacy study with much less animal numbers since the same animal can be repeatedly imaged in different days posttreatment when tumor grows or be ablated during the treatment.

On the other hand, various optical imaging techniques can be used in combination, either simultaneously or sequentially, to provide information from cancer cells and tumor tissues in living animal models. Among them, both fluorescence and bioluminescence imaging techniques have found wide applications for *in vivo* tumor optical imaging in mouse models. These approaches afford the collection of convenient, frequent visualization and measurement of tumor biomarkers in a real-time, sensitive, and noninvasive way.

One of the most successful imaging modes for preclinical imaging study is optical imaging technique which has been largely exploited to monitor the progress of practically all strategies for gene and cell therapy, regardless of vector or indication. The ability to image host responses to both vector administration and therapeutic interventions provides crucial data for clinical translation. By using a combination of bioluminescence and fluorescence, the efficacy of gene transfer as well as therapeutic response can be monitored simultaneously. A variety of vector studies which include

viral vectors and plasmids have been combined with bioluminescence for analyzing and understanding the effect of gene transfer to a variety of tissues and of physiological conditions such as tumors, autoimmune disorders, neurological conditions, etc.

One of the earliest uses of optical imaging probes for MI purposes involved the use of bioluminescent reporter probes. The fusion gene reporter system is an indirect imaging strategy where two or more genes are fused together under the control of a single common promoter giving rise to a single polypeptide or transcript. The output of fusion reporter genes is often in the form of fluorescence from a light-emitting gene such as GFP or RFP as well as other reporters such as the luciferase reporter which generates fluorescence upon addition of a substrate such as luciferin. This system provides the ability to noninvasively obtain molecular information in systems ranging from a single live cell to the multicellular environment of an animal which holds tremendous potential for studying cancer metastasis, efficacy of drug therapy, and many other applications in a preclinical imaging in laboratory animal. They can also be used for imaging of other disease models such as inflammation as demonstrated by Yaghoubi et al. (2007) using a mouse model of arthritis-induced inflammation; the authors were able to successfully image migration of T cells into the inflamed paws. Another use of reporter genes is the ability to study drug–protein interactions. A drug can activate a specific protein whose expression can be analyzed by optical reporters. Enzymes are a class of proteins, whose expression levels in various diseases models can be easily monitored by use of fusion reporters. Ray et al. developed a multimodality sensor that comprised a bioluminescent, a fluorescent, and a PET reporter gene linked together by a four-amino acid spacer (DEVD), a classical caspase 3 cleavage site. The activation of the caspase 3 through regulators of apoptosis could be easily visualized by all the three modalities including bioluminescence and fluorescence. This system could be used for screening of apoptosis activators/inhibitors in a preclinical drug development.

Activatable probes as described earlier are an interesting class of MI probes. These probes are designed to generate a signal in response to specific biological stimuli. The signal generated is amplified and observed by use of optical imaging agent, which is one of the most popular modalities utilized. Most activatable probes are based on fluorescence activation methods. The idea is that in close proximity, a donor (fluorophore) and a quencher (acceptor) silence each other optically because of a variety of quenching phenomena such as fluorescent resonance energy transfer (FRET), dark-quenching mechanisms, and nanoparticle-based surface energy transfer. These probes often described as “molecular beacons” are activated in the presence of specific biomolecules or chemical stimuli, which generate an amplified fluorescence signal. The distance between the donor and acceptor molecules thus plays an important role when designing useful imaging probes.

Proteases (used as probes for cancer imaging) are enzymes which can break down proteins by hydrolyzing the peptide bonds that link the amino acids in the polypeptide chains making up the protein. Proteases are known to be overexpressed under different pathologic conditions including cancer. To date, various activatable probes have been developed to detect and image representative cancer-related proteases, such as matrix metalloproteinases (MMPs), cathepsins, and caspases. A dual-labeled MMP-7 activatable peptide probe was designed for potential use in imaging MMP-7 activity in tumors. It comprised of a quencher NIRQ820 linked to a donor Cy5.5 through a linker (GVPLSLTMGC polypeptide chain) which is a well-known substrate for MMP-7. Another class of quenchers used for design of activatable probes involves use of dark quenchers which have no known fluorescence from the visible to the NIR spectrum. Black Hole Quencher (BHQ) dyes are one of the commercially available dark quenchers and are able to permit efficient quenching. Studies have demonstrated that the synthesis of an MMP-13 activatable peptide probe was designed using a combination of the known MMP-13 substrate GPLGMRGLGK and Cy5.5 and BHQ-3. The study showed a 32 times increase in the optical signal following incubation with MMP-13. The specificity of the peptide substrate was also observed following a decrease in optical signal when cotreated with an MMP-13 inhibitor. Tsien and colleagues developed a similar probe called as activatable cell-penetrating peptide. This system allows for the intracellular transportation of the probe following hydrolysis by extracellular peptidases. The transportation of the probe only occurs following protease-specific activation.

Apoptosis or programmed cell death plays an important role following cancer treatment. Noninvasive approaches to monitor apoptosis will allow for determining the efficacy of existing as well new antitumor therapies in a preclinical and clinical setting. Caspase families of enzymes are known mediators of apoptosis.

Activation of imaging probes linked with caspase target polypeptides could allow for imaging of apoptosis in cells.

Macromolecules such as polymer-based systems, dendrimer, and antibodies have been used for development of MI agents for cancer and other disease models. Enzyme incubation leads to a 12-fold increase in signal under *in vitro* conditions. Dendrimer and other macromolecules such as chitosan-based polymers have been used in a similar manner for MI of cathepsin as well as caspases under *in vitro* as well as *in vivo* conditions. Antibodies have also been studied for developing imaging moieties to specific targets in cancer. Antibody imaging can provide a sensitive, noninvasive means for characterizing cell surfaces. Kobayashi and colleagues have developed a series of antibody-conjugated optical imaging agents. Indocyanine green (ICG) was conjugated to a variety of monoclonal antibodies targeted toward a variety of receptors such as anti-CD25, antiepidermal

growth factor receptor 1 (EGFR1), and antihuman epidermal growth factor receptor 2 (anti-HER2). These conjugates are nonfluorescent as the fluorescence is quenched because of the proximity of the antibody and the fluorescent dye. But intracellular uptake of the antibody by tumor cells and further breakdown allow for fluorescence recovery of the dye for imaging tumor cells. This methodology has been used successfully to image tumors overexpressing certain markers under *in vivo* conditions. Integrins are class of proteins which have a wide variety of functions in mediating cell-cell communication and cell-extracellular matrix interactions. They have wide variety of roles including promoting angiogenesis, cell migration, and cell survival. They have been implicated in a wide variety of cancers such as glioblastomas, melanomas, breast cancer, and head and neck tumors, among others. The most extensively studied integrin is the $\alpha\beta3$ integrin. It is known to bind to a specific ligand, a 3-amino acid polypeptide chain, consisting of arginine-glycine-aspartic acid (RGD). This polypeptide is associated with many receptors such as vitronectin, fibronectin, and thrombospondin that have been associated with many cancers such as brain, ovarian, breast, prostate, etc. Imaging of $\alpha\beta3$ integrin can provide a lot of information about the tumor at the structural, functional, and even molecular level. *In vivo* studies demonstrated high tumor contrast of the targeted probe and good delineation of the tumor. Subsequent studies have shown the use of a wide variety of contrast agents conjugated with the cRGD peptide for imaging a variety of tumor models. Molecular modeling studies have identified preferential interaction of the cRGD sequence with the $\beta3$ subunit of integrins relative to the αv subunit, suggesting that most probes mainly target the $\beta3$ integrin but not the heterodimer $\alpha\beta3$. RGD peptides have also been used for imaging of ischemia in mouse models. In the cardiovascular setting, angiogenesis is triggered by hypoxia and ischemia, and its major consequence in tissues is the restoration of perfusion and oxygenation. RGD dendrimers have been developed for multimodal imaging of limb ischemia in mice (Sinusas et al., 2008). The newer research focuses on developing multimeric compounds for imaging tumors including multivalent ligand-receptor interactions for improved MI. Similar approach has been used for imaging other overexpressed biomarkers such as folate receptors for image-guided resection of ovarian cancer. Van Dam and colleagues reported the development of folate receptor-targeted fluorescent probe (van Dam et al., 2011). The folate molecule was conjugated to FITC for image of folate receptor-positive tumors (FR +ve). The newly designed MI probe was used under intraoperative conditions for identification of folate-positive tumor cells for possible resection. In the pilot study, the injected probe was found to be safe and offer specific and sensitive real-time identification of tumor tissue during surgery in patients with ovarian cancer and the presence of FR- α -positive tumors.

The use of targeted fluorescent agents could provide a paradigm shift in surgical imaging as it allows an engineered approach to improving tumor staging and the technique of cytoreductive surgery, thereby improving the outcome in various cancers.

Nanoparticles have also been utilized for developing agents for fluorescence imaging of cancer and other modalities. Metal-based nanoparticles such as quantum dots (QDs) and gold nanoparticles (AuNPs) have been investigated extensively as agents for optical imaging. These possess unique characteristics such as high photostability, low quenching properties, and high luminosity. QDs are a class of nanoparticles made from semiconductor metals which based on size and composition can generate light over wide spectrum ranging from visible to NIR.

An interesting concept would be to attach multiple fluorescent dyes on the surface of the QDs. The increasing number of the fluorescence molecules proportionally enhances the overlap integral between the QD and the fluorophore acceptor and can reduce their PL emission. The linker between the QD and fluorophore can be specifically designed to be the substrate of an enzyme of interest. Addition of the enzyme will lead to fluorescence recovery allowing for imaging of specific biomarkers. Antibodies conjugated to QDs have also been used for detection of tumors. Nie and colleagues have synthesized QDs targeted toward prostate cancer cells by use of antibodies recognition of the prostate-specific membrane antigen (PSMA). The fluorescence imaging indicated the high tumor targeting after the injection of QD-PSMA Ab conjugates. But there are issues with poor penetration and high autofluorescence following injection of the conjugates. Other approaches for MI involve use of bioluminescence-based detection and fluorescent proteins conjugated to QDs. Bioluminescence resonance energy transfer (BRET)-based imaging moiety was developed by Rao and colleagues. The authors' conjugated Renilla luciferase mutant (Luc8) fused with an MMP-2 substrate polypeptide sequence to QDs. The addition of MMP-2 to this preassembled QD-substrate-Luc8 reduced the BRET ratio and enabled the detection of MMP-2 activity with high sensitivity. QDs have also been used for mapping of sentinel lymph nodes, a site for metastatic tumorigenic cells following primary tumor development. Frangioni and Bawendi first reported the use of NIR QDs in labeling lymph nodes in living subjects. Experiments performed in small and large animals such as mouse and pigs allowed for high-resolution imaging of sentinel lymph nodes using NIR CdTe/CdSe QDs. The mapping allowed for resection of the lymph node up to a 1 cm depth indicating the usefulness of SLN mapping for fluorescence-based resection of tumorigenic lymph nodes. Peptide conjugated to QDs have allowed for active targeting and subsequent imaging of tumors. Ruoslahti and colleagues have reported the discovery of novel peptides targeting the endothelium and lymphatic vessels surrounding the tumor.

Peptides such as CGFECVRQCPERC (denoted as GFE) which binds to membrane dipeptidase on the endothelial cells and KDEPQRRSARLSAKPAPPKPEPKP KKAPAKK (denoted as F3) preferentially bind to blood vessels and tumor cells in various tumors. The peptide CGNKRTRGC (denoted as LyP-1) which recognizes lymphatic vessels and tumor cells in certain tumors conjugated with QDs has shown to be extremely encouraging for fluorescence-based imaging of various types of tumors. Chen et al. have shown the use of cadmium- as well as noncadmium-based integrin $\alpha\beta 3$ -targeted QDs for imaging of gliomas in mice models. These QDs showed high concentration in tumor sites, and immunofluorescence studies confirmed that the majority of the QD fluorescence signal in the tumor colocalized with the tumor vessels. The high efficiency of conjugation for the synthesis of QDs and peptides along with the good sensitivity and specificity of these conjugates to image tumors warrants the further development of such agents for *in vivo* molecular diagnosis of tumor, including early detection as well as image-guided resection of tumor via surgery.

AuNPs have been shown to be promising candidates for biomedical applications (Gad et al., 2012; Stern et al., 2015). They are biocompatible and easily functionalized for attachment of a variety of ligands such as DNA, peptides, proteins, and antibodies for developing target-specific agents. Their unique surface plasmon resonance absorbance allows for distinctive optical properties which can be used for sensing and imaging biomolecules. It has been noted in the literature that attachment of fluorescent molecules to the AuNP surface allows for unique interactions leading to the development of a highly quenched state. These have led to the development of a number of AuNP-based imaging probes. The construct was based on an MMP-2-sensitive peptide sequence conjugation Cy5.5, a fluorescence dye and AuNP. The quenching of Cy5.5 was accomplished through NSET interactions with the AuNPs as well as self-quenching through FRET. Incubation with the enzyme MMP-2 under *in vitro* conditions caused the cleavage of the peptide leading to fluorescence recovery. The probe showed similar results under *in vivo* conditions in mice bearing MMP-2-positive SCC7 tumors. Carbon nanotubes have also found use as potential optical imaging agents. Robinson et al. were one of the first to report the use of carbon nanotubes as potential dual imaging as well as therapeutic agent for cancer imaging and therapy. SWCNT distribution within tumors was tracked at a high spatial resolution due to their intrinsic optical properties. This was one of the first reported uses of carbon nanotubes as candidates for image-guided photothermal treatment of cancer. Magnetic nanoparticles functionalized with fluorophore Cy5.5 and loaded with siRNA have allowed for sensitive detection of siRNA delivery in animal models with both optical imaging and MRI. MRI also allows for monitoring of long-term therapeutic efficacy following injection of the nanoparticles. Other nanoparticle platforms such as

polymeric nanoparticles, silica nanoparticles, and liposomal formulations have all been reported for preclinical optical imaging of cancer, CNS, inflammation, cardiac diseases, and other disease states.

24.11.3 Preclinical Imaging of CNS Disease

Neurodegenerative diseases, such as Alzheimer's disease (AD) and Parkinson's disease (PD), are becoming an increasingly urgent public health concern, particularly among aging populations. PET and SPECT imaging of brain functions help illustration on mechanism and progress of AD and PD in molecular level.

PD is associated with nigral degeneration and striatal dopamine deficiency and is not known to occur in any species other than humans. Therefore, the most widely used animal model for this disease is the toxicity induced in the nigrostriatal pathway of C57BL6 mice following the administration of 1-methyl-4-1,2,3,6-tetrahydropyridine (MPTP). The ^{18}F -FDG PET reveals reduced lentiform nucleus glucose metabolism. Noninvasive functional imaging of dopaminergic change in the striatum was assessed by using PET to visualize and quantify the uptake of ^{18}F -dopamine in the brain of PD animal model.

Showing striatal dopamine terminal dysfunction with PET supports the diagnosis and rationalizes the use of dopaminergic medications. PET imaging can detect changes in striatal dopamine levels after ^{18}F -dopamine administration and correlate these functional changes to motor responses.

The neurobiologic processes support the use of metabolic imaging technique, namely, ^{18}F -FDG PET imaging, in the study of AD as brain perfusion imaging. Amyloid imaging has been used in studies seeking to elucidate the natural history of AD and early detection and monitoring of the treatment of AD.

One of the applications of optical imaging includes cell tracking following transplantation in cases of neurological disorders and stroke by use of bioluminescence. Kim et al. (2004) used BLI to track murine C17.2 NPCs after transplantation in a murine model of stroke. Their aim was not only to follow migration but also to study possible proliferation and quantification of the transplanted cells. It was found that the maximum photon emission was observed at the site of experimental injury. The signal was maintained for 7 days following injection of the cells. Migration was observed to the ischemic site following injection. This migration was not seen in healthy animals indicating clearance of the cells from the animals. QDs have also been used for stem cell trafficking with good success. Kawabori et al. have demonstrated the use of QDs with NIR fluorescence for *in vivo* tracking of cells within the brain. Using a permanent model of middle cerebral arterial occlusion in rats, the authors tested the potential of bone marrow stromal cells for functional recovery when transplantation occurred 7 days poststroke via

two different delivery routes. Intracerebral transplantation of cells was found to be the best mode of transplantation as fluorescence was detected for 7 days following transplantation. No fluorescence was observed following intravenous transplantation. The fluorescent probes mentioned earlier need to be excited by an external light source, which makes the excitation of deeply transplanted cells difficult. Two-photon or multiphoton excitation techniques have been proposed in the literature to overcome this issue. Chronic neurodegenerative diseases have also been imaged at a preclinical level using fluorescent agents. Beta-amyloid peptides were labeled with Cy5.5 and used for imaging transport in normal as well as P-glycoprotein/ABCB1 and BCRP/ABCG2 pump-deficient mice. The increased accumulation of beta-amyloid peptides is correlated to AD patients and elder nondemented patients. Such studies could be useful in evaluating kinetics of brain elimination of intracerebrally injected compounds for a variety of diseases.

24.11.4 Preclinical Imaging of Autoimmune Disease

Autoimmune diseases are a heterogeneous class of diseases characterized by chronic inflammation of the target organ and often requiring lifelong treatment. One of the most important progress in the study of autoimmune diseases is the development of modern MI techniques by the production of specific radioactive probe which contributes to the identification of immune process responsible for various autoimmune diseases, such as rheumatoid arthritis (RA), age-related muscular degeneration (AMD), type 1 diabetes mellitus (IDD), and Crohn's disease. These imaging studies described a novel detection system for determining the localization patterns of arthritogenic anti-GPI IgG in the joints of normal healthy mice, using rodent-scale PET (micro-PET). The micro-PET R4 scanner permits dynamic noninvasive high-resolution imaging of radiolabeled GPI-specific IgG in mice at multiple time points. The dynamic PET imaging results illustrated that anti-GPI IgG rapidly localized within minutes to distal joints of the front and rear limbs and remained there for at least 24 h.

The fluorescence imaging can be used for monitoring immune responses and demonstrating the tissue specificity of the construct in laboratory animal models. A similar approach was used for detection of reactive oxygen species (ROS) found commonly in patients with RA and malignant tumors. Under these conditions there is also overexpression of hyaluronidase which degrades hyaluronic acid (HA). Park et al. have demonstrated a probe to measure the ROS generation by conjugating an NIR fluorophore to a thiolated HA which can be attached to AuNPs. The newly synthesized nanoparticle allowed for ROS detection in laboratory animal models of RA and metastatic ovarian cancer. The signal intensity derived from the abnormal tissues was significantly higher than other systemic organs.

24.11.5 Imaging Animal Model of Infectious Disease

Infectious diseases are the second leading cause of death worldwide. As age progresses, the immune system undergoes numerous changes that may affect our susceptibility to infection. Laboratory animal imaging has become an important research tool in studies of infectious diseases and has significantly contributed to both our understanding of pathogenesis and preclinical investigations on drug development.

The mouse model of *infectious disease* has been extensively investigated worldwide. An important caveat to the epidemiological studies is that they usually combine several different routes of infection. Some studies have indicated that the route of infection with *Francisella* is an important determinant of bacterial dissemination as well as disease progression and outcome. In a PET study, $^{64}\text{Cu}(\text{II})$ -PTSM, a tracer for detecting hypoxia, was used to radiolabel *Francisella tularensis* and to evaluate the bacteria trafficking of different routes in mice. PET imaging study shows the dissemination of *F. tularensis*, the cause of tularemia, when administered intranasally (i.n.), intratracheally (i.t.), intragastrically (i.g.), intradermally (i.d.), intraperitoneally (i.p.), or intravenously (i.v.) in mice. The results demonstrated that *Francisella* rapidly disseminates within hours to multiple tissues via most routes of administration, although different trafficking patterns were observed. Infection via the pulmonary routes resulted in rapid spread to the lung and gastrointestinal (GI) tract. In fact, this direct radiolabeling and imaging strategy can be used for study of various laboratory animals of infectious diseases since the noninvasive imaging techniques permit enhanced information through longitudinal studies of the same animal during the infection process. The future development of multimodality MI studies for investigating the pathogenesis of infection will provide researchers multiple tools to have a positive impact on treatment of infectious diseases.

24.11.6 Preclinical Imaging of Cardiac Disease

Imaging-based approach on noninvasive evaluation of myocardial function plays more and more important role for preclinical imaging research and clinical practice, such as for myocardial metabolism, congestive heart failure, atherosclerosis, thrombosis, and stem cell therapy in myocardial infarction. PET and SPECT/CT imaging are most effective if motion artifacts can be filtered out. This is typically performed by using physiological monitoring equipment to introduce “gating” tags to the raw acquisition data, indicating the start of the cardiac and respiratory cycles by ^{201}Tl - and $^{99\text{m}}\text{Tc}$ -labeled radioactive probes.

In the use of a combination of bioluminescent cells and fluorescent probes, Weissleder et al. (2010) have been able to image the beating of heart cells at a cellular level in live mice. Combined with the intravital microscopy and

multiphoton imaging, the authors demonstrate the *in vivo* optical sectioning and dual-channel time-lapse fluorescence imaging of cardiac ischemia. The NIRF imaging has also been used for imaging of plaques in coronary arteries. Weissleder and colleagues (2010) have developed an NIRF catheter for intravascular imaging of protease activity. The catheter used for the study was of clinical grade and allowed for efficient visualization of plaques in a rabbit model using a commercially available NIRF agent called ProSense®. This study allows for the theranostic use of NIRF imaging of plaques in high-risk patients, thus potentially reducing incidences of acute coronary syndrome with interventional pharmaceutical or mechanical interventions.

Optical imaging technology has also played an important role in the risk minimization and improved clinical translation. The idea is to select fluorescent molecules that will allow for the maximum possibility of long-term clinical success. This will allow for accelerated use of fluorescence-based methodologies in humans. Besides camera-based fluorescence imaging methods, advanced photonic approaches including multispectral optoacoustic tomography (Ntzachristos and Razansky, 2010) or fiber-based confocal methods can further improve the theranostic and diagnostic potential of optical MI in laboratory animal models of human cardiovascular diseases and thus translate to future clinical practice.

All the techniques described to this point are considered and regulated as medical devices. Radiopharmaceuticals and imaging agents, which are injected parenterally into the body or administered orally (in the case of GI tract contrast agents such as barium), are considered to be and are regulated as drugs. As such, we will now consider required nonclinical safety evaluation.

24.12 NONCLINICAL SAFETY ASSESSMENT FOR IMAGING AGENTS

As our tour of imaging modalities and some of the established classes of imaging agents should make clear, this group of molecules covers a diverse range of structures. All are not intended to chemically or metabolically interact with the body to achieve their desired function, but nevertheless can have toxicity.

Characteristics which are considered relevant to potential safety include:

1. Mass dose—How much is administered?

Doses of imaging agents are usually small, though larger for a contrast agent than for a radiodiagnostic agent. As in all of toxicology, the lower the dose, the less likely are adverse events.

2. Route of administration

Imaging agents are administered by routes that best deliver them to their target tissues or regions of the body.

While this most commonly means a parenteral, some are administered by routes that decrease the likelihood of systemic adverse events (such as barium sulfate as contrast medium to the GI tract). Routes determine the potential for off-target hits or physiologic and immunologic effects which greatly influence the potential for toxicity.

3. Frequency of use

Most imaging agents are administered once or infrequently over a period of time, limiting the possibility of either cumulative/chronic effects or of inducing adaptive immune responses. That is, decreasing the CT concentration and time aspect of dose as a metric of systemic exposure.

Accordingly, the nonclinical safety assessment programs need not include longer-term (3-month or longer) repeat-dose studies. When repeat administration of an imaging agent is possible in clinical use (such as when being used to monitor disease or recovery progression), preclinical repeat-dose studies of 14–28-day duration should be performed (FDA, 2004).

Biologic imaging agents are frequently immunogenic with repeat intermittent administration leading to development of antibodies which can alter safety, pharmacokinetics, and performance (Awe et al., 2010).

Such studies should evaluate standard systemic toxicity parameters as well as pharmacokinetics and the formation of appropriate antibodies as well as whole-body distribution of same at the completion of repeat dosing. Human clinical data should also be assessed following repeat-dose clinical studies of biologic imaging (or diagnostic) agents but prior to application for licensure (FDA, 2004; Awe et al., 2010).

4. Biological, physical, and effective half-lives

As should be expected, it is possible that there are several half-lives to be considered. Radionuclides generally have short physical half-lives, which limit their effective half-lives. Persistence (PK) half-life may be different than effective half-life, which will also be the case for potential biologic effect half-life.

For our purposes here, imaging agents will be considered to include both contrast agents and diagnostic radiopharmaceuticals.

Contrast agents serve to improve our ability to visualize tissues, organs, and cells and physiologic/biochemical processes by sharpening the differences in imaging signal intensities between these and surrounding portions of an organism.

Diagnostic radiopharmaceuticals are intended for use in the diagnosis or monitoring of a disease, in whole or in terms of some monitorable part or manifestation of it.

These act by spontaneous disintegration of unstable nuclei emitting a photon or other nuclear particle.

These usually are composed of a radionuclide component (technetium-99, iodine-123, indium-111, or such) and a nonradioactive component that is found to the radionuclide and serves to deliver it to a specific target cell, tissue, organ, or region of the body.

Generally to open an IND for a nonbiologic imaging agent, the requirements are less extensive than for most drugs:

1. Safety pharmacology—Cardiovascular, respiratory, and central nervous system, plus any organ system that the agent is intended to visualize (Gad, 2015).
2. Expanded acute single-dose studies in two species (at least one nonrodent).
3. *In vitro* genotoxicity through conducting the complete three-assay battery as per ICH S2(R1) is advised.
4. Tissue tolerance and special studies as dictated by the route of administration.

Table 24.3 presents the current complete program up to NDA approval. Note that longer (>90-day)-term repeat-dose studies, carcinogenicity studies, and reproductive/developmental studies can be omitted. If larger doses are administered, osmolarity effects should be considered.

Diagnostic radiopharmaceuticals are recommended to have the following modifications due to their special characteristics and the way that they are used:

- Long-term repeat-dose toxicity studies in animals typically can be omitted.
- Long-term rodent carcinogenicity studies typically can be omitted.
- Reproductive toxicology studies can be waived when adequate scientific justification is provided (ICH S2A).
- Genotoxicity studies should be conducted on the non-radioactive component because the genotoxicity of the nonradioactive component should be identified separately from that of the radionuclide. Genotoxicity studies can be waived if adequate scientific justification is provided (ICH S2B).

There are special safety considerations for diagnostic radiopharmaceuticals, including verification of the mass dose of the radiolabeled and unlabeled moiety; assessment of the mass, toxic potency, and receptor interactions for any unlabeled moiety; assessment of potential pharmacologic or physiologic effects due to molecules that bind with receptors or enzymes; and evaluation of all components in the final formulation for toxicity (e.g., excipients, reducing drugs, stabilizers, antioxidants, chelators, impurities, and residual solvents). Specific safety considerations include an analysis of particle size (for products containing particles) and an assessment of instability manifested

TABLE 24.3 Timing of Nonclinical Studies for Nonbiological Products Submitted to an IND

| Study Type | Before Phase 1 | Before Phase 2 | Before Phase 3 | Before NDA |
|--|--|-----------------------------------|---|------------|
| Safety pharmacology | Major organs ^a and organ systems the drug is intended to visualize | | | |
| Toxicokinetic and pharmacokinetic | See ICH guidances | | | |
| Expanded single-dose toxicity | Expanded acute single dose ^b | | | |
| Short-term (2–4-week) multiple-dose toxicity | | Repeat-dose toxicity ^c | | |
| Local tissue tolerance and special | Conduct as necessary based on route irritancy, blood compatibility, protein flocculation, misadministration, and extravasation | | | |
| Radiation dosimetry | If applicable | | | |
| Genotoxicity | <i>In vitro</i> per ICH S9 | As per ICH S9 | | |
| Immunotoxicity | | | May be needed based on molecular structure, biodistribution pattern, class concern, or clinical or nonclinical signal | |
| Reproductive and developmental toxicity | | | Needed or waiver obtained | |
| Drug interaction | | | | As needed |
| Other based on data results | | | | As needed |

^aSee the guidances S7A (“Safety Pharmacology Studies for Human Pharmaceuticals”) and S7B (“Safety Pharmacology Studies for Assessing the Potential for Delayed Ventricular Repolarization (QT Interval Prolongation) by Human Pharmaceuticals”) (note that S7B allows for phase evaluation of the required studies).

^bSee the guidance “Single Dose Acute Toxicity Testing for Pharmaceuticals.”

^cWhen repeat-dose toxicity studies have been performed, but single-dose toxicology studies have not, dose selection for initial human studies will likely be based on the results of the no-observed-adverse-effect level (NOAEL) obtained in the repeat-dose study. The likely result will be a mass dose selection for initial human administration that is lower than if the dose selection had been based on the results of acute single-dose toxicity studies.

by aggregation or precipitation. We also recommend that an individual component be tested if specific toxicological concerns are identified or if toxicological data for that component are lacking. However, if toxicological studies are performed on the combined components of a radiopharmaceutical and no significant toxicity is found, toxicological studies of individual components are seldom required.

Additionally, sufficient data must be submitted from nonclinical and clinical studies to facilitate reasonable calculation of both whole-body and critical organ radiation absorbed dose (see 21 CFR 312.33(a)(10)(ii)).

If the radiopharmaceutical is intended for pediatric use, the radiation absorbed dose should be provided for each of the age groups for which use is intended (neonate, 1-year-old, 5-year-old, 10-year-old, and 15-year-old are potential use classes) using appropriate anthropomorphic phantoms established in the literature (e.g., by the Medical Internal Radiation Dose (MIRD) Committee of the Society of Nuclear Medicine).

24.13 RADIOPHARMACEUTICALS

In research uses, both imaging agents and radiopharmaceuticals are generally excluded from IND requirements (21 CFR 361.1). If, however, there is intention to use these clinically immediately for diagnosis or treatment, an IND is required, and the nonclinical safety testing requirements laid out in this chapter are operative.

Therapeutic radiopharmaceuticals not only have the receptor interaction toxicodynamics concerns already addressed for nonradioactive molecules in this volume but also carry concerns for potential late (delayed) radiation toxicity effects. The FDA has promulgated a specific guidance (FDA, 2011) for the assessment of these concerns. Radiobiologics are specifically excluded from coverage by this guidance and indeed currently have no specific assessment guidances rather than those applying to biologics in general.

Therapeutic radiopharmaceuticals are typically administered systemically to treat cancer. For cancer therapy with curative intent, the radiation absorbed doses delivered by

therapeutic radiopharmaceuticals are comparable to those delivered with external beam radiotherapy (XRT) and are orders of magnitude higher than doses delivered by diagnostic radiopharmaceuticals. At therapeutic doses of radiation, the late radiation toxicities commonly associated with XRT (e.g., renal, pulmonary, neurologic, late bone marrow failures) can also be seen with therapeutic radiopharmaceuticals. With XRT, if the total dose given to an organ is less than its tolerance dose, the probability of symptomatic late radiation toxicity to that organ will be minimal (Halperin et al., 2013). This type of toxicity should not be confused with the radiation-induced secondary malignancies for which the risk is known and accepted as unavoidable. The tolerance doses of most human organs for conventionally fractionated XRT are known and are routinely used to direct the administration of XRT at a dose and schedule that minimizes late toxicity. In the FDA's experience, however, there are few clinical data from which to estimate organ tolerance doses for therapeutic radiopharmaceuticals.

Organ tolerance doses for systemically administered therapeutic radiopharmaceuticals can differ significantly from the published tolerance doses for conventionally fractionated high-dose-rate XRT. With XRT, the dose received by an organ is determined by its proximity to the primary radiation beam and the tumor. Organs within the primary radiation beam and in close proximity to the tumor are at greatest risk. In the case of systemically administered radiopharmaceuticals, the dose received by each organ is determined primarily by the pharmacokinetics and biodistribution of the radiopharmaceutical agent. In addition, the range and type of the radiations emanating from the source organ are critical.

Radiolabeled drug-based therapy is an emerging and complex field with many potential dose-modifying factors such as dose rate and fractionation. Experience with external beam therapy demonstrates that with therapeutic doses of radiation, a relatively small percentage change in total dose could lead to a large change in the probability of complications after the tolerance limit of an organ has been reached. The organ tolerance doses for XRT are based on conventionally fractionated high-dose-rate therapy. Fractionation allows for repair of radiation damage between fractions. In contrast, therapeutic radiopharmaceuticals usually deliver a single dose of radiation at a low dose rate, where damage and repair of that damage occur simultaneously as competing processes. Therefore, organ tolerance doses for systemically administered therapeutic radiopharmaceuticals are not directly comparable to those for XRT. Radiation toxicity has been observed with therapeutic radiopharmaceuticals at estimated organ doses that were below the XRT tolerance doses for the target organs. The entity of low-dose hypersensitivity may account for this discrepancy, as well as could local region concentration of active emitters.

Ionizing radiation causes injury to cells and tissues mainly by damaging nuclear DNA (Hall, 2000), although non-DNA targets have been described (Coppes et al., 2005). Most damaged cells continue to function normally until they die while attempting to undergo mitosis. Thus the time frame in which radiation injury becomes clinically apparent is determined in part by cell turnover time.

In organs with a rapid cell turnover (early reacting normal tissue) (e.g., bone marrow, epidermis, small intestine, and oropharyngeal mucosa), symptoms of radiation injury (e.g., bone marrow failure, desquamation, nausea, vomiting and diarrhea, and oral mucositis) appear within days to weeks of an acute dose of radiation. Radiation injury to these organs is called early or acute radiation toxicity and is often self-limiting and reversible. However, in organs with a slow cell turnover rate (late-responding normal tissue) (e.g., the brain, spinal cord, heart, lungs, liver, kidneys, bone, and bladder), symptoms of radiation injury (e.g., brain necrosis, paralysis, pericardial and myocardial fibrosis with left ventricular failure, interstitial pneumonitis and pulmonary fibrosis, liver or kidney failure, osteoradionecrosis, and hemorrhagic cystitis) do not occur until after a latency period of several months to years during which relatively normal organ function continues. Radiation injury to these organs is referred to as late radiation toxicity and is usually progressive and irreversible (Halperin et al., 2013).

Because acute radiation toxicity becomes apparent within a short time period after administration, proximity in time to radiation exposure can be used as an important criterion in determining whether the radiopharmaceutical is the cause of a particular complication or adverse effect. Such toxicities will become apparent early in a clinical trial, and the trial can be revised or terminated, as appropriate. In contrast, late radiation toxicity in organs such as the kidneys and liver or CNS will not become apparent until months or years after treatment, necessitating longer-term follow-up of treated patients.

With XRT, radiation injury is often limited to organs within the radiation beams, because more distant organs receive much lower doses. With radiopharmaceutical therapy, the risk of radiation injury to an organ is determined by both its intrinsic radiosensitivity and the concentration time–activity curve of the agent in that organ. For example, late radiation effects can occur if the kidneys receive a significant radiation absorbed dose from radiopharmaceuticals that are removed from the systemic circulation by glomerular filtration. The kidneys are known to have a relatively low radiation tolerance dose (23 Gray (Gy) for conventionally fractionated XRT); therefore, late radiation nephritis may be a dose-limiting toxicity for many therapeutic radiopharmaceuticals. Although the bladder tolerance dose is considerably higher (65 Gy), hemorrhagic cystitis can occur as a late effect unless the bladder is adequately irrigated to reduce residency time.

24.14 NONCLINICAL LATE RADIATION TOXICITY STUDIES

24.14.1 Study Goals

For treatment with therapeutic radiopharmaceuticals with curative intent, radiation absorbed doses comparable to doses delivered by XRT must be delivered to the tumor. Because similarly high doses may be unavoidably delivered to normal tissue, radiation toxicities commonly associated with XRT may also be seen with radiopharmaceutical therapy. Because the prescribed radioactivity is given with a small mass dose of the carrier, radiation toxicity, rather than pharmacological toxicity associated with the cold (nonradioactive) drug substance (formulation), is often dose limiting. Historically, nonclinical toxicity studies have been conducted mainly with the cold formulation. Although these studies usually have shown that the no-observed-adverse-effect levels (NOAELs) are many times the clinical mass dose, such studies assess the toxicity of the cold formulation only. Therefore, to assess the potential risk of late radiation toxicity in humans, it is important to conduct late radiation toxicity studies in animals. Such studies may allow the sponsor to:

- Perform controlled experiments that are not ethically feasible in humans
- Identify organs at risk for late radiation toxicity
- Establish a NOAEL for late-occurring, irreversible radiation effects in an appropriate animal species to help select the clinical doses
- Compare the biological effects and tolerance doses of radiation delivered with radiopharmaceutical therapy to those of radiation delivered by XRT in specific organs
- Examine the pathologic changes and possible mechanism of injury
- Distinguish the toxicity of radiopharmaceutical therapy from that of other concomitant therapies
- Determine the amount of organ sparing that could be obtained by fractionating the radiopharmaceutical dose

24.15 STUDY DESIGN

There are challenges associated with the design and conduct of nonclinical late radiation toxicity studies. Therapeutic doses of radiopharmaceuticals require the administration of large amounts of radioactivity. The animals and animal waste will be radioactive, requiring radiation precautions to protect personnel and the general public. Precautions will also be necessary for the disposal of radioactive waste. Despite these challenges, such studies have been conducted and are recommended to optimize dosing and thus ensure

safe clinical trials and patient care. Before initiating late radiation toxicity studies, the sponsor should discuss the specifics of the study design with the applicable review division and consider the following factors.

24.15.1 Good Laboratory Practices

Late radiation toxicity studies conducted for the safety evaluation of a radiopharmaceutical drug product should be conducted in accordance with preexisting requirements under the regulations for good laboratory practices (21 CFR part 58) and the Animal Welfare Act (7 U.S.C. 2131 et seq.).

24.15.2 Species Selection

The sponsor should take into consideration the similarity in dosimetry, biodistribution, and PK profile of the radiopharmaceutical in the selected species and in humans. Suitable animal models to study late radiation toxicity are available. Rats have been shown to develop late radiation nephropathy and pulmonary fibrosis after external beam irradiation. Radiation-induced myocardial fibrosis has been shown to occur in rabbits and dogs. The sponsors should discuss with the applicable review division alternative developmental programs when appropriate animal models are not available to study late radiation toxicity.

24.15.3 Timing of Study

We recommend that the animal studies be scheduled to facilitate the conduct of clinical trials, including the selection of appropriate safety monitoring methods based on findings in such studies. To select the most appropriate species, human dosimetry and PK data using tracer doses should be obtained before initiation of the late radiation toxicity study. Several factors should be considered when assessing the relevance and the timing of the nonclinical studies: (i) the availability of human data following sufficient long-term follow-up in treated patients that might obviate the need for such studies and (ii) the recognition that therapeutic radiopharmaceuticals are sometimes developed to treat patients with no other viable treatment options or for patients who will not survive long enough to be affected by late radiation toxicity. Ideally, the studies should be completed before the start of phase II dose-escalation clinical trials, because late radiation toxicity may not be seen in the first dose cohort until after the entire trial has been completed. However, a phase II trial can, based on risk/benefit considerations, be initiated before completion of the late radiation toxicity study.

24.15.4 General Study Design

The study design should capture acute (occurring within the first few weeks after irradiation) as well as delayed (occurring after a prolonged latency) radiation effects.

Clinically, late radiation toxicity is not observed until at least several months to years following the radiotherapy. In animals, late radiation toxicity usually occurs on a shorter time scale than in humans. For example, the latent period for radiation nephritis in rats ranges from 3 to 7 months. In dogs, renal dysfunction is observed by 10 months. Therefore, to obtain a reasonable estimate of the incidence of specific adverse effects, animals should be monitored for late radiation toxicity for at least 1 year after dosing.

To the extent feasible, the nonclinical study design should closely mimic the design of the anticipated clinical trials including similar amount of injected radioactivity, number of doses, frequency of dosing, and dosing interval, as well as the relative tissue turnover rate and the relative biodistribution and pharmacokinetics in the animal species and human. If both single and fractionated dosing will be studied in clinical trials, a two-arm study design evaluating late radiation toxicity after single as well as fractionated dosing may be necessary. If planned radiation doses in humans will require hematopoietic growth factor support or bone marrow rescue, it may be necessary to support or rescue the irradiated animals so that they will survive comparable doses to allow for late radiation toxicity observations.

Parameters that should be monitored are similar to those evaluated in expanded single- or repeat-dose toxicity studies. These parameters include clinical observations, food consumption, body weight, ophthalmologic examination, hematology, clinical chemistry, urinalysis, and postmortem investigations (e.g., necropsy, organ weights, macroscopic and microscopic examinations).

24.15.5 Dose Levels

Late radiation toxicity studies in animals should include at least four dose levels to identify the NOAEL and dose-related mild-to-severe late radiation toxicity. The study should also include the cold formulation (ideally, the cold isotope equivalent to the highest mass dose) as a control group to distinguish specific radiation effects from potential pharmacological effects of the cold formulation. The dose-limiting toxicities will be severe but are usually reversible (e.g., acute radiation toxicity related to the GI tract and bone marrow). Therefore, the highest dose selected should produce acute radiation toxicity. This dose should be at least twice the maximum planned human dose or radiation tolerance dose for the critical organ (TD5/5 external beam radiation) identified as a possible dose-limiting factor in clinical trials. The dose multiples should be expressed in terms of body surface area and radiation absorbed dose to the critical organs, when critical organs have been identified. The number of animals in each group should be sufficient to ensure survival of an adequate number to perform proper analysis at the completion of study.

24.15.6 Clinical Pathology

Hematology, urinalysis, and clinical chemistries should be performed before dosing, 2 weeks after dosing, then once every 3 months afterward, and at termination. In addition to a standard battery of hematology and clinical chemistry parameters, the study should also include the assessment of relevant biomarkers, if available, to identify late radiation toxicity for the target organ. For example, urinary glutathione *S*-transferase isoenzyme levels can be monitored in addition to blood urea nitrogen and creatinine levels as markers for renal injury. We recommend that the study design be developed in consultation with the FDA to ensure that appropriate long-term toxicity indices are monitored.

24.15.7 Necropsy and Histopathology

Necropsy, including organ weights and macroscopic examination of various organs, should be performed for all animals in the study, including those that died during the study observation period. Detailed histopathologic and microscopic evaluation should be performed at termination.

REFERENCES

- Ashton J, West JL, Bodea CT. (2015) *In vivo* small animal micro-CT using nanoparticle contrast agents. *Front. Pharmacol.* 6:256.
- Awe S, Biade S, Hargus SJ, Kokate T, Laniyonu A, Ouyang Y. (2010) Nonclinical product developmental strategies, safety considerations, and toxicity profiles of medical imaging and radiopharmaceuticals products. In: *Molecular Imaging: Principles and Practice* (WRRG, Ed.). PMPH-USA, Shelton, CT, pp. 589–603.
- Bruggen N, Roberts T. (2003) *Biomedical Imaging in Experimental Neuroscience*. CRC Press, Boca Raton, FL.
- Cai W, Chen X. (2010) Multimodality agents. In: *Molecular Imaging: Principles and Practice* (WRRG, Ed.). PMPH-USA, Shelton, CT, pp. 445–470.
- Cali JJ, Ma D, Wood MG, Meisenheimer PL, Klaubert DH. (2012) Bioluminescent assays for ADME evaluation: dialing in CYP selectivity with luminogenic substrates. *Expert Opin. Drug Metab. Toxicol.* 8:1115–1130.
- Coppes RP, Meter A, Latumalea SP, Roffel AF, Kampinga HH. (2005) Defects in muscarinic receptor-coupled signal transduction in isolated parotid gland cells after *in vivo* irradiation: evidence for a non-DNA target of radiation. *Br. J. Cancer.* 92:539–546.
- van Dam GM, Themelis G, Crane LMA, Harlaar NJ, Pleijhuis RG, Kelder W, Sarantopoulos A, de Jong JS, Arts HJG, van der Zee AGJ, Bart J, Low PS, Ntziachristos V. (2011) Intraoperative tumor-specific fluorescence imaging in ovarian cancer by folate receptor- α targeting: first in-human results. *Nat. Med.* 17:1315–1319.
- FDA (2004) *Guidance for Industry: Medical Imaging Drug and Biologics Products—Part 1: Conducting Safety Assessment*.

- Center for Drug Evaluation and Research (CDER), Center for Biologics Evaluation and Research (CBER), Rockville MD.
- FDA (2011) *Guidance for Industry: Nonclinical Evaluation of Late Radiation Toxicity of Therapeutic Radiopharmaceuticals*. Center for Drug Evaluation and Research (CDER), Silver Spring, MD.
- Gad SC. (2015) Cardiovascular toxicology and its evaluation. In: *Mammalian Toxicology* (Abou-Donia M, Ed.). John Wiley & Sons, Inc., Hoboken, NJ, pp. 425–452.
- Gad SC, Sharp KL, Montgomery C, Payne JD, Goodrich GP. (2012) Evaluation of the toxicity of intravenous delivery of auroshell particles (gold-silica nanoparticles). *Int. J. Toxicol.* 31:584–594.
- Giepmans BNG, Adams SR, Ellisman MH, Tsien RY. (2006) The fluorescent toolbox for assessing protein location and function. *Science*. 312:217–224.
- Hall EJ. (2000) *Radiobiology for the Radiologist*, 5th Ed. Lippincott Williams & Wilkins, Philadelphia, PA.
- Halperin EC, Brady LW, Perez CA, Brady LW. (2013) *Peres & Brady's Principles and Practice of Radiation Oncology*, 6th Ed. Lippincott, Philadelphia, PA.
- Hargreaves RJ. (2008) The role of molecular imaging in drug discovery and development. *Clin. Pharmacol. Ther.* 83(2):349–353.
- Jacoby M. (2015) Medical imaging turns to oft-neglected part of light spectrum. *Chem. Eng. News*. 93(44):10–14.
- Kaddoura S. (2002) *ECHO Made Easy*. Churchill Livingstone, Philadelphia, PA.
- Kiessling F, Pichler BJ. (2011) *Small Animal Imaging*. Springer, Berlin.
- Kim DE, Schellingerhout D, Ishii K, Shah K, Weissleder R. (2004) Imaging of stem cell recruitment to ischemic infarcts in a murine model. *Stroke*. 35:952–957.
- Kowalsky RJ, Falen SW. (2011) *Radiopharmaceuticals in Nuclear Pharmacy and Nuclear Medicine*, 3rd Ed. American Pharmacists Association, Washington, DC.
- Lindner JR. (2004) Molecular imaging with contrast ultrasound and targeted microbubbles. *J. Nucl. Cardiol.* 11:215–221.
- Mattoon JS, Nyland TG. (2015) *Small Animal Diagnostic Ultrasound*, 3rd Ed. Taylor & Francis, Boca Raton, FL.
- Morone MIF, Betterhauser C. (2015) Personalizing nanomedicine. *Chem. Eng. News*. 93:5.
- Ntziachristos V, Razansky D. (2010) Molecular imaging by means of multispectral optoacoustic tomography (MSOT). *Chem. Rev.* 110:2783–2794.
- Reske SN, Kotzerke J. (2001) FDG-PET for clinical use. *Eur. J. Nucl. Med.* 28(11):1707–1723.
- Rogge MC, Taft DR. (2010) *Preclinical Drug Development*, Drugs and Pharmaceutical Sciences, 2nd Ed, Vol. 187. Informa Healthcare, New York.
- Schwaiger M, Dinkelborg L, Schweinfurth H. (Eds.) (2013) *From Morphological Imaging to Molecular Targeting: Implications to Preclinical Development*. Springer, Berlin.
- Sinusas A, Bengel F, Nahrendorf M, Epstein F, Wu J, Villanueva F, Fayad Z, Gropler R. (2008) Multimodality cardiovascular molecular imaging, part I. *Circ. Cardiovasc. Imaging*. 1(3):244–256.
- Stern MS, Solomonov KVV, Sazykina E, Schwartz JA, Gad SC, Goodrich GG. (2015) Initial evaluation of the safety of nanoshell-directed photothermal therapy in the treatment of prostate disease. *Int. J. Toxicol.* 35(1):38–46.
- Szabo TL. (2004) *Diagnostic Ultrasound Imaging: Inside Out*. Elsevier, New York.
- Wang J, Maurer L. (2005) Positron emission tomography: applications in drug discovery and drug development. *Curr. Top. Med. Chem.* 5:1053–1075.
- Wang ZJ, Chang T-TA, Slauter R, Porter SR, Srivatsan A, Chen X. (2015) Molecular imaging techniques in laboratory animals. In: *Animal Models in Toxicology* (Gad SC, Ed.). Informa, Boca Raton, FL.
- Weissleder R, Ross BD, Rehemtulla A, Gambhir SS. (2010) *Molecular Imaging: Principles and Practice*. Peoples Medical Publishing, Mellon, CT.
- Wernick MN, Aarsvold JN. (Eds.) (2004) *Emission Tomography: The Fundamentals of SPECT and PET*. Elsevier, San Diego, CA, pp. 270–292.
- Willmann JK, Bruggen NV, Dinkelborg LM, Gambhir SS. (2008) Molecular imaging in Drug Development. *Nat. Rev. Drug Discov.* 7:591–607.
- Yaghoubi SS, Creusot RJ, Ray P, Fathman CG, Gambhir SS. (2007) Multimodality imaging of T-cell hybridoma trafficking in collagen-induced arthritic mice: image-based estimation of the number of cells accumulating in mouse paws. *J. Biomed. Opt.* 12:064025.

OCCUPATIONAL TOXICOLOGY IN THE PHARMACEUTICAL INDUSTRY

25.1 INTRODUCTION

Most of the assessment of toxicology and safety of therapeutics is focused on the patients who are to benefit from the new medicine. However, there are two other groups of individuals (each of which has different exposure profiles) that one must be concerned about—the healthcare providers (nurses, pharmacists, and physicians) who provide and/or administer the drugs and the individuals involved in manufacturing them. The concerns here are in the realm of occupational toxicology.

Modern toxicology has its roots in the occupational environment. The earliest recorded observations relating exposure to chemical substances and toxic manifestations were made about workers. These include Agricola's identification of the diseases of miners and Pott's investigation of scrotal cancer incidence among chimney sweeps. Occupational toxicology, as its name implies, concerns itself with the toxicological implications of exposure to chemicals in the work environment (EEC, 1992; ACGIH, 2015).

Here we will examine occupational toxicology as it applies to and is currently practiced in the pharmaceutical industry and its support services. This industry, which by definition involves biologically active compounds, has been a driving force in the development of the science of toxicology. The need for a thorough safety evaluation of potential therapeutics prior to marketing approval has driven the continued evolution of toxicological testing methods and the identification of mechanisms of toxic action. The area of occupational toxicology gained great momentum in the pharmaceutical industry from the early 1980s into the twenty-first century including participation in developing TLVs (ACGIH, 2015). This is made clear by the increased number of companies that have implemented occupational toxicology programs during this period. Yet, occupational-related

activities generally represent only a small fraction of activity in safety assessment in the pharmaceutical and biotechnology industries. It is difficult to gauge the level of activity in this occupational area due to the paucity of publications on the subject (Teichman et al., 1988). This is probably the result of the fact that most occupational toxicologists function in an administrative environment and thus experience less pressure to publish and that they in general deal with information relating to new chemical processes that may not be protected by patents. The lack of general knowledge about the function of the occupational toxicologist that has resulted could lead one to conclude that (i) the thorough evaluation of drugs to obtain marketing approval makes an investigation of their potential hazards to manufacturing employees unnecessary and (ii) because they are used therapeutically, pharmaceutical agents are safe under any and all circumstances in occupational settings.

25.2 OCCUPATIONAL TOXICOLOGY VERSUS DRUG SAFETY EVALUATION

While pharmaceutical products are indeed created to treat disease, they cannot necessarily be considered nonhazardous in all situations. The clinician must evaluate the benefits to the patient in light of any side effects or adverse reactions that may be associated with drug usage and any other toxicological properties uncovered in animal studies. Examples range from antibiotics, which have clastogenic properties in mice (for which, to the patient, its activity in suppressing life-threatening infections presents a clear and overriding benefit), to an antineoplastic, which has extreme renal toxicity but which effectively kills established tumors. The occupational toxicologist must look at pharmaceutical

agents in a completely different light. These same risk/benefit analyses do not apply in an occupational setting. Even agents with minimal clinical adverse reactions and whose pharmacological activity could be considered generally beneficial in the clinical setting may present certain employees with health hazards in the manufacturing or healthcare provider settings.

Table 25.1 presents some of the basic differences in the way preclinical and occupational toxicologists must approach their work. Preclinical development of a pharmaceutical product requires exhaustive testing of drug candidates under the requirements of the US Food and Drug Administration (FDA) or equivalent national agency to help predict and evaluate clinical findings and to preclude serious or chronic hazards that would not ordinarily be observed in clinical studies limited in duration and population (FDA, 1990). By contrast, the occupational toxicologist must evaluate the potential of a compound to cause toxicity from unintended exposures via a variety of routes of administration and a wide range of exposure levels of varying lengths and from the perspective that in those that make and administer the drugs, there is no benefit and therefore any biological effect is undesirable. In drug safety evaluations, studies are designed to approximate the clinical setting, particularly in terms of routes of administration and dosage. Thus, most preclinical studies generally focus on oral and/or parenteral administration with dosages that either are comparable to or exceed therapeutic levels. However, neither oral nor parenteral administration is a likely route of exposure among employees. Rather, during manufacturing operations employees are more likely to be exposed via inhalation or direct contact with the skin or eyes. In addition to the route of exposure itself, the effects occurring following direct contact or inhalation exposure may be of a nature not predictable by the studies undertaken for preclinical safety evaluation. Most important among these effects are irritation and sensitization. Dermal, ocular, and respiratory irritation potential generally cannot be predicted from preclinical studies utilizing oral or parenteral administration. Similarly, sensitization, which has the potential to significantly add to the difficulty of conducting manufacturing operations safely, is difficult to evaluate even with

current specific methods and models. For most pharmaceutical agents, testing to ascertain the potential to induce dermal sensitization reactions is not conducted during a typical preclinical development program.

Other important differences lie in the length of treatment and the dosages involved. Therapeutic use of pharmacological agents may be of acute or limited duration, such as in the administration of anti-infectives, or chronic, as with antihypertensive agents. Occupational exposure may also be of varying length, limited by shift or batch manufacturing methods. It is possible, however, that the manufacture of certain high-volume products such as antibiotics may result in daily exposure, if only in limited doses, over a significant portion of a working lifetime. The levels to which employees may be exposed are, in general, potentially lower than those that are used therapeutically, although exposures will vary with the type of operations performed.

The area of occupational toxicology has received a great deal of attention in the chemical industry. Historically, the chemical industry has focused on the occupational environment and developed many of our current toxicological methods to address health and safety concerns. However, since the mid-1970s the chemical industry has increasingly become subject to testing requirements relevant to the protection of the environment and the public at large, as mandated by Environmental Protection Agency (EPA) regulations in the United States (EPA, 1976, 1979). Data development for occupational health hazard evaluation has seldom been sought by the Occupational Safety and Health Administration (OSHA). Consequently, few new test methods have been developed, and those in current use are generally modifications of methods introduced in the 1930s and 1940s. Some of the differences in the issues addressed by occupational toxicologists in the pharmaceutical and chemical industries are highlighted in Table 25.2. Among these, a major difference lies in the physical and functional nature of the substances involved. Pharmaceutical agents are generally handled as solids, while chemical industry products are generally processed as liquids or vapors. Although there is a greater focus on the consequences of occupational exposure, the occupational toxicologist in the chemical industry rarely

TABLE 25.1 Comparison of Occupational and Preclinical Toxicology

| | Occupational | Preclinical |
|--------------------------|---|---|
| Purpose | Potential for effects from unintended exposure | Predict and evaluate clinical findings and preclude serious hazards from clinical use |
| Routes of administration | Inhalation, direct contact with skin or eyes | Oral and/or parenteral most likely |
| Dose level | Not really predictable | Relative to estimated therapeutic dose or maximum tolerated dose |
| Duration of exposure | Extremely variable; depends on campaign/batch, procedure, and so on; may be short daily exposure for working lifetime | Dependent on therapeutic use and test model |

TABLE 25.2 Toxicological Testing Requirements under EC Seventh Amendment (Directive 92/32/EC-Notification of New Substances)

| Quantity Imported to, or Manufactured in, EC | Test Results to Be Submitted |
|---|--|
| <1000 kg year ⁻¹ or <5000 kg total | None needed unless the compound is considered toxic (oral LD ₅₀ 25–200 mg kg ⁻¹) or very toxic (oral LD ₅₀ <25 mg kg ⁻¹) |
| >1000 kg year ⁻¹ or 5000 kg total | Acute oral/dermal LD ₅₀ Acute inhalation LC ₅₀ Skin irritation Eye irritation Skin sensitization 28-day subacute toxicity Mutagenicity (bacteriological and nonbacteriological tests) Acute toxicity to fish (LC ₅₀) Acute toxicity to <i>Daphnia</i> |
| >10000 kg year ⁻¹ or 50000 kg total | Additional tests may be required depending on results, including: Fertility—one or two generation (males or females) Teratology—additional species Subchronic or chronic—90 days–2 years Carcinogenicity Acute and subacute on additional species |

has the available wealth of information that exists in the pharmaceutical industry. New drug dossiers include toxicological information as well as data on the pharmacology, pharmacokinetics, and mechanism of action, and, most importantly, much of this information has been gathered from clinical trials on human beings. Even though the data are not developed for the purpose of evaluating the occupational environment, they can be invaluable for this purpose. Clinical studies, even if the doses and routes of administration may be different from those used in clinical trials/therapeutically, provide insight into the unique responses of the human body. Another important difference in the parameters involved in the chemical and pharmaceutical industries concerns biological activity. The chemical industry strives to minimize it, while pharmaceutical agents are specifically designed to be biologically active. Recent advances in our understanding of molecular and cellular processes have led to the development of agents with improved specificity for unique receptor or molecular targets. These potent agents may present increased hazards for employees and a great challenge to the occupational toxicologist in the pharmaceutical industry.

At the same time, several types of data necessary to ensure proper management of occupational risks associated with a drug substance are not generally useful in evaluating

potential patient risks. So the necessary tests—those to evaluate ocular and skin irritation, sensitization, and inhalation toxicity, as well as assessment of the hazards of by-products and impurities that do not get incorporated into the final therapeutic product—are not performed in the normal course of development.

25.3 REGULATORY PRESSURES IN THE UNITED STATES AND THE EUROPEAN COMMUNITY

The safety and health of workers in the United States is regulated under the Occupational Safety and Health Act of 1970, which established OSHA. Since its inception, OSHA has promulgated a variety of health standards, including compound-specific regulations, permissible exposure limits (PELs), and rules for providing access to medical records (OSHA, 1986) and for communicating to employees the hazards of the materials they handle. The last regulation, the Hazard Communication Standard (OSHA, 1987), is a standard that specifically requires manufacturers or importers to carry out an evaluation of the toxicological properties of chemicals. This standard outlines specific criteria for evaluating substances as hazardous or nonhazardous. However, there still is no US regulatory requirement for testing a compound of unknown toxicity (Gad, 2001). Rather, such a compound could be classified as nonhazardous, based on the unavailability of data. Pharmaceutical agents are generally considered hazardous under the standard since they meet the criterion of having a biological effect on humans. Any adverse reaction observed during clinical use will be construed as toxicity, however irrelevant to the occupational environment. The main result of classification as hazardous is a requirement to develop a Safety Data Sheet (SDS) as the main vehicle for providing to employees. The model SDS suggested for use in complying with the OSHA Hazard Communication Standard contains a great deal of information about the physical properties and hazards and the procedures necessary to deal with the accidental spill, fire, explosion, or accidental contact with hazardous material. The standard requires that all information regarding adverse effects in human beings and most animal toxicity data be included in the SDS, however irrelevant this information may be to the work environment. The resulting SDS can be a highly technical document that may not be the optimal vehicle for conveying this type of information to manufacturing employees handling pharmaceutical agents.

In the environmental area, the EPA's Toxic Substances Control Act (TSCA) (EPA, 1976) regulations for filing of premanufacture notification (PMN) (EPA, 1979) have resulted in the development of toxicological information on many new industrial chemicals. New chemical entities generated for use as pharmaceutical agents are exempted from PMN requirements. This exemption may also extend to all

TABLE 25.3 Comparison of Occupational Toxicology in Pharmaceutical and Chemical Industries

| | Pharmaceutical | Chemical |
|---------------------|---|---|
| Compounds | | |
| Physical state | Generally solids | Liquids, vapors, polymers, solids |
| Biological activity | Designed for biological activity | Strive for biological inertness |
| Toxicology data | | |
| Development | Focus on preclinical evaluation | Focus on occupational and general environment |
| Study length | Acute to chronic for final products, acute for intermediates | Acute to chronic (depending on volume) |
| Human data | | |
| “ADME” ^a | Generally available for oral/parenteral routes | Not generally available |
| Mechanism of action | Targeted during drug development | Not generally studied |
| Adverse effects | Extensive clinical trial studies for final products from oral or parenteral route | Generally only known as a result of overexposure, accident, and so on |

^aADME: absorption, distribution, metabolism, and excretion.

intermediates generated during chemical synthesis. Many pharmaceutical companies have instituted toxicological testing of these compounds even though there is no specific US regulatory impetus to develop such information. The European Community (EC) has implemented several directives that parallel and exceed the OSHA Hazard Communication Standard and TSCA regulations. The European Community Directive 80/1107 (European Economic Community (EEC), 1980) requires employee communication of hazards of chemical substances as well as biological materials. Another EC directive, which was the impetus for the development of New Substances Notification Regulations in several member nations, requires the development of toxicological data on new compounds and does not exempt pharmaceutical agents or isolated intermediates (EEC, 1979). Notification and testing must be conducted in accordance with the amount of the substance manufactured in, or imported into, Europe yearly (20.2). These regulations will support the development of toxicological data on entities manufactured or processed in Europe that can in turn be applied to occupational health hazard evaluations. Pharmaceutical companies are, as demonstrated in Table 25.3, very different from chemical companies in their handling of occupational toxicology.

25.4 ORGANIZATIONAL STRUCTURE

The occupational toxicology function is organized and structured in very different ways across the industry, particularly with so much of the sector now composed of small companies that are not. The function exists in many of the major PhrMA member companies. In most of these companies the occupational toxicology function is located within the Employee Safety/Industrial Hygiene area, while in some it resides within the Research and Development (R&D), Toxicology, or Employee Health/Medical Services

areas. How the function fits into the organization greatly depends on its mission. Occupational toxicology will function well with the R&D environment if the evaluation of occupational health hazards is considered an integral requirement in the development and approval process. In such an organization there would likely be a greater emphasis on developing toxicological data on novel compounds and their synthetic intermediates, rather than on existing processes or such other activities as training. Most occupational toxicology departments are located within the Employee Safety/Industrial Hygiene areas. This organization provides great opportunity for interaction and cooperation with those disciplines that are charged with implementing the toxicologists' recommendations. Good interaction between the occupational toxicologist and the industrial hygienist can be particularly useful in developing and implementing solutions to potential health hazards. However, poor understanding of the limitations of scientific data by the more engineering-oriented safety specialists may lead to unrealistic expectations for easy solutions or answers. The last existing arrangement is for the occupational toxicologist to report into the Medical Services area. This arrangement provides perhaps the easiest interactions for the toxicologist, who shares a common language and understanding of biological systems with the occupational physician. However, in order to effect any changes in the work environment, it is necessary to enlist the aid of the Employee Safety/Industrial Hygiene group, an act that may incur the potential problems mentioned earlier. Clearly, wherever the function is located, the occupational toxicologist must be able to interact well with a variety of disciplines, including R&D, Safety/Industrial Hygiene, Medical Services, Legal Services, Regulatory Affairs, Technical Services, and, of course, Operations Management.

Staffing of industrial toxicology programs varies among the different programs, including groups with a single or a few full-time Ph.D.s who spend all of their time on

occupational issues and those with one or two Ph.D.- or masters-level staff members who may have part-time responsibility for occupational-related issues along with R&D responsibilities. The level of staffing depends, of course, on the activities assigned to the occupational toxicology group, and these may vary from one organization to another. It is impossible to generalize or recommend an adequate staffing level, since that will be dictated by the emphasis placed on specific activities. Whatever the mission of the occupational toxicology function, a high level of education or professional credentials is desirable. A doctoral degree and/or board certification in toxicology should be imperative to be able to interact effectively with many of the other disciplines mentioned earlier, particularly R&D management.

25.5 ACTIVITIES

The scope of activities of occupational toxicologists may be quite different from one organization to another, depending on its specific mission, resources available, and corporate culture. In general, their activities can be divided into four broad areas: data development, data evaluation and dissemination, hazard assessment, and employee training.

25.5.1 Data Evaluation and Dissemination

It is first important to establish who will use the toxicological information provided and how this information will be applied. Unlike the preclinical toxicologist who provides information to other toxicologists, to the regulatory agencies, or to physicians for evaluation of potential therapeutic liabilities, the occupational toxicologist is providing information to a variety of individuals and functions. First, the information will be provided to the industrial hygienist or safety specialist who must evaluate the quality of the work environment and the appropriateness of personal protective equipment. Second, the information will be given to the occupational physician who must evaluate the potential causes of any symptoms reported by employees who may have been exposed to the material. Third, R&D, plant management, and/or manufacturing services must evaluate the need to implement engineering or other controls and weigh these costs against the commercial viability of the product. Last, but not least, the information must be provided to the production employees who will be handling the compound and who need to know of its hazards. Clearly, there is a need to provide the necessary information in such a way that it can be clearly understood by nonscientists. With an audience of such a potentially wide-ranging educational level and understanding, multiple communication vehicles may be necessary.

To be effective the toxicology evaluation must meet several criteria: it must be (i) thorough, (ii) clear and concise,

(iii) in a form appropriate to its target audience, and (iv) include a conclusion or recommendation.

Thoroughness can be achieved through an exhaustive search of the published literature using the available computerized databases. There is a risk, particularly when dealing with pharmaceutical agents, that the most relevant information to occupational toxicology can be overlooked in the great number of clinical case reports, many of which are not relevant to the work environment. In general, little information has been published on the occupational hazards of pharmaceutical agents. A thorough review does not mean a listing of every reported clinical adverse reaction. This type of information is more likely to confuse the reader, and may lead him or her to ignore the important occupational hazards. The toxicologist must, therefore, be extremely selective in performing this evaluation. A review of the available clinical information, however, may yield data that can be used in evaluation, particularly if the product has been tested for dermal administration. An integral part of assuring the thoroughness of the evaluation must be a process of updating the information on a regular basis. In general, it is unlikely that new clinical data will significantly change a review for an established pharmaceutical agent. However, new therapeutic entities should be reviewed more frequently since new data may be published on potential adverse reactions not identified in clinical trials, and these data may impact the occupational evaluation.

When providing information to technical personnel, it is best to use language that does not require the use of a medical dictionary. It is tempting to use medical terms, particularly when quoting from the clinical literature. However, use of these terms may result in poor understanding of the information and may also evoke unnecessary anxiety in the reader. A good rule of thumb is to think of what the reader will do with the information: if the biological effect will require more than a few words to be clearly explained in plain language and it is irrelevant or unimportant to the work environment, it is best left off any communication to the field.

It is not always possible to reach a conclusion regarding the degree of hazard of exposure to a compound, particularly if the data are not directly relevant to the work environment. There is often a temptation to provide a thorough evaluation, setting out all necessary information in plain language but leaving the formulation of a conclusion to the reader. However, if it is difficult for the trained toxicologist to reach such a conclusion, it must be even more difficult for the layperson. If an estimate of the hazard cannot be reached, then the evaluation must be concluded with a recommendation of the type of exposures that may increase hazard or the type of effect that is most likely to occur should there be an overexposure. These may at least give the industrial hygienist or physician a useful reference point. At the same time, it is important to express to the reader the inherent limitations of such a conclusion. The audience may expect black-and-white

answers; if this is not possible, they should be made to understand why.

Perhaps the most difficult part of the communication equation is the ability to match the information to the audience. This may be best illustrated using an example. Over the past 10 years, one company has developed several vehicles for communicating information to various audiences. One instrument is the toxicology review. In general, this is a one- to two-page document that reviews the published literature on the compound. A reference list is prepared and maintained on file for future reference. This review is provided to safety, medical, and industrial hygiene personnel and, if appropriate, research chemists. These individuals have received training to help them understand the terms used and the effects outlined. A second method of communication involves a computerized database. This personal computer-based system, which provides only bottom-line information, is available online via a modem to safety, industrial hygiene, nursing, and research personnel (Sussman and Gáler, 1990). It includes only that information specifically relevant to the work environment and necessary for compliance with OSHA Hazard Communication Standard or EC Directive 80/1107. A third vehicle was developed jointly with an industrial hygiene department and consists of a short paragraph highlighting the specific hazards of the compound followed by safe handling recommendations. A fourth commonly used method is the SDS. The toxicology department prepares the toxicology section of the SDS. The appropriate other disciplines complete the remaining sections, and the completed SDS is then reviewed and approved by a committee. Lastly, for certain compounds, on-site training programs, such as those described in the following text, can be presented by the toxicologist on the hazards of the chemical. These various formats for the same information were developed to serve the information all needs and educational levels of various audiences. This is one approach to filling the need to communicate toxicological information to a variety of groups. The appropriate vehicle for each company will, of course, depend on the available resources and corporate culture. Even a large number of formats may not suffice. The occupational toxicologist should determine, through discussions and follow-up communications, how the information is received and if it is understood. The communication of toxicological information may represent approximately 50% of the occupational toxicologists' responsibilities, thus explaining the level of commitment to developing appropriate formats. The SDS alone is often insufficient for the successful communication of health hazard information to employees.

25.5.2 Data Development

The motivation for conducting toxicological tests for pharmaceutical, chemical intermediates, and impurities arises

from the need to ensure the health of employees by preventing the occurrence of adverse reactions from occupational exposure. Employers thus secondarily minimize the associated potential for work interruption. Programs in place at many larger companies routinely test new drug candidates and/or isolated synthetic intermediates for the purpose of occupational health hazard evaluation.

The development of a toxicological testing program for occupational health hazard evaluation requires consideration of (i) the compounds to be tested, (ii) the stage of drug development at which testing occurs, (iii) the specific tests to be conducted, and (iv) the means for funding.

These four issues are, of course, interdependent, and it is not always possible to deal with one without affecting the others. As indicated previously, drug candidates undergo extensive toxicological testing to ensure an adequate margin of safety for patients. Additional tests are usually required to obtain information specific to the work environment. By contrast, synthetic intermediates are generally not subject to testing for drug safety evaluation. These compounds, if they have the potential to present an exposure hazard to employees, may warrant evaluation. Clearly, it is neither feasible nor necessary to conduct the same level of testing required for drug marketing approval. However, a toxicological assessment can often be developed to determine whether these isolated intermediates have the potential to elicit toxicity from exposures that could occur during work.

Compounds should be selected for testing on the basis of an evaluation of potential exposure and likelihood of them causing adverse effects. The first evaluation is best achieved by including the research chemist, industrial hygienist, and/or safety specialist in the decision-making process. They are in the best possible position for judging potential sources of employee exposure. Including these disciplines in the pretesting stage ensures not only their commitment to the program but also that the studies will be designed with careful consideration of the work experience. The second evaluation, an estimate of toxic effects, may be obtained from a comparison of the compounds in question to known toxic agents, also known as a structure-activity relationship (SAR) evaluation, and the sophistication of available software programs for obtaining a quantitative estimate of toxicity using SAR models. However, it is most likely that an SAR evaluation will be achieved by simple comparison to the final product, similar pharmacological agents, or raw materials that have known toxic properties. Information on potential exposures and toxic effects can, thus, be utilized to decide which compounds to test or to assign priorities to compounds selected for testing.

The timing of these studies depends greatly on the developmental track for the test compounds and may vary for intermediates and final products. Discovery early in the development process that an isolated intermediate poses a significant health hazard may prompt a change in the chemical synthesis or process or in the implementation of

engineering controls or personal protective equipment. Thus it is generally useful to test intermediates at an early stage. This approach presents several practical problems. First, in a long development program, such as occurs in the pharmaceutical industry, there are many opportunities for changing the synthetic route for reasons other than toxicity. Thus, a large percentage of the intermediates tested during the early development stages may be replaced in the ultimate manufacturing process. Second, only a fraction of new drug candidates actually reach the drug approval process. Therefore, the majority of intermediates tested early in development process will likely never reach large-scale manufacture. Conducting a toxicological assessment of intermediates at a later stage in the development process presents a comparable set of advantages and disadvantages: it is more likely that the compounds tested will be manufactured on a large scale, but the ability to make fundamental changes in the chemical process will be greatly diminished. Testing of new drug candidates for occupational health hazards can be an integral part of the drug safety evaluation process. Acute oral toxicity is frequently evaluated as the first step in the drug's safety assessment. Adding acute dermal toxicity and, thus, skin irritation evaluation at the same time can often be accomplished with a minimum impact on the development schedule. This additional information can then be used not only to protect employees manufacturing supplies of the chemical but also those laboratory employees handling test doses of the substance. Eye irritation testing (*in vivo* or *in vitro*) requires minimal amounts of test compounds and could also be accomplished at the same time. Sensitization testing requires a greater commitment in terms of time and the quantity of compounds needed. Therefore, investigation of a compound's allergenic properties is often postponed until sufficient toxicological information is available to permit a decision as to whether the compounds will advance to the next stage in the development process.

Practical considerations of funding and the selection of the testing laboratory need to be addressed when developing an occupational toxicology testing program, and such are addressed in the chapter on contracting out of studies. As indicated earlier, if the activity is located within the R&D department, it may be simple to include the cost of conducting these tests within the new drug's development budget. There is a possible risk in this situation, however, that the safety and industrial hygiene communities may be inadvertently omitted from the prioritization process and the information loop. Explaining the necessity of testing programs to nonscientific management personnel may be challenging. Solutions to these barriers may be found with R&D funding of testing or designation of testing cost, thus possibly including these programs in research funds.

While there are no set regulations on what tests should be done or when these should be conducted and indeed no activity to preclude such testing on animal welfare grounds,

there is general agreement as to the type of effects that need to be addressed: skin and eye irritation, sensitization, and acute oral and dermal toxicity. Testing for these effects generally involves studies of short duration. Thus, results can often be obtained relatively quickly. Additional tests for inhalation toxicity and/or sensory irritation are conducted by several companies.

Although there is general agreement on the effects to be investigated, the methods used have not necessarily been consistent. Several companies have developed protocols uniquely tailored to the needs of their workplace health hazard evaluation and their in-house testing resources. The most common protocols utilized for occupational health hazard evaluation are briefly described on Table 25.4 (Gad and Chengelis, 1998).

These modifications included a combined protocol to assess acute dermal toxicity as well as skin irritation in rabbits and a stepwise approach to acute oral toxicity determination rather than a classic LD₅₀ (Gáler, 1989). Doses are selected based on regulatory criteria, such as those that are required for classification as a toxic under the OSHA Hazard Communication Standard and/or EC 80/1107. Testing for eye irritation involves a modification of current methods using rabbits. While in some views there is no justification for testing cosmetic products in live animals, eye irritation information pertaining to unique pharmacological chemicals is important to protect employees from accidental exposures. There should be careful consideration of whether a version of the rabbit eye irritation test or one of the currently available *in vitro* alternatives will best service the needs of the occupational toxicologist. Current protocols include refinements to the original method, including a reduction in the number of animals used, the application of topical anesthetics to decrease animals' sensation, and rinsing with distilled water following the instillation of the test compound to allow evaluation of the benefits of washing the eye as a first aid measure. Another refinement that may be utilized is a reduction in the amount of material instilled into the eye (Griffith and Yam, 1989). In general, modifications of this type have not affected the reliability of this test (Hatoum et al., 1990) and may, in fact, better simulate possible workplace accidents and provide additional information.

The battery of tests shown in Table 25.4 can provide useful information to complete a workplace hazard assessment. However, they are not the only tools that may be used to determine the toxic potential of a workplace contaminant. Additional tests may be required to provide more rigorous recommendations. Depending on the results of initial tests, a second stage of testing may be initiated to address specific needs. For example, sensory irritation tests may be conducted for compounds that are found to have irritant properties. The sensory irritation test, developed by (Alarie, 1966), is used to develop a parameter—the RD₅₀—that has been directly correlated with threshold limit values (TLVs) for a certain class

TABLE 25.4 Summary of Protocols Used for Current Test Methods

| Test | Species | Method | Dose | Data Application | Additional Data |
|------------------|----------------|---|---|---|--|
| Acute toxicity | | | | | |
| Oral | Rat or mouse | $N=5 \text{ sex}^{-1} \text{ dose}^{-1}$, 14-day observation period, necropsy with/without histopathology | Chosen as limit (0.5 or 5 g kg^{-1}) or for LD_{50} | Classify compounds as harmful/toxic/highly toxic | |
| Dermal | Rabbit | $N=5 \text{ sex}^{-1} \text{ dose}^{-1}$, 24 h dermal application under occlusion, 14-day observation period, necropsy with/without histopathology, dermal irritation scores | Chosen as limit (2 g kg^{-1}) or for LD_{50} | Classify compounds as harmful/toxic/highly toxic, selection of protective equipment | Irritation potential class, target organ information |
| Inhalation | Rat or mouse | $N=5 \text{ sex}^{-1} \text{ dose}^{-1}$, 4 h nose only or whole-body exposure, 14-day observation period, necropsy with/without histopathology | Chosen as limit (20 g m^{-3}) or for LD_{50} | Classify compounds as harmful/toxic/highly toxic, selection of protective equipment | Respiratory irritation potential, target organ information |
| Skin irritation | Rabbit | $N=3$; 4 h application under semioclusive binder to abraded and nonabraded skin, irritation scores at $\frac{1}{2}$, 1, 24, 48, and 72 h | 500 mg site^{-1} | Classify irritation potential | |
| Eye irritation | Rabbit | $N=3 \text{ group}^{-1}$; into right eye, compare to untreated eye; test only compounds with $\text{pH} < 12$ or > 2 ; score at 1, 24, 48, and 72 h and up to 21 days for corneal opacity, conjunctivitis, and iritis; may use a rinse with some animals | 100 mg or 0.1 mL in standard protocol or 10 mg or 0.01 mL in low-volume protocol | Classify eye irritation potential, selective equipment | Evaluate first aid methods, ocular toxicity |
| Sensitization | | | | | |
| Buehler method | Guinea pigs | $N=10-15$; topical applications one to three times per week for 3 weeks, 2-week rest then challenge at a naïve site; concurrent negative/vehicle and positive controls | Up to 500 mg or 0.5 mL dose^{-1} | Classify as sensitizer—most sensitive for moderate to strong sensitizers, selection of protective equipment, evaluation of allergic reactions | Repeated dermal dosing, additional data on skin irritation and dermal absorption |
| Maximization | Guinea pigs | $N=10-15$; combines two intradermal +/- adjuvant and one topical occlusive administration for induction, 2-week rest then topical application for challenge at naïve site; concurrent positive and negative/vehicle controls | 0.1 mL of compound in solution for induction, nonirritating concentration for challenge | May be sensitive to mild sensitizing agents but may also overpredict severity | |
| Local lymph node | Mouse (female) | $N=5$, combines dermal exposure on ear with IV challenge by tail vein on 6th day with tritiated methyl thymidine | $25 \mu\text{L}$ of compound in solvent on each ear for 3 days | Does produce some false positives | No rechallange possible |

of compounds (Kane et al., 1979; Alarie, 1981). However, the usefulness of this test for solid compounds, which include most pharmaceutical agents, has not been determined. Results of genotoxicity tests may present a need for testing in additional systems to assess genotoxic potential. Mechanistic studies may also be appropriate for certain compounds, such as intermediates, in the synthesis of inhibitors of specific enzymes or receptor agonists/antagonists. The information available from clinical, pharmacological, or pharmacokinetic studies on the final drug can be useful in determining possible avenues for investigation. A particularly interesting type of study, yet to be developed, might involve determination of the absorption and the bioavailability of compounds from occupational exposures that could then be related to similar parameters developed in clinical or preclinical pharmacokinetic studies. The need to conduct additional testing will depend on the application of the information by the individual toxicologist and the resources available. The cost of additional tests should be weighed against the cost of applying the most conservative interpretation of the data to the work environment. In some cases implementing stricter controls based on preliminary tests may be less costly than conducting more extensive confirmatory testing.

25.5.3 Occupational Exposure Limits (OELs)

In the manufacture of both active pharmaceutical ingredients (API) and final formulated drug products, before initiation of the manufacturing process for the materials, it is essential to determine what degree of protection is required for the workers during the manufacture and, from there, what steps must be taken (ranging from modest personal protective gear to having a fully contained—and separated from the operators—process).

The end result will be placing the subject material in a category. Such classification systems assign materials to categories ranging from 1 (the lowest level of concern) to either 5 or 6 (depending on the “system” used and the hazard level of concern). The most widely used system is SafeBridge (classes 1–5), but our firm and many other use a 1–6 system (see Table 25.6).

Uncertainty factors are used in response to data quality. The key point here is that production workers theoretically do not derive any health benefit from drug exposure and thus any biological effect is to be avoided.

There has recently been increasing pressure from governmental agencies and animal rights advocates to reduce the number of animals used in toxicological testing. As alternative toxicological methods become more accurate and sophisticated, they should be considered for incorporation into the occupational toxicology battery. Additional tools such as computer-aided quantitative structure–activity relationship (QSAR) evaluations may also be considered as additions or alternatives to animal tests (Jurs et al., 1985;

Klopman, 1985; Frierson et al., 1986; Enslein, 1988). As indicated previously, QSAR methods may be particularly well suited to aid in the selection and/or prioritization of chemicals for testing, particularly in the case of intermediates. Alternative test methods currently under investigation, such as those being proposed for replacement of the Draize eye irritation test, do not appear to be well suited to the testing of pharmaceuticals or their synthetic intermediates (Booman et al., 1988, 1989). An intensive program of testing the available alternative models with compounds in this class is required to determine the ultimate usefulness of these alternative testing methods.

Occupational toxicologists from several companies supported a program to evaluate several experimental models as alternatives to the rabbit eye irritation test (Gáler et al., 1993; Sina et al., 1994). As a result of this cooperative study, several of the participating companies have implemented the routine use of several of these alternative models in their test batteries (R.G. Sussman and J. Sina, personal communications, 2004), thus effectively increasing the number and classes of compounds evaluated in alternative models.

25.5.4 Hazard Assessment

Individual companies have, nonetheless, developed methodologies for establishing occupational exposure limits (OELs) based on the type of data available and other resources available to the company/public. One method involves a formula for extrapolating to an 8 h time-weighted average from the therapeutic dose of the drug using safety factors (Sargent and Kirk, 1988). A group composed of occupational toxicologists from several companies presented a monograph at the second annual Occupational Toxicology Roundtable, held in November 1989, regarding the development of OELs (Gáler et al., 1989, 1992). Several methods are available for developing OELs: analogy, correlation, safety and uncertainty factors, and low-dose extrapolation (Table 25.5). The

TABLE 25.5 Methods for Setting Occupational Exposure Limits (OELs)

| Method | Formula |
|------------------------|--|
| Analogy | $OEL_i = OEL_j$ |
| Correlation | $OEL_j = (PP_i/PP_j) \times OEL_j$ |
| Safety factors | $OEL = \text{reference dose} / UF_1 \times UF_2 \times SF \times BR$ |
| Low-dose extrapolation | $OEL = (\text{rodent RSD} \times (BW_H/BW_R)^{-1/3}) / BR$ or, if PBPK is available $OEL = (\text{human reference dose}) / BR$ |

Source: Adapted from Gáler et al. (1989).

BR, breathing rate for 8-h workday; BW, body weight; OEL, occupational exposure limit; PBPK, physiologically based pharmacokinetic model; PP, physical property; RSD, risk specific dose; SF, safety factor; UF_1 , uncertainty in extrapolation to NOEL; UF_2 , uncertainty from interspecies extrapolation.

appropriate method must be selected on the basis of the appropriateness of the available data. For example, low-dose extrapolation may be appropriate only if sufficient pharmacokinetic data are available to build a suitable physiologically based pharmacokinetic (PBPK) model. Analogy is a method by which the OEL for one compound is adopted for a second, based on the two compounds' structural and functional similarity. This method is suitable only if the two compounds are similar in every aspect, including therapeutic or toxic dose, physical properties, and the like. Correlation is similar to analogy in that it compares similar compounds. However, the OEL is chosen based on a key property of the chemical that influences its toxicological properties. An example would be the use of the relative potency of two drugs as the key property used to adjust the reference OEL. The most commonly used method is that of applying safety and/or uncertainty factors to a reference dose, which may be the lowest therapeutic dose. In using this method, it is important to choose the reference dose and end point with great care. In general, the most sensitive end point should be chosen. Uncertainty factors are selected to approximate levels from effective doses and to account for interspecies differences. A safety factor is selected based on the overall toxicological evaluation of the compound. Because it is necessary to look at the complete toxicological profile, it is generally inappropriate to assign specific values to each type of toxic effect.

Most commonly employed now is the approach suggested initially by Sargent and Kirk (1988), where the OEL for airborne pharmaceutical materials is calculated as

$$\text{OEL} = \frac{\text{NOAEL}(\text{mg kg}^{-1})(\text{BW in kg})}{(\text{UF})(V \text{ in } \text{m}^3)(\alpha)(s)}$$

where NOAEL is the no-effect level in the most sensitive species, BW is body weight, UF is an uncertainty factor (usually 10), V is the volume of air breathed by a worker in 8 h (usually 10 m³ is used), α is an adjustment for bioavailability between routes (if the animal data is from a route other than inhalation), and s is for already slate plasma levels if known (Binks, 2003). The FDA HED factor is commonly used in place of plasma levels if the steady state is unknown.

The OEL process has essentially become an industry standard. While US OSHA regulations do not require that manufacturers establish OELs, European governmental agencies do. The first country to require this activity is the United Kingdom, under its Control of Substances Hazardous to Health (COSHH) Regulations (Health and Safety Executive, 1988; Agius, 1989). A similar requirement is included in the EC Directive 80/1107, which was promulgated into law by other EC member nations.

In creating a program for establishing OELs, several disciplines will generally be included in the development or approval process. In those programs that are currently in

place, Safety/Industrial Hygiene, Manufacturing or Technical Services, Medical Services, Legal Services, R&D, and, of course, Occupational Toxicology may take part in the process. While the ability to make the OEL level in the workplace does not drive the process, the OEL may often be issued as an interim guideline to provide manufacturing locations an opportunity to bring their operations into compliance and for the development of a suitable industrial hygiene sampling and analysis method.

25.5.5 Employee Training

The ultimate client for the services of the occupational toxicologist is the manufacturing or research employee who must be informed of and protected from the potential hazards of chemicals present in the work environment. Providing employees with health hazard information directly through presentations or training programs accomplishes this task better than most written communications and also provides an excellent way to build confidence in the organization and its safety and health programs. The trust gained in this manner can be an invaluable asset when a company is challenged with the manufacture of especially toxic or potent compounds. Face-to-face communication will also promote discussions with line employees and give the occupational toxicologist the opportunity to learn of those adverse health effects that might otherwise go unnoticed or uninvestigated.

There are several areas for which it may be useful to consider developing specific training programs. The Hazard Communication Standard requires that employees be trained to understand the hazards of chemicals as they are outlined in the SDS. There is an obvious need for the occupational toxicologist to be involved in the development of an internal training program or the selection of a commercial program to address this need. In addition to this required training, it may be useful to consider a more in-depth program on basic concepts involved in health hazard evaluation, particularly the dose-response relationship and the different types of chronic health hazards. It may be particularly important to promote an understanding of health hazard information obtained at work as well as through the news media. There are several commercial training programs available that may be useful for this purpose, including computer-based self-training programs and videos.

Table 25.6 presents what is currently the most commonly used classification system for pharmaceutical occupational risks. Based on the assigned categorization from this risk, appropriate worker protection methods are selected (Olsen et al., 2002).

As most pharmaceutical manufacturing is now performed on contract (in "toll" manufacturers), categorization is usually performed on contract (by a firm such as SafeBridge) or by a consultant.

TABLE 25.6 Banding Decision Matrix

| | Band 1 | Band 2 | Band 3 (Default) | Band 4 | Band 5 | Band 6 |
|---|---------------------------|---|----------------------------|--------------------------|---------------------------|---------------------------|
| OEL (TWA, 8h) | > 1 mg m ⁻³ | 1 mg m ⁻³ to 10 µg m ⁻³ | 10–1 µg m ⁻³ | 1–0.1 µg m ⁻³ | < 0.1 µg m ⁻³ | << 0.1 µg m ⁻³ |
| Potency (mg day ⁻¹) | > 100 mg | 1–100 mg | 0.1–1 mg | 0.01–0.1 mg | < 0.01 mg | << 0.01 mg |
| Severity of acute (life-threatening) effects | Low | Low/moderate | Moderate | Moderate/high | High | Extreme |
| Acute warning properties | Excellent | Good | Fair | Fair/poor | Poor | None |
| 28D TDLo (mg kg ⁻¹) | > 500 mg kg ⁻¹ | 200–500 mg kg ⁻¹ | 50–200 mg kg ⁻¹ | 5–50 mg kg ⁻¹ | 0.5–5 mg kg ⁻¹ | < 0.5 mg kg ⁻¹ |
| 90D TDLo (mg kg ⁻¹) | > 200 mg kg ⁻¹ | 80–200 mg kg ⁻¹ | 20–80 mg kg ⁻¹ | 2–20 mg kg ⁻¹ | 0.2–2 mg kg ⁻¹ | < 0.2 mg kg ⁻¹ |
| Onset of acute oral toxicity (rat) | > 500 mg kg ⁻¹ | 200–500 mg kg ⁻¹ | 50–200 mg kg ⁻¹ | 5–50 mg kg ⁻¹ | 0.5–5 mg kg ⁻¹ | < 0.5 mg kg ⁻¹ |
| Onset of warning symptoms | Immediately | Immediately/ minutes–hours | Immediately/hour | Delayed | Indeterminate | Indeterminate |
| Medically treatable symptoms | Yes | Yes | Yes | Borderline | No | No |
| Need for medical intervention | Not required | Not required | May be required | Required | Required | Required |
| Likelihood of chronic effects (e.g., cancer, reproductive, developmental) | Unlikely | Unlikely | Possible | Known | immediately known | immediately known |
| Severity of chronic (life-shortening) effects | None | None | Low | Moderate | High | Extreme |
| Cumulative effects | None | None | Low to moderate | High | Extreme | Extreme |
| Irritation | Not an irritant | Slight to moderate irritant | Moderate | Severe | Corrosive | Extremely corrosive |
| Sensitization | Not expected | Mild sensitizer | Moderate sensitizer | Strong sensitizer | Extreme sensitizer | Extreme sensitizer |
| Reversibility | Reversible | Reversible | Reversible | Irreversible | Irreversible | Irreversible |
| Alteration of quality of life (disability) | Unlikely | Unlikely | Possible | Probably | Unknown | Unknown |

Specific training on compounds of interest can also be useful, particularly before the beginning of a manufacturing campaign, and is particularly effective if coupled with industrial hygiene training on appropriate safe handling techniques. If a testing program is in place, it is good policy to present an evaluation of the information gained in the compound's testing program to the research or manufacturing chemists involved.

Issues continuing to gain in importance for pharmaceutical industry occupational toxicologists are those relating to the new, more potent drugs currently being designed. Current drugs have therapeutic dose levels ranging down to the sub-microgram (and in some cases, even nanogram) levels. Hazard assessment and OEL development, already difficult, may be nearly impossible or inappropriate when dealing with drugs active at pico- or femtogram levels. Alternative methods of evaluating occupational exposures and assuring a safe work environment may need to be developed. Then there are biotechnology products and their associated issues.

A third of all new drugs come from the biotechnology pipeline, most commonly peptides and proteins with significant allergenic potential in an occupational setting. The potential occupational health hazards of this class of potent but high molecular weight products have not been fully evaluated. Because of the inherent functional and structural differences, the extrapolation of testing methods from traditional pharmaceutical products to biotechnology-derived compounds may be fraught with many difficulties. Hypersensitivity and other immunologically based toxicities are particularly of concern for protein- and peptide-based therapeutics.

25.6 CONCLUSION

The field of occupational toxicology in the pharmaceutical industry presents continuing challenges to the industry, particularly as it shifts increasingly to an outsourced function in companies. The occupational toxicologist finds that he or she must become an "expert" in several fields and not be limited to the scientific area. Unlike the preclinical toxicologist, the occupational practitioner functions under less stringent regulatory requirements and minimal precedents. Additionally, as new classes of therapeutic agents enter development and commerce, new concerns and challenges will accompany them.

REFERENCES

ACGIH (American Conference of Governmental Industrial Hygienists). (2015) *Documentation of Threshold Limit Values and Biological Exposure Indices*, 7th with 2002–2015 supplements Ed. American Conference of Governmental Industrial Hygienists, Cincinnati, OH.

- Agius R. (1989) Occupational exposure limits for therapeutic substances. *Ann. Occup. Hyg.* 33:555–562.
- Alarie Y. (1966) Irritating properties of airborne materials to the upper respiratory tract. *Arch. Environ. Health.* 13:433–449.
- Alarie Y. (1981) Bioassay for evaluating the potency of airborne sensory irritants and predicting acceptable levels of exposure in man. *Food Cosmet. Toxicol.* 19:623–626.
- Binks SP. (2003) Occupational toxicology and the control of exposure to pharmaceutical agents at work. *Occup. Med.* 53:363–370.
- Booman KA, Cascieri TM, Demetruilas J, Diedger A, Griffith JF, Grochosky GT, Kong B, McCormick WC, North-Root H, Rozen MG, Sedlak RI. (1988) *In vitro* methods for estimating eye irritancy of cleaning products. Phase I: preliminary assessment. *J. Toxicol. Cutan. Ocul. Toxicol.* 7:173–185.
- Booman KA, DeProspero J, Demetruilas J, Diedger A, Griffith JF, Grochosky G, Kong B, McCormick WC, North-Root HI, Rozen MG, Sedlak RI. (1989) The SDA alternatives program: comparison of *in vitro* data with Draize test data. *J. Toxicol. Cutan. Ocul. Toxicol.* 8:35–49.
- Enslein K. (1988) An overview of structure-activity relationships as an alternative to testing in animals for carcinogenicity, mutagenicity, dermal and eye irritation and acute oral toxicity. *Toxicol. Indust. Health.* 4:479–498.
- Environmental Protection Agency (EPA) (1976) Toxic Substances Control Act, Public Law 94–469, 94th Congress.
- Environmental Protection Agency (EPA). (1979) Toxic Substances Control Act premanufacture testing of new chemical substances. Guidance for premanufacture testing: discussion of policy issues, alternative approaches and tests methods. *Fed. Reg.* 44:16240–16292.
- European Economic Community (EEC). (1979) Council Directive 67/548/EEC as amended by Directive 79/831/EEC. *Off. J. Eur. Communities No. L259.* 15(10):10.
- European Economic Community (EEC). (1980) Council Directive 80/1107/EEC of 27 November 1980 on the protection of workers from the risks related to exposure to chemical, physical and biological agents at work. *The European Community Directive, Brussels.*
- European Economic Community (EEC). (1992) Council Directive 92/32/EC of 27 November, 1980, on the protection of workers from the risks related to exposure to chemical, physical, and biological agents at work. *Off. J. Eur. Communities No. L327.* 3(12):8.
- Food and Drug Administration (1990) *Applications for FDA Approval to Market a New Drug or an Antibiotic Drug*. 21 CFR Part 314. U.S. Government Publishing Office, Washington, DC.
- Frierson MR, Klopman G, Rosenkranz HS. (1986) Structure-activity relationships (SARs) among mutagens and carcinogens: a review. *Environ. Mutagen.* 8:283–327.
- Gad SC. (2001) *Regulatory Toxicology*, 2nd Ed. Taylor & Francis, Philadelphia, PA.
- Gad SC, Chengelis CP. (1998) *Acute Toxicology Testing*, 2nd Ed. Academic Press, San Diego, CA.
- Gáler DM. (1989) A testing battery for evaluating occupational health hazards. *J. Am. Coll. Toxicol.* 8:1215.

- Gáler DM, Leung H-W, Sussman RG, and Trzos RJ. (1989) *Scientific and practical considerations for developing occupational exposure limits for chemical substances*. Presentation to the 2nd Annual Occupational Toxicology Roundtable, November 21, 1989, Rahway, NJ.
- Gáler DM, Leung HW, Sussman RG, Trzos RJ. (1992) Scientific and practical considerations for the development of occupational exposure limits (OELs) for chemical substances. *Regul. Toxicol. Pharmacol.* 15:291–306.
- Gáler DM, Curren R, Gad SC, Gautheron P, Leong B, Miller K, Sargent E, Shah PV, Sina J, Sussman RG. (1993) A 10-company collaborative evaluation of alternatives to the eye irritation test using chemical intermediates. *Altern. Methods Toxicol.* 9:237.
- Griffith JF, Yam J. (1989) The low-volume eye irritation test. Current status of use and acceptance. *J. Am. Coll. Toxicol.* 8:1215.
- Hatoum NS, Leach CL, Talmsa DM, Gibbons RD, Garvin PF. (1990) A statistical basis for using fewer rabbits in dermal irritation testing. *J. Am. Coll. Toxicol.* 9:49–59.
- Health and Safety Executive—U.K. (1988) *The Control of Substances Hazardous to Health Regulations*, Statutory Instruments 1988, No. 1657, HMSO, London, pp. 43.
- Jurs PC, Stouch TP, Czerwinski M, Narvaez JN. (1985) Computer-assisted studies of molecular structure–activity relationships. *J. Chem. Inf. Comput. Sci.* 25:296–308.
- Kane LE, Barrow CS, Alarie Y. (1979) A short-term test to predict acceptable levels of exposure to airborne sensory irritants. *Am. Ind. Hyg. Assoc. J.* 40:207–229.
- Klopman G. (1985) Predicting toxicity through a computer automated structure evaluation program. *Environ. Health Perspect.* 61:269–274.
- Occupational Safety and Health Administration (OSHA) (1986) *Administrative Rule Concerning OSHA Access to Employee Medical Records*. 29 CFR Part 1913. U.S. Government Publishing Office, Washington, DC.
- Olsen MJ, Slaman CW, Guerleo FJ. (2002) An integrated approach to hazard evaluation for worker safety assessment of pharmaceuticals and synthetic intermediates. *Toxicol. Sci.* 66:328–340.
- OSHA (1987) *Hazard Communication Standard*. 29 CFR Part 1910, Subpart Z, Section 1910.1200. U.S. Government Publishing Office, Washington, DC.
- Sargent EV, Kirk GD. (1988) Establishing airborne exposure control limits in the pharmaceutical industry. *Am. Ind. Hyg. Assoc. J.* 49:309–313.
- Sina JF, Gáler DM, Sussman RG, Gautheron PD, Sargent EV, Leong B, Shah PV, Curren R, Miller K. (1994) A collaborative evaluation of seven alternatives to the Draize eye irritations test using pharmaceutical intermediates. *Fund. Appl. Toxicol.* 26:20–31.
- Sussman R, Gáler D. (1990) A database for communicating toxicology information to safety personnel. *Toxicologist.* 10:87.
- Teichman RF, Fallon LF, Brandt-Rauf P. (1988) Health effects of workers in the pharmaceutical industry: a review. *J. Soc. Occup. Med.* 38:55–57.

STRATEGY AND PHASING FOR NONCLINICAL DRUG SAFETY EVALUATION IN THE DISCOVERY AND DEVELOPMENT OF PHARMACEUTICALS

26.1 INTRODUCTION

The preclinical assessment of the safety of potential new pharmaceuticals represents a special case of the general practice of nonclinical safety assessment (Gad, 1996, 2000, 2009) possessing its own peculiarities and special considerations and differing in several ways from the practice of toxicology in other fields—for some significant reasons. Because of the economics involved and the essential close intertwining with other activities (e.g., clinical trials, chemical process optimization, formulation development, regulatory reviews, etc.), the development and execution of a crisp, timely, and flexible, yet scientifically sound, program is a prerequisite for success. The ultimate aim of preclinical assessment also makes it different. A good pharmaceutical safety assessment program seeks to efficiently and effectively move safe, potential therapeutic agents into, and support them through, the clinical evaluation, then to registration, and, finally, to the market. This requires the quick identification of those agents that are either fundamentally not safe or that may require a paradigm shift in their formulation or use. At the same time, the very biological activity which makes a drug efficacious also acts to complicate the design and interpretation of safety studies—especially true for biologics.

Such evaluations occur on different time scales because of different objectives. There is the traditional Big Pharma case and different versions used predominantly by the much more numerous small pharmaceutical cases. These small (or virtual) pharma cases may be either the short case (do only what is required to initiate and support initial clinical trials) or the midgame (a variety where studies to support further clinical trials are preformed but spread out).

Pharmaceuticals, unlike industrial chemicals, agricultural chemicals, and environmental agents, are intended to have human exposure and biological activity. And, unlike these materials and food additives, pharmaceuticals are intended to have biological effects on the people that receive them. Frequently, the interpretation of results and the formulation of decisions about the continued development and eventual use of a drug are based on an understanding of both the potential adverse effects of the agent (its safety) and its likely benefits, as well as the dose separation between these two (the “therapeutic index”). This makes a clear understanding of dose–response relationships critical, so that the actual risk/benefit ratio can be identified. It is also essential that the pharmacokinetics be understood and that “doses” (plasma tissue levels) at target organ sites are known (Scheuplein et al., 1990). Integral evaluation of pharmacokinetics is essential to any effective safety evaluation program.

The development and safety evaluation of pharmaceuticals have many aspects specified by regulatory agencies, and this has also tended to make the process more complex (until recently, as the International Conference on Harmonization (ICH) has tended to take hold) as markets have truly become global. An extensive set of safety evaluations is absolutely required before a product is ever approved for market. There are even novels on the subject (see Zbinden, 1992). Regulatory agencies have increasingly come to require not only the establishment of a “clean dose” in two species with adequate safety factors to cover potential differences between species but also an elucidation of the mechanisms underlying such adverse effects as are seen at higher doses and are not well understood. These regulatory requirements are compelling for the pharmaceutical toxicologist (Traina, 1983; Smith, 1992). There is not, however, a set menu of what must be

done. Rather, much (particularly in terms of the timing of testing) is open to professional judgment and is tailored for the specific agent involved and its therapeutic claim.

The discovery, development, and registration of a pharmaceutical are an immensely expensive operation and represent a rather unique challenge. PhRMA estimated that for every 9000–10000 compounds specifically synthesized or isolated as potential therapeutics, one (on average) will actually reach the market. Other estimates suggest that a more realistic figure is one in every 200 that enter clinical trials will make it to market. This process is illustrated diagrammatically in Figure 26.1. Each successive stage in the process is more expensive, making it of great interest to identify as early as possible those agents that are likely not to go the entire distance, allowing a concentration of effort on the compounds that have the highest probability of reaching the market. Compounds “drop out” of the process primarily for three reasons:

1. Toxicity or (lack of toxicity tolerance)
2. (Lack of) efficacy
3. (Lack of) bioavailability of the active moiety in man

Early identification of poor or noncompetitive candidates in each of these three categories is thus extremely important (Fishlock, 1990), forming the basis for the use of screening in pharmaceutical discovery and development. How much and which resources to invest in screening, and each successive step in support of the development of a potential drug, are matters of strategy and phasing that are detailed in a later section of this chapter. *In vitro* methods are increasingly providing new tools for use in both early screening and the understanding of mechanisms of observed toxicity in preclinical and clinical studies (Gad, 1989b, 2001, 2009), particularly with the growing capabilities and influence of

genomic and proteomic technologies. This is increasingly important as the societal concern over drug prices has grown (Littlehales, 1999). Additionally, the marketplace for new drugs is exceedingly competitive. The rewards for being either early (first or second) into the marketplace or achieving a significant therapeutic advantage are enormous in terms of eventual market share. Additionally, the first drug approved sets agency expectations for those drugs which follow. In mid-2014, there were 508 pharmaceutical products awaiting approval (120 of these biotech products)—the “oldest” having been in review for 7 years—and some 22000 additional agents in the Investigation New Drug (IND) stage (Food and Drug Administration (FDA) website). Not all of these (particularly the oldest) will be economically successful.

The successful operation of a safety assessment program in the pharmaceutical industry requires that four different phases of the product-related operation be simultaneously supported. These four phases of pharmaceutical product support (discovery support, IND support, clinical and registration support, and product support) constitute the vast majority of what is done by the safety assessment groups in the pharmaceutical industry. The constant adjustment of balance of resources between these four areas is the greatest management challenge in pharmaceutical safety assessment. An additional area, occupational toxicology, is conducted in a manner similar to that for industrial environments and is the subject of Chapter 14 of this volume. In most companies, occupational toxicology is the responsibility of a separate group.

The usual way in which transition (or “flow”) between the different phases is handled in safety assessment is to use a tiered testing approach. Each tier generates more specific data (and costs more to do so) and draws on the information generated in earlier tiers to refine the design of new studies. Different tiers are keyed to the support of successive decision

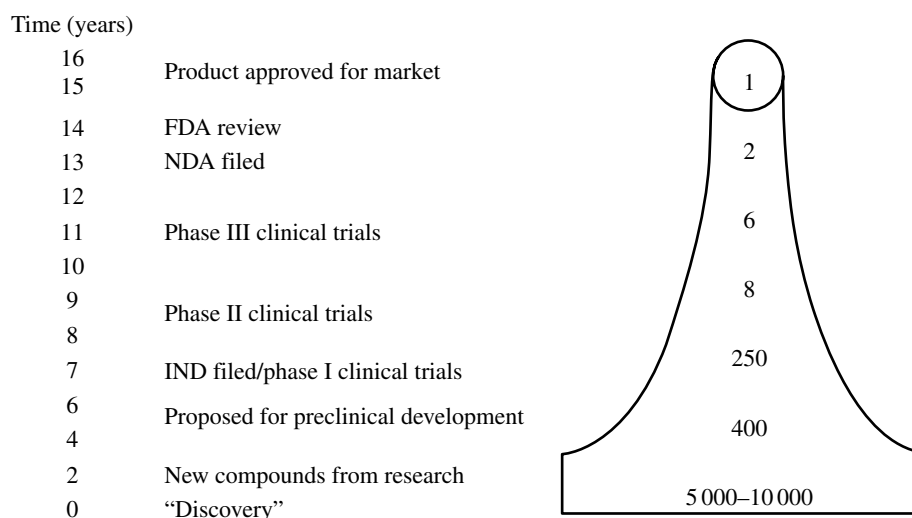


FIGURE 26.1 Attrition during development of new molecules with promise of therapeutic potential. Over the course of taking a new molecular entity through scale-up, safety and efficacy testing, and, finally, to market, typically only one out of every 9000–10000 will go to the marketplace.

points (go/no go points) in the development process, with the intent of reducing risks as early as possible.

The first real critical decisions concerning the potential use of a compound in humans are the most difficult. They require an understanding of how well particular animal models work in predicting adverse effects in man (usually very well, but there are notable lapses; e.g., giving false positives and false negatives) and an understanding of what initial clinical trials are intended to do. Though an approved IND grants one entry into limited evaluations of drug effects in human, flexibility in the execution and analysis of these studies offers a significant opportunity to also investigate efficacy (O'Grady and Linet, 1990).

Once past the discovery and initial development stages, the safety assessment aspects of the process become extremely tightly connected with the other aspects of the development of a compound, particularly the clinical aspects. These interconnections are coordinated by project management systems. At many times during the early years of the development process, safety assessment constitutes the rate-limiting step—it is, in the language of project management, on the critical path.

Another way in which pharmaceutical safety assessment varies from toxicology as practiced in other industries is that it is a much more multidisciplinary and integrated process. This particularly stands out in the incorporation of the evaluation of absorption, distribution, metabolism, and excretion (ADME) aspects in the safety evaluation process. These pharmacokinetic/metabolism (PKM) aspects are evaluated for each of the animal model species (most commonly the rat and dog or primate) utilized to evaluate the preclinical systemic toxicity of a potential drug prior to evaluation in man. Frequently, *in vitro* characterizations of metabolism for model (or potential model) species and man are performed to allow optimal model selection and understanding of findings. This allows for an early appreciation of both the potential bioavailability of active drug moieties and the relative predictive values of the various animal models. Such data early on are also very useful (in fact, sometimes essential) in setting dose levels for later animal studies and in projecting safe dose levels for clinical use. Unlike the case in most other areas of industrial toxicology, one is not limited to extrapolating the relationships between administered dose and systemic effects. Rather, one has significant information on systemic levels of the therapeutic moiety; typically, total area under the curve (AUC), peak plasma levels (C_{\max}), and plasma half-lives, at a minimum. Chapter 18 looks at these aspects in detail.

The state of the art for preclinical safety assessment has now developed to the point where the resulting products of the effort (reports, IND/new drug application (NDA) summaries, and the overall professional assessment of them) are expected to reflect and integrate the best effort of all the available scientific disciplines. Actual data and discussion should thus come from toxicology, pharmacology, pathology, and metabolism, at a minimum. The success of current premarket efforts to develop and ensure that only safe drugs

make it to market is generally good, but clearly not perfect. This is reflected in popular (Arnst, 1998; Raeburn, 1999) and professional (Lazaron et al., 1998; Moore, 1998) articles looking at both the number of recent marketed drug withdrawals for safety (summarized in Table 26.1) and at rates of drug-related adverse drug events and deaths in hospital patients. It is hoped that this system can be improved, and there are continuing efforts to improve or optimize drug candidate selection and development (Lesko et al., 2000). Indeed, the entire pharmaceutical development paradigm is clearly in need of a complete change—a synthesis, as opposed to a continual application of small corrections.

26.2 REGULATORY REQUIREMENTS

Minimum standards and requirements for safety assessment of new pharmaceuticals are established by the need to meet regulatory requirements for developing, and eventually gaining approval to market, the agent. Determining what these requirements are is complicated by (i) the need to compete in a global market, which means gaining regulatory approval in multiple countries that do not have the same standards or requirements, and (ii) the fact that the requirements are documented as guidelines, the interpretation of which is subject to change as experience alters judgments. The ICH process has much improved this situation, as detailed in Chapter 2. ICH (M3)(R2) (2009) clearly denotes what nonclinical studies are required to support clinical drug development. Unfortunately, since then additional requirements (safety pharmacology and immunotoxicology) have been promulgated by additional guidelines. Accordingly, M3 is not all current or inclusive in its guidance.

Standards for the performance of studies (which is one part of regulatory requirements) have as their most important component Good Laboratory Practices (GLPs). GLPs largely dictate the logistics of safety assessment—training, adherence to other regulations (such as those governing the requirements for animal care), and (most of all) the documentation and record-keeping that are involved in the process. There are multiple sets of GLP regulations (in the United States alone, agencies such as the FDA and EPA each have their own) that are not identical; however, adherence to US FDA GLPs (FDA, 1987a) will rarely lead one astray.

Not all studies that are done to assess the preclinical safety of a new pharmaceutical need to be done in strict adherence to GLPs. Those studies that are “meant to support the safety of a new agent” (i.e., are *required* by regulatory guidelines) must be so conducted or run a significant risk of rejection. However, there are also many other studies of an exploratory nature (such as range finders and studies done to understand the mechanisms of toxicity) that are not required by the FDA, which may be done without strict adherence to GLPs. A common example is those studies performed early on to support research in selecting candidate agents. Such studies do not meet the requirements for having a validated analytical

TABLE 26.1 Postapproval Adverse Side Effects and Related Drug Withdrawals Since 1990

| Year | Drug | Indication/Class | Causative Side Effect |
|-----------|---|--------------------------------------|---|
| 1991 | Enkaid (4 years on market) | Antiarrhythmic | Cardiovascular (sudden cardiac death) |
| 1992 | Temafloracin | Antibiotic | Blood and kidney |
| 1997 | Fenfluramine*/dexafluramine (combo used since 1984) (*24 years on market) | Diet pill | Heart valve abnormalities |
| 1998 | Posicor (mibefradil) (1 year on market) | Ca ²⁺ channel blocker | Lethal drug interactions (inhibited liver enzymes) |
| | Duract (bronfemic sodium) (early preapproval warnings of liver enzymes) | Pain relief | Liver damage |
| 1999 | Trovan (use severely restricted) | Antibiotic | Liver/kidney damage |
| | Raxar | Quinolone antibiotic | QT internal prolongation/ventricular arrhythmias (deaths) |
| | Hismanal | Antihistamine | Drug–drug interactions |
| | RotaShield | Rotavirus vaccine | Bowel obstruction |
| 2000 | Renzulin (approved December 1996) | Type 2 diabetes | Liver damage |
| | Propulsid | Heartburn | Cardiovascular irregularities/deaths |
| | Lotronex | Irritable bowel syndrome | Ischemic colitis/death |
| 2001 | Phenylpropanolamine (PPA) | OTC ingredient | Hemorrhagic stroke |
| | Baychlor | Cholesterol reducing (statin) | Rhabdomyolysis (muscle-weakening) (deaths) |
| 2002/2003 | None | | |
| 2004 | Serzone | Antidepressant | Liver failure and injury |
| | Vioxx | Arthritis (COX-2 inhibitor) | Heart attack/cardiovascular (thrombosis) |
| 2005 | Tysabri | Multiple sclerosis (MS) | Progressive multifocal leukoencephalopathy (PML) |
| | Bextra | Arthritis (COX-2 inhibitor) | Skin reaction (sometimes fatal) |
| 2006 | Dolophine (methadone hydrochloride) | Treatment of moderate to severe pain | Respiratory depression and cardiac arrhythmias |
| 2007 | Zelnorm | Constipation | Cardiovascular safety |
| | Permax | Parkinson's Disease | Heart-valve damage |
| 2008 | Trasylol | Bleeding | Increased risk of death |
| | Acomplia | Weight loss | Severe depression/suicide |
| 2009 | Raptiva | Psoriasis | PML |
| 2010 | Meridia | Weight loss | Cardiovascular safety |
| | Avandia | Diabetes | Cardiovascular safety |
| | Darvon&Darvocet | Pain | Addiction |
| | Mylotarg | AML | Veno-occlusive disease |
| 2011 | Viva Globin | Primary immune deficiencies | Thrombolytic adverse extents (CV) |

Fifty-one percent of approval drugs had serious postapproval identified side effects.
FDAMA passed in 1997.

method to verify the identity, composition, and stability of materials being assayed, yet they are essential to the processes of discovery and development of new drugs. All such studies must eventually be reported to the FDA if an IND application is filed, but the FDA does not in practice “reject” such studies (and therefore the IND) because they are “non-GLP.”

There is a second set of “standards” of study conduct that are less well defined. These are “generally accepted practice,” and though not written down in regulation, are just as important as GLPs for studies to be accepted by the FDA and the scientific community. These standards, which are set by what is generally accepted as good science by the scientific community, include techniques, instruments utilized, and interpretation of results. Most of the chapters in this book will reflect these generally accepted practices in one form or another.

Guidelines establish which studies must be done for each step in the process of development. Though guidelines

supposedly are suggestion (and not requirements), they are in fact generally treated as minimums by the promulgating agency. The exceptions to this are special cases where a drug is to meet some significant need (a life-threatening disease such as acquired immunodeficiency syndrome (AIDS) or ALS) or where there are real technological limitations as to what can be done (as with many of the new biologically derived (or biotechnology) agents, where limitations on compound availability and biological responses make traditional approaches inappropriate).

There are some significant differences in guideline requirements between the major countries (see Gad (2010) for an excellent international overview of requirements), though this source is now becoming dated. The core of what studies are generally done is those studies conducted to meet US FDA requirements. These are presented in Table 26.2. As will be discussed in Chapter 2, these guidelines are giving

TABLE 26.2 Synopsis of General Guidelines for Animal Toxicity Studies (US FDA, *Total Drug Quality*)

| Category | Duration of Human Administration ^a | Phase ^b | Subacute or Chronic Toxicity ^c | Special Studies |
|----------------------------------|---|---------------------------|--|--|
| Oral or parenteral | Several days (up to 3) Up to 2 weeks | I, II, III, NDA | 2 species: 2 weeks | For parenterally administered drugs; compatibility with blood and local tolerance at injection site where applicable |
| | | I | 2 species: 4 weeks | |
| | | II | 2 species: up to 4 weeks | |
| | | III, NDA | 2 species: up to 3 months | |
| | Up to 3 months | I, II | 2 species: 4 weeks | |
| | | III | 2 species: 3 months | |
| | | NDA | 2 species: up to 6 months | |
| | 6 months to unlimited | I, II | 2 species: 3 months | |
| | | III | 2 species: 6 months or longer | |
| | | NDA | 2 species: 12 months in rodents | |
| | | | 9 months in nonrodents | |
| | | | +2 rodent species for CA | |
| | | | 18 months (mouse)—may be met by use of a transgenic model | |
| | | | 24 months (rat) | |
| Inhalation (general anesthetics) | Single administration | I, II, III, NDA | 4 species: 5 days (3 h day ⁻¹) | |
| Dermal | Single application | I | 1 species: single 24 h exposure followed by 2-week observation | Sensitization |
| | Single or short-term application | II | 1 species: 20-day repeated exposure (intact and abraded skin) | |
| | Short-term application | III | As aforementioned | |
| | Unlimited application | NDA | As aforementioned, but intact skin study extended up to 6 months | |
| Ophthalmic | Single application | I | | Eye irritation tests with graded doses |
| | Multiple application | I, II, III | 1 species: 3 weeks, daily applications, as in clinical use | |
| | | NDA | 1 species: duration commensurate with period of drug administration | |
| Vaginal or rectal | Single application | I | | Local and systematic toxicity after vaginal or rectal application in two species |
| | Multiple application | I, II, III, NDA | 2 species: duration and no. of applications determined by the proposed use | |
| Drug combinations ^d | I | | | Lethality by appropriate route, compared to components run concurrently in one species |
| | II, III, NDA | 2 species: up to 3 months | | |

^aPhase I dosing of females if childbearing potential requires a segment II study in at least one species; phase III dosing of this population requires a segment I study and both segment II studies.

^bPhases I, II, and III are defined in Section 130.0 of the new drug regulations.

^cAcute toxicity should be determined in three species; subacute or chronic studies should be conducted by the route to be used clinically. Suitable mutagenicity studies should also be performed.

Observations:

| | | |
|-----------------------------------|----------------------------|---------------------------------|
| Body weights | Food consumption | Behavior |
| Metabolic studies | Ophthalmologic examination | Fasting blood sugar |
| Gross and microscopic examination | Hemogram | Liver and kidney function tests |
| Coagulation tests | Others as appropriate | |

^d Where toxicity data are available on each drug individually.

way to the ICH guidelines. However, while the length and details of studies have changed, the nature and order of studies remain the same.

The major variations in requirements for other countries still tend to be in the area of special studies. The United States does not formally require any genotoxicity studies, but common practice for US drug registration is to perform at least a bacterial gene mutation assay (Ames test), a mammalian cell mutation assay, and a clastogenicity assay, while Japan requires specific tests, including a gene mutation assay in *Escherichia coli*. Likewise, the European Economic Community (EEC) has a specified set of requirements, while individual countries have additional special requirements (e.g., Italy requires a mutagenicity assay in yeast). As detailed in Chapter 6, the new ICH genotoxicity guidelines have come to meet multinational requirements. Japan maintains a special requirement for an antigenicity assay in guinea pigs. The new safety pharmacology requirements are likely to be adopted over a period of time by different adherents.

It is possible to interact with the various regulatory agencies (particularly the FDA) when peculiarities of science or technology leave one with an unclear understanding of what testing is required. It is best if such discussions directly involve the scientists who understand the problems, and it is essential that the scientists at the FDA be approached with a course of action (along with its rationale) that has been proposed to the agency in advance.

The actual submissions to a regulatory agency that request permission either to initiate (or advance) clinical trials of a drug or to market a drug are not just bundles of reports on studies. Rather, they take the form of summaries that meet mandated requirements for format, accompanied by the reports discussed in these summaries. In the United States, these summaries are the appropriate section of the IND and NDA, which have now been harmonized in the ICH CTD format. The formats for these documents have recently been revised (FDA, 1987b). The EEC equivalent is the expert report, as presented in EEC Directive 75/319. Similar approaches are required by other countries. In each of these cases, textual summaries are accompanied by tables that also serve to summarize significant points of study design and of study findings.

What all of these approaches have in common is that they are to present integrated evaluations of the preclinical safety data that are available on a potential new drug. The individual studies and reports are to be tied together to present a single, cohesive overview of what is known about the safety of a drug.

Leber (1987) presents an excellent overview of the regulatory process involved in FDA oversight of drug development and gives the historical perspective for the evolution of the conservative process that is designed to ensure that any new pharmaceutical is both safe and efficacious.

There are other regulatory, legal, and ethical safety assessment requirements beyond those involved in the selection and marketing of a drug as a product entity. The actual drug product must be manufactured and transported in a safe manner, and any waste associated with this manufacture disposed of properly. Chapter 14 of this volume specifically addresses this often overlooked aspect of safety assessment programs.

26.3 ESSENTIAL ELEMENTS OF PROJECT MANAGEMENT

It is important to keep in mind that safety assessment is only one of many components involved in the discovery and development of new pharmaceuticals. The entire process has become enormously expensive, and completing the transit of a new drug from discovery to market has to be as efficient and expedition a process as possible. Even the narrow part of this process (safety assessment) is dependent on many separate efforts. Compounds must be made, analytical and bioanalytical methods developed, and dosage formulations developed, to name a few. One needs only to refer to Beyer (1978), Hamner (1990), Matoren (1984), Sneader (1986) (a good short overview), or Spilker (1994), Guarino (2009), or Blass (2015) for more details on this entire process and all of its components.

The coordination of this entire complex process is the province of project management, the objective of which is to ensure that all the necessary parts and components of a project match up. This discipline in its modern form was first developed for the Polaris missile project in the 1960s. Its major tool that is familiar to pharmaceutical scientists is the “network” or program evaluation and review technique (PERT) chart, as illustrated in Figure 26.2. This chart is a tool that allows one to see and coordinate the relationships between the different components of a project. One outcome of the development of such a network is the identification of the rate-limiting steps, which, in aggregate, comprise the critical path (see Table 26.3 for a lexicon of the terms used in project management).

A second graphic tool from project management is the Gantt chart, as illustrated in Figure 26.3. This chart allows one to visualize the efforts under way in any one area, such as safety assessment, for all projects that are currently being worked on.

Figure 26.4 shows a hybrid of the PERT and Gantt charts, designed to allow one to visualize all the resources involved in any one project.

As understanding of the key concepts of project management and their implications is critical of strategic planning and thinking for safety assessment, Kliem (1986) and Knutson (1980) offer excellent further reading in the area of project management.

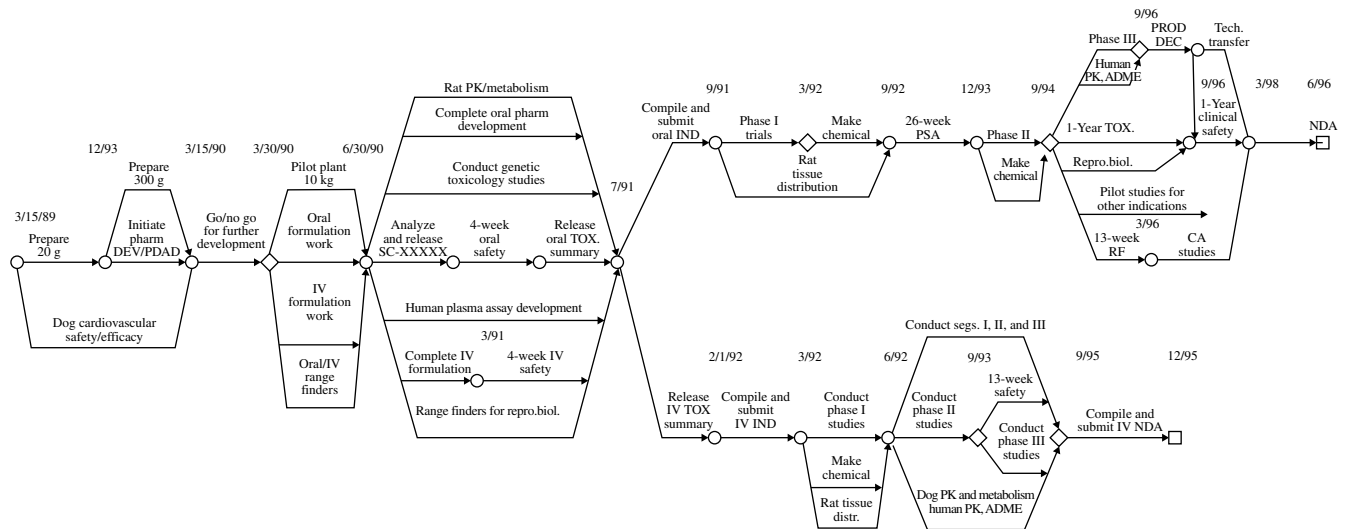


FIGURE 26.2 An example of program evaluation and review technique (PERT) chart of the development of a new pharmaceutical through to the filing of a new drug application (NDA). Circles are “nodes” indicating completion of activities. Diamonds are initiation points for tasks that have starting points independent of others. This “network” serves to illustrate the relationships between different activities and to evaluate effects of changes on project timing.

TABLE 26.3 Glossary of Project Management Terms

| | |
|------------------------------------|---|
| Activity | The work or effort needed to complete a particular event. It consumes time and resources |
| Average daily resource requirement | The likely amount of resources required to complete an activity or several activities on any workday during a project. The average daily labor requirement is one example |
| CPM | Acronym for critical path method. A network diagramming technique that places emphasis on time, cost, and the completion of events |
| Critical path | The longest route through a network that contains activities absolutely crucial to the completion of the project |
| Dummy arrow | A dashed line indicating an activity that uses no time or resources |
| Duration | The time it takes to complete an activity |
| Earliest finish | The earliest time an activity can be completed |
| Earliest start | The earliest an activity can begin if all activities before it are finished. It is the earliest time that an activity leaves its initiation node |
| Event | A synonym for node. A point in time that indicates the accomplishment of a milestone. It consumes neither time nor resources and is indicated whenever two or more arrows intersect |
| Free float | The amount of time that an activity can be delayed without affecting succeeding activities |
| Gantt chart | A bar chart indicating the time interval for each of the major phases of a project |
| Histogram | A synonym for bar chart |
| Latest finish | The latest time an activity can be completed without extending the length of a project |
| Latest start | The latest time an activity can begin without lengthening a project |
| Leveling | The process of “smoothing” put labor, material, and equipment requirements to facilitate resource allocation. The project manager accomplishes this by “rescheduling” noncritical activities so that the total resource requirements for a particular day match the average daily resource requirements |
| Most likely time | Used in PERT diagramming. The most realistic time estimate for completing an activity or project under normal conditions |
| Node | A synonym for event |
| Optimistic time | Used in PERT diagramming. The time the firm can complete an activity or project under the most ideal conditions |
| PERT | Acronym for program evaluation and review technique. A network diagramming technique that places emphasis on the completion of events rather than cost or time |
| Pessimistic time | Used in PERT diagramming. The time the firm can complete an activity or project under the worst conditions |
| Project | The overall work or effort being planned. It has only one beginning node and ending node. Between those nodes are countless activities and their respective nodes |
| Project phase | A major component, or segment, of a project. It is determined by the process known as project breakdown structuring |
| Total float | The total amount of flexibility in scheduling activities on a noncritical path. Hence, it provides the time an activity could be prolonged without extending a project’s final completion date |

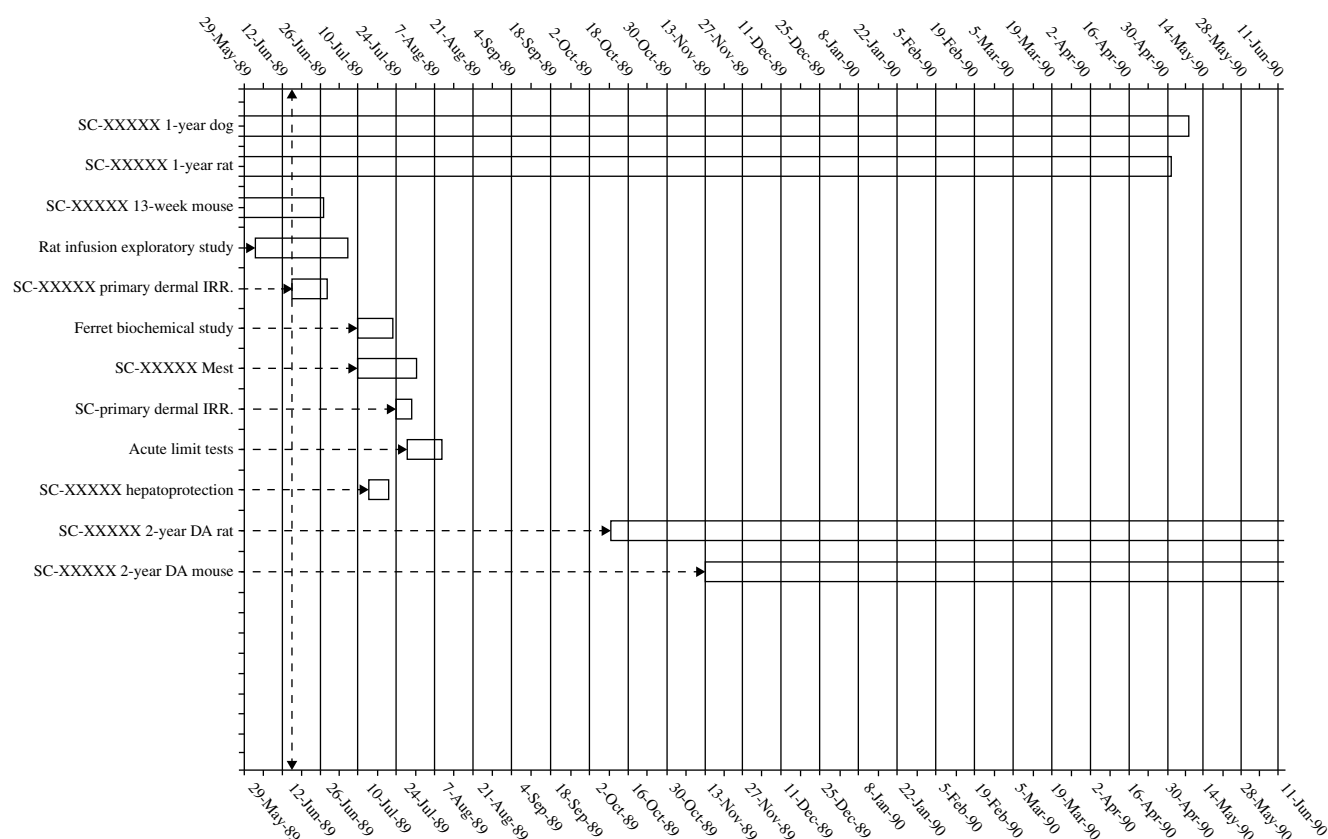


FIGURE 26.3 Gantt, or bar, chart showing scheduling of major safety assessment activities (studies) involved in pharmaceutical development project.

26.4 SCREENS: THEIR USE AND INTERPRETATION IN SAFETY ASSESSMENT

Much (perhaps even most) of what is performed in safety assessment can be considered screening—trying to determine if some effect is or is not (to an acceptable level of confidence) present (Zbinden et al., 1984). The general concepts of such screens are familiar to toxicologists in the pharmaceutical industry because the approach is a major part of the activities of the pharmacologists involved in the discovery of new compounds. But the principles underlying screening are not generally well recognized or understood. And such understanding is essential to the proper use, design, and analysis of screens (Gad, 1988, 1989a). Screens are the biological equivalent of exploratory data analysis (EDA) (Tukey, 1977).

Each test or assay has an associated activity criterion, that is, a level above which the activity of interest is judged to be present. If the result for a particular test compound meets this criterion, the compound may pass to the next stage. This criterion could be based on statistical significance (e.g., all compounds with observed activities significantly greater than the control at the 5% level could be tagged). However,

for early screens, such a formal criterion may be too strict, resulting in few compounds being identified as “active.”

A useful indicator of the efficacy of an assay series is the frequency of discovery of truly active compounds. The frequency is related to the probability of discovery and to the degree of risk (hazard to health) associated with an active compound passing a screen undetected. These two factors in turn depend on the distribution of activities in the series of compounds being tested and the chances of rejecting or accepting compounds with given activities at each stage.

Statistical modeling of the assay system may lead to the improvement of the design of the system by reducing the interval between discoveries of active compounds. The objectives behind a screen and considerations of (i) costs for producing compounds and testing and (ii) the degree of uncertainty about test performance will determine desired performance characteristics of specific cases. In the most common case of early toxicity screens performed to remove possible problem compounds, preliminary results suggest that it may be beneficial to increase the number of compounds tested, decrease the numbers of animals per group, and increase the range and number of doses. The result will be less information on more structure, but there will be an

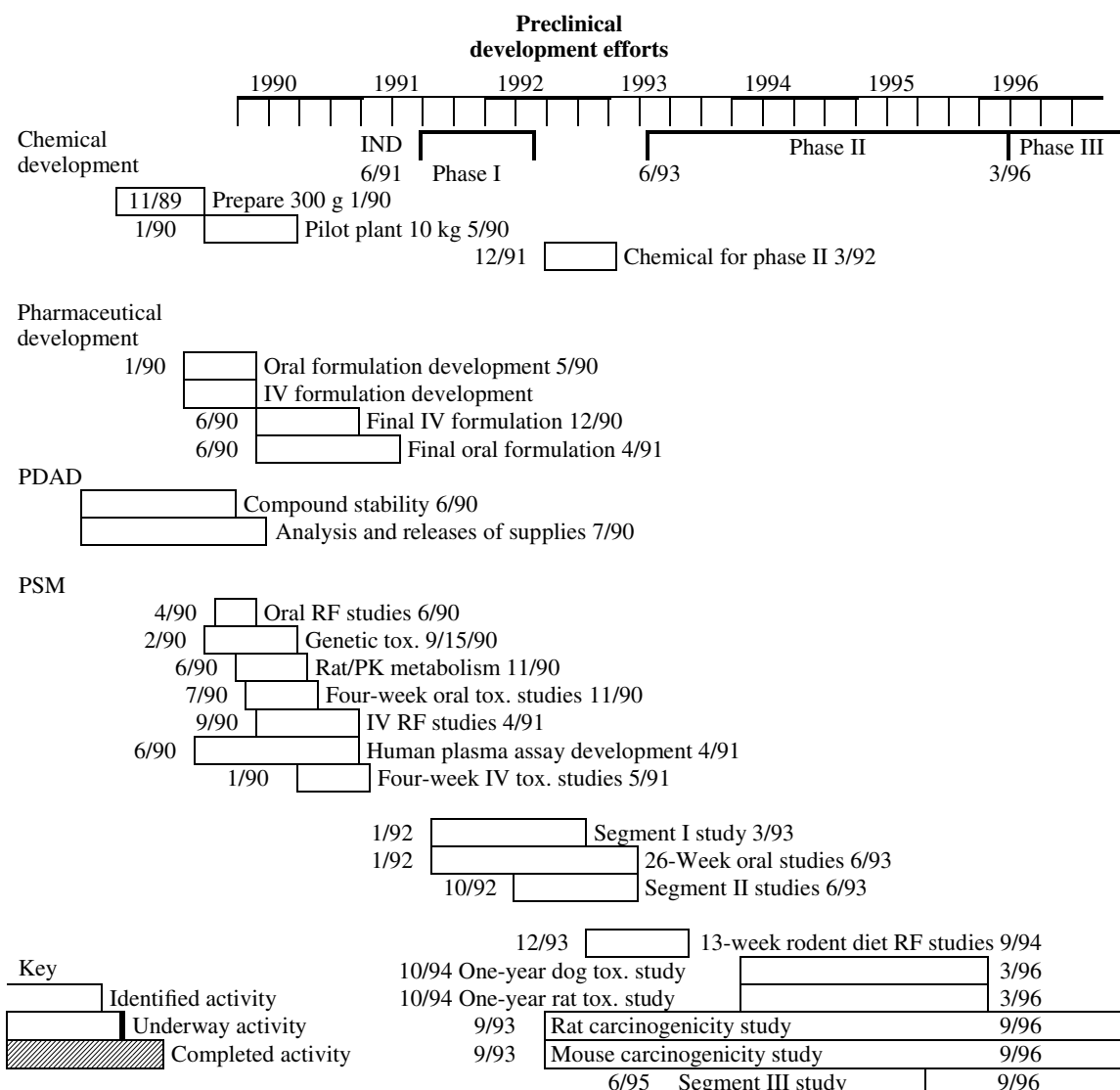


FIGURE 26.4 Hybrid project Gantt chart which identifies work of each development function ("line operation") in development of new compound and how it matches phase of development.

overall increase in the frequency of discovery of active compounds (assuming that truly active compounds are entering the system at a steady rate).

The methods described here are well suited to analyzing screening data when the interest is truly in detecting the absence of an effect with little chance of false negatives. There are many forms of graphical analysis methods available, including some newer forms that are particularly well suited to multivariate data (the type that is common in more complicated screening test designs). It is intended that these aspects of analysis will be focused on in a later publication.

The design of each assay and the choice of the activity criterion should, therefore, be adjusted, bearing in mind the relative costs of retaining false positives and rejecting false negatives. Decreasing the group sizes in the early assays

reduces the chance of obtaining significance at any particular level (such as 5%), so that the activity criterion must be relaxed, in a statistical sense, to allow more compounds through. At some stage, however, it becomes too expensive to continue screening many false positives, and the criteria must be tightened accordingly. Where the criteria are set depends on what acceptable noise levels in a screening system.

26.4.1 Characteristics of Screens

An excellent introduction to the characteristics of screens is Redman's (1981) interesting approach, which identifies four characteristics of an assay. Redman assumes that a compound is either active or inactive and that the proportion of activities in a compound can be estimated from past experience.

After testing, a compound will be classified as positive or negative (i.e., possessing or lacking activity). It is then possible to design the assay so as to optimize the following characteristics:

1. Sensitivity: The ratio of true positives to total activities
2. Specificity: The ratio of true negatives to total inactives
3. Positive accuracy: The ratio of true to observed positives
4. Negative accuracy: The ratio of true to observed negatives
5. Capacity: The number of compounds that can be evaluated
6. Reproducibility: The probability that a screen will produce the same result at another time (and, perhaps, in some other lab)

An advantage of testing many compounds is that it gives the opportunity to average activity evidence over structural classes or to study quantitative structure–activity relationships (QSARs). QSARs can be used to predict the activity of new compounds and thus reduce the chance of *in vivo* testing on negative compounds. The use of QSARs can increase the proportion of truly active compounds passing through the system.

To simplify this presentation, data sets drawn only from neuromuscular screening activity were used. However, the evaluation and approaches should be valid for all similar screening data sets, regardless of source. The methods are not sensitive to the biases introduced by the degree of interdependence found in many screening batteries that use multiple measures (such as the neurobehavioral screen):

1. Screens almost always focus on detecting a single end point of effect (such as mutagenicity, lethality, neurotoxicity, or development toxicity) and have a particular set of operating characteristics in common.
2. A large number of compounds are evaluated, so ease and speed of performance (which may also be considered efficiency) are very desirable characteristics.
3. The screen must be very sensitive in its detection of potential effective agents. An absolute minimum of active agents should escape detection; that is, there should be very few false negatives (in other words, the type II error rate or beta level should be low). Stated yet another way, the signal gain should be way up.
4. It is desirable that the number of false positives be small (i.e., there should be a low type I error rate or alpha level).
5. Items 2–4, which are all to some degree contradictory, require the involved researchers to agree on a set of compromises, starting with the acceptance of a relatively high alpha level (0.10 or more), that is, an increased noise level.

6. In an effort to better serve Item 2, safety assessment screens are frequently performed in batteries so that multiple end points are measured in the same operation. Additionally, such measurements may be repeated over a period of time in each model as a means of supporting Item 3.
7. This screen should use small amounts of compound to make Item 1 possible and should allow evaluation of materials that have limited availability (such as novel compounds) early on in development.
8. Any screening system should be validated initially using a set of blind (positive and negative) controls. These blind controls should also be evaluated in the screening system on a regular basis to ensure continuing proper operation of the screen. As such, the analysis techniques used here can then be used to ensure the quality or modify performance of a screening system.
9. The more that is known about the activity of interest, the more specific the form of screen that can be employed. As specificity increases, so should sensitivity.
10. Sample (group) sizes are generally small.
11. The data tend to be imprecisely gathered (often because researchers are unsure of what they are looking for) and therefore possess extreme within-group variability. Control and historical data are not used to adjust for variability or modify test performance.
12. Proper dose selection is essential for effective and efficient screen design and conduct. If insufficient data are available, a suitably broad range of doses must be evaluated (however, this technique is undesirable on multiple grounds, as has already been pointed out).

The design, use, and analysis of screens are covered in detail in Chapter 4 of this volume.

26.5 STRATEGY AND PHASING

Regulatory requirements and our understanding of the pharmacology, marketing, and clinical objectives for a potential product provide a framework of requirements for the safety assessment of potential new pharmaceuticals. How one meets these requirements is not fixed, however. Rather, exactly what is done and when are reflections of the philosophy and managerial climate of the organization that is doing the discovery and development. It should be kept in mind that establishing and maintaining an excellent information based on the biological basis for a compound's expected therapeutic activity and safety is essential but often left undone. This subject is addressed in Chapter 2 of this volume.

There are multiple phases involved in the safety assessment portion of the discovery, development, and marketing process.

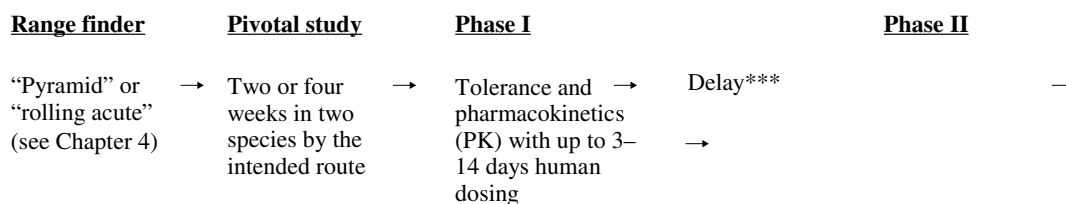
The actual conduct of the studies in each phase forms the basis of the bulk of the chapters in this book. However, unless the pieces are coordinated well and utilized effectively (and completed at the right times), success of the safety assessment program is unlikely or very expensive.

First, support needs to be given to basic research (also called discovery, biology, or pharmacology in different organizations) so that it can efficiently produce a stream of potential new product compounds with as few overt toxicity concerns as possible. This means that there must be early and regular interaction between the individuals involved and that safety assessment must provide screening services to rank the specific safety concerns of the compounds. These screens may be *in vitro* (both for genetic and nongenetic end points) or *in vivo* (designed on purpose for a single end point, such as effects on reproductive performance, promotion activity, etc.). There must also be an ongoing work to elucidate the mechanisms and structure–activity relationships behind those toxicities that are identified (Gad, 1989b).

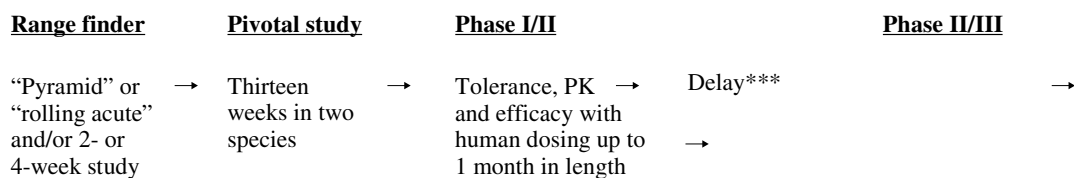
Second, it is the traditional core of safety assessment that is viewed as development. Development includes providing

the studies to support compounds getting into the clinic (an IND application being filed and accepted); evaluating a compound to the point at which it is considered safe, able to be absorbed, and effective (clinical phase II); and, finally, registration (filing an NDA and having it approved). Various organizations break this process up differently. Judgments are generally made on the likelihood of compounds failing (“dying”) at different stages in the clinical development process, and the phasing of preclinical support is selected and/or adjusted accordingly. If an organization has a history of many compounds failing early in the clinic (such as in the initial phase I tolerance trials, where there may be only 3–10 days of human dosing), then initial “pivotal” preclinical studies are likely to be only 4-week-long studies. If compounds tend to fail only in longer efficacy trials, then it is more efficient to run longer initial preclinical trials. Figure 26.5 shows several variations on these approaches. Additionally, the degree of risk involved in study design (particularly in dose selection) is also an organizational characteristic. Pivotal studies can fail on two counts associated with dose selection. Either they cannot identify a “safe”

Plan 1: Clinical decision point** is short-term tolerance or human pharmacokinetics



Plan 2: Clinical decision point is an indication of efficacy in man



Plan 3: Plan for success or resources are not a constraint

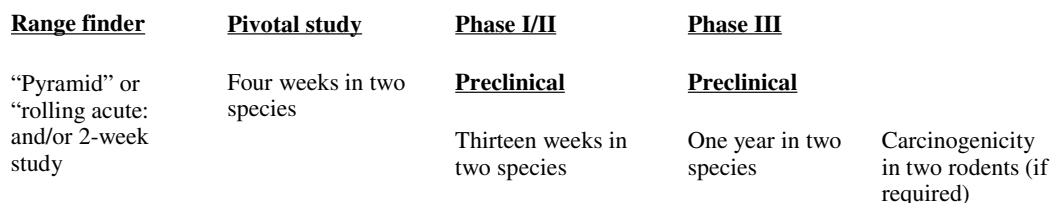


FIGURE 26.5 Three different approaches to matching preclinical safety efforts to support clinical development of new drug. Which is the best one for any specific case depends on the considerations of resource availability and organizational tolerance of “risk.” In plan 1, little effort will be “wasted” on projects that fail during early (phase I) clinical trials, but if phase I trials are successful, there will be major delays. In plan 3, clinical development will never be held up waiting for more safety work, but a lot of effort will go into projects that never get past phase I. Plan 2 is a compromise. Delays are to allow additional preclinical (animal safety) studies to support longer clinical trials in accordance with FDA or other applicable guidelines.

(no-effect) dose, or they can neglect to find a dose that demonstrates a toxic effect (and therefore allows identification of potential target organs). Therefore, picking the doses for such studies is an art that has been risky because, traditionally, only three different dose groups have been used, and before clinical trials are conducted, there is at best a guess as to what clinical dose will need to be cleared. The use of four (or five) dose groups only marginally increases study cost and, in those cases where the uncertainty around dose selection is great, provides a low-cost alternative to repeating the study.

Pivotal studies can also be called *shotgun tests*, because it is unknown in advance what end points are being aimed at. Rather, the purpose of the study is to identify and quantitate all potential systemic effects resulting from a single exposure to a compound. Once known, specific target organ effects can then be studied in detail if so desired. Accordingly, the generalized design of these studies is to expose groups of animals to controlled amounts or concentrations of the material of interest and then to observe for a measure as many parameters as practical over a period past or during the exposure. Further classification of tests within this category would be the route by which test animals are exposed/dosed or by the length of dosing. "Acute," for example, implies a single exposure interval (of 24h or less) or dose of test material. Using the second scheme (length of dosing), the objectives of the successive sets of pivotal studies could be defined as follows:

Acute or dose range-finding (DRF) studies:

1. Set maximum doses for next studies.
2. Identify very or unusually toxic agents.
3. Estimate upper limit of tolerability.
4. Identify organ system affected.

Either 2-week studies:

1. Set doses for next studies.
2. Identify organ toxicity.
3. Identify very or unusually toxic agents.
4. Estimate lethality potential.
5. Evaluate potential for accumulation of effects.
6. Get estimate of kinetic properties (blood sampling/urine sampling).

Or 4-week studies:

1. Set doses for next studies.
2. Identify organ toxicity.
3. Identify very or unusually toxic agents.
4. Estimate lethality potential.
5. Evaluate potential for accumulation of effects.

6. Get estimate of kinetic properties (blood sampling/urine sampling).
7. Elucidate nature of specific types of target organ toxicities induced by repeated exposure.

Thirteen-week studies (now commonly included as an interim necropsy and report in chronic studies):

1. Set doses for next studies.
2. Identify organ toxicity.
3. Identify very or unusually toxic agents.
4. Evaluate potential for accumulation of effects.
5. Evaluate pharmacokinetic properties.
6. Elucidate nature of specific types of target organ toxicities induced by repeated exposure.
7. Evaluate reversibility of toxic effects.

Chronic studies:

1. Elucidate nature of specific types of target organ toxicities induced by prolonged repeated exposure.
2. Identify potential carcinogens.

The problems of scheduling and sequencing toxicology studies and entire testing programs have been minimally addressed in print. Though there are several books and many articles available that address the question of scheduling multiple tasks in a service organization (French, 1982) and an extremely large literature on project management (as briefly overviewed earlier in this chapter), no literature specific to a research testing organization exists.

For all the literature on project management, however, a review will quickly establish that it does not address the rather numerous details that affect study/program scheduling and management. There is, in fact, to my knowledge, only a single article (Levy et al., 1977) in the literature that addresses scheduling, and it describes a computerized scheduling system for single studies.

There are commercial computer packages available for handling the network construction, interactions, and calculations involved in what, as will be shown later, is a complicated process. These packages are available for use on both mainframe and microcomputer systems.

Scheduling for the single study case is relatively simple. One should begin with the length of the actual study and then factor in the time needed before the study is started to secure the following resources:

- Animals must be on hand and properly acclimated (usually for at least 2 weeks prior to the start of the study).
- Vivarium space, caging, and animal care support must be available.

- Technical support for any special measurements such as necropsy, hematology, urinalysis, and clinical chemistry must be available on the dates specified in the protocol.
- Necessary and sufficient test material must be on hand.
- A formal written protocol suitable to fill regulatory requirements must be on hand and signed.

The actual study (from first dosing or exposure of animals to the last observation and termination of the animals) is called the in-life phase, and many people assume the length of the in-life phase defines the length of a study. Rather, a study is not truly completed until any samples (blood, urine, and tissue) are analyzed, slides are prepared and microscopically evaluated, data are statistically analyzed, and a report is written, proofed, and signed off. Roll all of this together, and if you are conducting a single study under contract in an outside laboratory, an estimate of the least time involved in its completion should be equal to (other than in the case of an acute or single and point study) no more than

$$L + 6\text{weeks} + \frac{1}{2} L$$

where L is the length of the study. If the study is a single end-point study and does not involve pathology, then the least time can be shortened to $L + 6$ weeks. In general, the best that can be done is $L + 10$ weeks.

When one is scheduling out an entire testing program on contract, it should be noted that if multiple tiers of tests are to be performed (such as an acute, 2-week, 13-week, and lifetime studies), then these must be conducted sequentially, as the answer from each study in the series defines the design and sets the doses for the subsequent study.

If, instead of contracting out, one is concerned with managing a testing laboratory, then the situation is considerably more complex. The factors and activities involved are outlined later. Within these steps are rate-limiting factors that are invariably due to some critical point or pathway. Identification of such critical factors is one of the first steps for a manager to take to establish effective control over either a facility or program.

Before any study is actually initiated, a number of prestudy activities must occur (and, therefore, these activities are currently under way—to one extent or another—for the studies not yet under way but already authorized or planned for this year for any laboratory):

- Test material procurement and characterization

- Development of formulation and dosage forms for study
- If inhalation study, development of generation and analysis methodology, chamber trials, and verification of proper chamber distribution

- Development and implementation of necessary safety steps to protect involved laboratory personnel

- Arrangement for waste disposal
- Scheduling to assure availability of animal rooms, manpower, equipment, and support services (pathology and clinical)
- Preparation of protocols
- Animal procurement, health surveillance, and quarantine
- Preparation of data forms and books
- Conduct of prestudy measurements on study animals to set baseline rates of body weight gain and clinical chemistry values

After completion of the in-life phase (i.e., the period during which live animals are used) of any study, significant additional effort is still required to complete the research. This effort includes the following:

- Preparation of data forms and books, preparation of tissue slides, and microscopic evaluation of these slides
- Preparation of data tables
- Statistical analysis of data
- Preparation of reports

There are a number of devices available to a manager to help improve the performance of a laboratory involved in these activities. One such device (cross-training) is generally applicable enough to be particularly attractive.

Identification of rate-limiting steps in a toxicology laboratory over a period of time usually reveals that at least some of these are variable (almost with the season). At times, there is too much work of one kind (say, inhalation studies) and too little of another (say, dietary studies). The available staff for inhalation studies cannot handle this peak load, and since the skills of these two groups are somewhat different, the dietary staff (which is now not fully occupied) cannot simply relocate down the hall and help out. However, if, early on, one identifies low- and medium-skill aspects of the work involved in inhalation studies, one could cross-train the dietary staff at a convenient time so that it could be redeployed to meet peak loads.

It should be kept in mind that there are a number of common mistakes (in both the design and conduct of studies and in how information from studies is used) that have led to unfortunate results, ranging from losses in time and money and the discarding of perfectly good potential products to serious threats to people's health. Such outcomes are indeed the great disasters in product safety assessment—especially since many of them are avoidable if attention is paid to a few basic principles.

It is quite possible to design a study for failure. Common shortfalls include:

1. Using the wrong animal model.
2. Using the wrong route or dosing regimen.

3. Using the wrong vehicle or formulation of test material.
4. Using the wrong dose level. In studies where several dose levels are studied, the worst outcome is to have an effect at the lowest dose level tested (i.e., the safe dosage in animals remains unknown). The next worst outcome is to have no effect at the highest dose tested (generally meaning that the signs of toxicity remain unknown, invalidating the study in the eyes of many regulatory agencies).
5. Making leaps of faith. An example is to set dosage levels based on others' data and to then dose all test animals. At the end of the day, all animals in all dose levels are dead. The study is over; the problem remains.
6. Using the wrong concentration of test materials in a study. Many effects (e.g., both dermal and gastrointestinal irritation) are very concentration dependent.
7. Failing to include a recovery (or rebound) group. If one finds an effect in a 90-day study (say, gastric hyperplasia), how does one interpret it? How does one respond to the regulatory question, "Will it progress to cancer?" If an additional group of animals were included in dosing and then were maintained for a month after dosing had been completed, recovery (reversibility) could be both evaluated and (if present) demonstrated.

Additionally, there are specialized studies designed to address end points of concern for almost all drugs (carcinogenicity, reproductive, or developmental toxicity) or concerns specific to a compound or family of compounds (e.g., local irritation, neurotoxicity, or immunotoxicity). When these are done, timing also requires careful consideration. It must always be kept in mind that the intention is to ensure the safety of people in whom the drug is to be evaluated (clinical trials) or used therapeutically. An understanding of special concerns for both populations should be considered essential.

Safety evaluation does not cease being an essential element in the success of the pharmaceutical industry once a product is on the market. It is also essential to support marketed products and ensure that their use is not only effective but also safe and unclouded by unfounded perceptions of safety problems. This requires not only that clinical trials be monitored during development (Spector et al., 1988) but also that experience in the marketplace be monitored.

The design and conduct of safety assessment studies and programs also require an understanding of some basic concepts:

1. The studies are performed to establish or deny the safety of a compound rather than to characterize the toxicity of a compound.

2. Because pharmaceuticals are intended to affect the functioning of biological systems and safety assessment characterizes the effects of higher-than-therapeutic doses of compounds, it is essential that one be able to differentiate between hyperpharmacology and true (undesirable) adverse effects.
3. Focus of the development process for a new pharmaceutical is an essential aspect of success but is also difficult to maintain. Clinical research units generally desire to pursue as many or as broad claims as possible for a new agent and frequently also apply pressure for the development of multiple forms for administration by different routes. These forces must be resisted because they vastly increase the work involved in safety assessment and they may also produce results (in one route) that cloud evaluation (and impede Institutional Review Board (IRB) and regulatory approval) of the route of main interest.

26.6 CRITICAL CONSIDERATIONS

In general, what the management of a pharmaceutical development enterprise wants to know at the beginning of a project are three things: what are the risks (and how big are they), how long will it take, and how much (money and test compound) will it take?

The risks question is beyond the scope of this volume. The time question was addressed earlier in this chapter. How much money is also beyond the scope of this volume. But calculating projected compound needs for studies is a fine challenge in the design and conduct of a safety evaluation program. The basic calculation is simple. The amount needed for a study is equal to

$$N W I L D$$

where

N = the number of animals per group

W = the mean weight per animal during the course of the study (in kg)

I = the total number of doses to be delivered (such as in a 28-day study, 28 consecutive doses)

L = a loss or efficiency factor (to allow for losses in formulation and dose delivery, a 10% factor is commonly employed, meaning a value of 1.1 is utilized)

D = the total dose factor (This is the sum of all the dose levels. For example, if the groups are to receive 1000, 300, 100, and 30 mg kg⁻¹, then the total dose factor is 1000 + 300 + 100 + 30 or 1430 mg kg⁻¹.)

As an example, let's take a 28-day study in rats where there are 10 males and 10 females per group and the dose

levels employed at 1000, 3000, 100, and 30 mg kg⁻¹. Over the course of the 28 days the average weight of the rats is likely to be 300 g (or 0.3 kg). This means our values are:

$$N=20$$

$$W=0.3 \text{ kg}$$

$$I=28$$

$$L=1.1$$

$$D=1430 \text{ mg kg}^{-1}$$

and therefore our total compound needs will be (20)(0.3)(28)(1.1)(1430 mg)=2642.64 mg or 2.642 g.

This is the simplest case but shows the principles.

A governing principle of pharmaceutical safety assessment is the determination of safety factors—the ratio between the therapeutic dose (that which achieves the desired therapeutic effect) and the highest dose which evokes no toxicity. This grows yet more complex (but has less uncertainty) if one bases these ratios on plasma levels rather than administered doses. Traditionally based on beliefs as to differences of species sensitivity, it has been held that a minimum of a fivefold (5×) safety factor should be observed based on toxicity findings in nonrodents and a 10-fold (10×) based on rodents.

The desire to achieve at least such minimal therapeutic indices and to also identify levels associated with toxicity (and the associated toxic effects) forms the basis of dose selection for systemic (and most other *in vivo*) toxicity studies.

26.7 SPECIAL CASES IN SAFETY ASSESSMENT

It may seem that the course of preclinical safety assessment (and of other aspects of development) of a pharmaceutical is a relatively linear and well-marked route, within some limits. This is generally the case—but not always. There are a number of special cases where the pattern and phasing of development (and of what is required for safety assessment) do not fit the usual pattern. Four of these cases are:

1. When the drug is intended to treat a life-threatening disease, such as AIDS
2. When the drug is actually a combination of two previously existing drug entities
3. When the drug actually consists of two or more isomers
4. When the drug is a peptide produced by a biotechnology process

Drugs intended to treat a life-threatening disease for which there is no effective treatment are generally evaluated against less rigorous standards of safety when making decisions about advancing them into and through clinical

testing. This acceptance of increased risk (moderated by the fact that the individuals involved will die if not treated at all) is balanced against the potential benefit. These changes in standards usually mean that the phasing of testing is shifted: animal safety studies may be done in parallel or (in the case of chronic and carcinogenicity studies) after clinical trials and commercialization. But the same work must still eventually be performed.

Combination drugs, at least in terms of safety studies up to carcinogenicity studies, are considered by regulatory agencies as new drug entities and must be so evaluated. The accordingly required safety tests must be performed on a mixture with the same ratio of components as is to be a product. Any significant change in ratios of active components means one is again evaluating, in regulatory eyes, a new drug entity.

Now that it is possible to produce drugs that have multiple isomers in the form of single isomers (as opposed to racemic mixtures), for good historical reasons, regulatory agencies are requiring at least some data to support any decision to develop the mixture as opposed to a single isomer. One must, at a minimum, establish that the isomers are of generally equivalent therapeutic activity and, if there is therapeutic equivalence, that any undesirable biological activity is not present to a greater degree in one isomer or another.

26.8 SUMMARY

It is the belief of this author that the entire safety assessment process that supports pharmaceutical research and development is a multistage process of which no single element is overwhelmingly complex. These elements must be coordinated and their timing and employment carefully considered on a repeated basis. Focus on the objective of the process, including a clear definition of the questions being addressed by each study, is essential, as is the full integration of the technical talents of each of the many disciplines involved. A firm understanding of the planned clinical development of the drug is essential. To stay competitive requires that new technologies be identified and incorporated effectively into safety assessment programs as they become available. It is hoped that this volume will provide the essential knowledge of the key elements to allow these goals to be realized.

REFERENCES

- Arnst C (1998) Danger: Read the Label. *Business Week*, May 11:100.
- Beyer K. (1978) Discovery, development and delivery of new drugs. In: *Monographs in Pharmacology and Physiology*, No. 12. Spectrum, New York.

- Blass BE. (2015) *Basic Principles of Drug Discovery and Development*. Academic Press, San Diego, CA.
- FDA (Food and Drug Administration) (1987a) *Good Laboratory Practice Regulations: Final Rule. 21 CFR Part 58, Federal Register*, September 4, 1987. Food and Drug Administration (FDA), Silver Spring, MD.
- FDA (Food and Drug Administration) (1987b) *New Drug, Antibiotic, and Biologic Drug Produce Regulations. 21 CFR Parts 312, 314, 511, and 514, Federal Register*, 52 (53): 8798–8857.
- Fishlock D (1990) Survival of the fittest drugs. *Financial Times*, April 24:16–17.
- French S. (1982) *Sequencing and Scheduling*. Halsted Press, New York.
- Gad SC. (1988) An approach to the design and analysis of screening studies in toxicology. *J. Am. Coll. Toxicol.* 7(2): 127–138.
- Gad SC. (1989a) Principles of screening in toxicology: with special emphasis on applications to neurotoxicology. *J. Am. Coll. Toxicol.* 8(1):21–27.
- Gad SC. (1989b) A tier testing strategy incorporating *in vitro* testing methods for pharmaceutical safety assessment. *Hum. Innov. Alter. Anim. Exp.* 3:75–79.
- Gad SC. (1996) Preclinical toxicity testing in the development of new therapeutic agents. *Scand. J. Lab. Anim. Sci.* 23: 299–314.
- Gad SC. (2000) *Product Safety Evaluation Handbook*, 2nd Ed. Marcel Dekker, New York.
- Gad SC. (2001) *In Vitro Toxicology*, 2nd Ed. Taylor & Francis, Philadelphia, PA.
- Gad SC. (2009) *Drug Safety Evaluation*, 2nd Ed. John Wiley & Sons, Inc, Hoboken, NJ.
- Gad SC. (2010) *International Regulatory Safety Evaluation of Pharmaceuticals and Medical Devices*. Springer, Berlin.
- Guarino RA. (2009) *New Drug Approval Process*, 5th Ed. Informa, New York.
- Hamner CE. (1990) *Drug Development*, 2nd Ed. CRC Press, Boca Raton, FL.
- ICH (2009) *M3(R2) Non-clinical Safety Studies for the Conduct of Human Clinical Trials for Pharmaceuticals*. International Conference on Harmonisation, Geneva, Switzerland.
- Kliem RL. (1986) *The Secrets of Successful Project Management*. John Wiley & Sons, Inc, New York.
- Knutson JR. (1980) *How to Be a Successful Project Manager*. American Management Associations, New York.
- Lazaron J, Parmeranz BH, Corey PN. (1998) Incidence of adverse drug reactions in hospitalized patients. *J. Am. Med. Assoc.* 279:1200–1209.
- Leber P. (1987) FDA: The federal regulations of drug development. In: *Psychopharmacology: The Third Generation of Progress* (Meltzer HY, Ed.). Raven Press, New York, pp. 1675–1683.
- Lesko LJ, Rowland M, Peck CC, Blaschke TF. (2000) Optimizing the science of drug development: opportunities for better candidate selection and accelerated evaluation in humans. *Pharm. Res.* 17:1335–1344.
- Levy AE, Simon RC, Beerman TH, Fold RM. (1977) Scheduling of toxicology protocol studies. *Comput. Biomed. Res.* 10:139–151.
- Littlehales C (1999) The price of a pill. *Modern Drug Discovery*, Jan/Feb 1999:21–30.
- Matoren GM. (1984) *The Clinical Process in the Pharmaceutical Industry*. Marcel Dekker, New York, pp. 273–284.
- Moore TJ, Patsy D, Furnberg J. (1998) Time to act on drug safety. *J. Am. Med. Assoc.* 279:1971–6.
- O'Grady J, Linet OI. (1990) *Early Phase Drug Evaluation in Man*. CRC Press, Boca Raton, FL.
- Raeburn P (1999) Drug safety needs a second opinion. *Business Week*, September 20:72–74.
- Redman C. (1981) Screening compounds for clinically active drugs. In: *Statistics in the Pharmaceutical Industry* (Buncher CR and Tsay J, Eds.). Marcel Dekker, New York, pp. 19–42.
- Scheuplein RJ, Shoal SE, Brown RN. (1990) Role of pharmacokinetics in safety evaluation and regulatory considerations. *Annu. Rev. Pharmacol. Toxicol.* 30:197–218.
- Smith CG. (1992) *The Process of New Drug Discovery and Development*. CRC Press, Boca Raton, FL.
- Sneader W. (1986) *Drug Development: From Laboratory to Clinic*. John Wiley & Sons, Inc, New York.
- Spector R, Park GD, Johnson GF, Vessell ES. (1988) Therapeutic drug monitoring. *Clin. Pharmacol. Ther.* 43:345–353.
- Spilker B. (1994) *Multinational Pharmaceutical Companies*, 2nd Ed. Raven Press, New York.
- Traina VM. (1983) The role of toxicology in drug research and development. *Med. Res. Rev.* 3:43–72.
- Tukey JW. (1977) *Exploratory Data Analysis*. Addison-Wesley, Reading, MA.
- Zbinden G. (1992) *The Source of the River Po*. Haag and Herschen, Frankfurt.
- Zbinden G, Elsner J, Boelsterli UA. (1984) Toxicological screening. *Regul. Toxicol. Pharmacol.* 4:275–286.

THE APPLICATION OF *IN VITRO* TECHNIQUES IN DRUG SAFETY ASSESSMENT

27.1 INTRODUCTION

The key assumptions underlying modern toxicology are (Zbinden, 1987) (i) that other organisms can serve as accurate predictive models of toxicity in man, (Gad, 1996, 2015) (ii) that selection of an appropriate model to use is the key to accurate prediction of potential hazard in man, and (Lijinsky, 1988) (iii) that understanding the strengths and weaknesses of any particular model is essential to translating potential hazards identified in these models to assess relevant hazard in man and in the subsequent management of actual risks. Historically higher animals (mammals) have been used in studies with significant focus on collecting descriptive data (reinforced by being able to translate methods such as clinical chemistry and pathology from human medicine). This approach to model selection and data collection in toxicological research became the subject of critical scientific review starting in the early 1980s. Usually in toxicology, when we refer to “models,” we really have meant test organisms or systems, although, in fact, the manners in which parameters are measured (and which parameters are measured to characterize an end point of interest) are also critical parts of the model (or, indeed, may actually constitute the “model”).

Although there have been accepted principles for test organism selection, these have not generally been the actual final basis for such selection. It is a fundamental hypothesis of both historical and modern toxicology that adverse effects caused by chemical entities in higher animals are generally the same as those induced by those entities in man. There are many who point to individual exceptions to this and conclude that the general principle is false. Yet, as our understanding of molecular biology advances and we learn more about the similarities of the structure and function of higher

organisms at the molecular level, the more it becomes clear that the mechanisms of chemical toxicity are largely identical in all higher life forms, including humans. The target sites are molecular, and differences in responses are all about similarities in receptor populations, in receptor population distribution in organ systems, and in the manner and means of getting toxicophores to these sites or preventing them from reaching these sites. In this sense, it is now the age of translational toxicology. This increased understanding has caused some of the same people who question the general principle of predictive value to in turn suggest that our state of knowledge is such that mathematical models or simple cell culture systems could be used just as well as intact higher animals to predict toxicities in man. This last suggestion has unfortunately missed the point that the final expressions of toxicity in man or animals are frequently the summations of extensive and complex interactions on cellular, biochemical, and molecular (even gene) levels. Zbinden (1987) and Gad (1996) published extensively in this area, including a very advanced defense of the value of animal models. Lijinsky (1988) has reviewed the specific issues about the predictive value and importance of animals in carcinogenicity testing and research. Although it was once widely believed, and may still be believed by many animal rights activists, that *in vitro* mutagenicity tests would entirely replace animal bioassays for carcinogenicity, this is clearly not the case on scientific, public health, or regulatory grounds (despite the limitations of the current bioassay models). Although there are differences in the responses of various species (including man) to carcinogens, the overall predictive value of such results, when tempered by judgment, is clear. At the same time, a well-reasoned use of *in vitro* or other alternative test model systems is essential to the continued development of a product safety assessment

program that is effective, efficient, and relevant to human safety (Gad, 1990a, 2000, 2009, 2015).

The subject of intact animal models and their proper selection and use have been addressed elsewhere (Gad, 2015) and will not be further addressed here. However, alternative models which use other than intact higher organisms are seeing increasing use in toxicology for a number of reasons.

The first and most significant factor behind the interest in the so-called *in vitro* systems has clearly been philosophical and political—an unremitting campaign by a wide spectrum of individuals concerned with the welfare and humane treatment of laboratory animals (Gad, 2009)—though some are also clearly simply antiscience and antitechnology. In 1959 Russell and Burch first proposed what have come to be called the three Rs of humane animal use in research—replacement, reduction, and refinement. These have served as the conceptual basis for reconsideration of animal use in research. Efforts continue along these lines (Singer, 1975; May et al., 2009; Birnbaum and Stokes, 2010).

Replacement means utilizing methods that do not use intact animals in place of those that do. For example, veterinary students may use a canine cardiopulmonary resuscitation simulator, Resusci-Dog, instead of living dogs; cell cultures may replace mice and rats that are fed new products to discover substances poisonous to humans. In addition, using the preceding definition of animal, an invertebrate (e.g., a horseshoe crab) could replace a vertebrate (e.g., a rabbit) in a testing protocol.

Reduction refers to the use of fewer animals. These are wide variations in estimates of animals in research, ranging as high as 127 million in a year (Boo and Knight, 2009). For instance, changing practices allows toxicologists to estimate the lethal dose of a chemical with as few as one-tenth the number of animals used in traditional tests. In biomedical research, long-lived animals, such as primates, may be used in multiple sequential protocols, assuming that they are not deemed inhumane or scientifically conflicting. Designing experimental protocols with appropriate attention to statistical inference can lead to decreases or to increases in the number of animals used. Through coordination of efforts among investigators, several tissues may be simultaneously taken from a single animal. Reduction can also refer to the minimization of any unintentionally duplicative experiments, perhaps through improvements in information resources.

Refinement entails the modification of existing procedures so that animals are subjected to less pain and distress. Refinements may include administration of anesthetics to animals undergoing otherwise painful procedures, administration of tranquilizers for distress, humane destruction prior to recovery from surgical anesthesia, and careful scrutiny of behavioral indices of pain or distress, followed by cessation of the procedure or the use of appropriate analgesics. Refinements also include the enhanced use of

noninvasive imaging technologies that allow earlier detection of tumors, organ deterioration, or metabolic changes and the subsequent early euthanasia of test animals.

Progress toward these first three Rs has been previously reviewed (Gad, 1990b; Salem, 1995; Salem and Katz, 1998; Indans, 2002; Gribaldo, 2007). However, there is a fourth R—responsibility—which was not in Russell and Burch's initial proposal (Russell and Burch, 1959). To toxicologists this is the cardinal R. They may be personally committed to minimizing animal use and suffering and to doing the best possible science of which they are capable, but at the end of it all, toxicologists must stand by their responsibility to be conservative in ensuring the safety of the people using or exposed to the drugs and chemicals produced by our society.

Since 1980, issues of animal use and care in toxicological research and testing have become one of the fundamental concerns of both science and the public. Are our results predictive of what may or may not be seen in man? Are we using too many animals, and are we using them in a manner that gets the answer we need with as little discomfort to the animals as possible? How do we balance the needs of man against the welfare of animals?

During the same time frame, interest and progress in the development of *in vitro* test systems for toxicity evaluations have also progressed. Early reviews by Hooisma (1982), Neubert (1982), and Williams et al. (1983) record the proceedings of conferences on the subject, but Rofe's (1971) review was the first found by this author. Although it is hoped that in the long term some of these (or other) *in vitro* methods will serve as definitive tests in place of those that use intact animals, at present it appears more likely that their use in most cases will be as screens. Frazier (1992) and Gad and Chengelis (1997) give recent overviews of the general concepts and status of *in vitro* alternatives.

The entire product safety assessment process, in the broadest sense, is a multistage process in which none of the individual steps are overwhelmingly complex, but the integration of the whole process involved fitting together a large complex pattern of pieces. The single most important part of this product safety evaluation program is, in fact, the initial overall process of defining and developing an adequate data package on the potential hazards associated with the product life cycle (the manufacture, sale, use, and disposal of a product and associated process materials). To do this, one must ask a series of questions in a highly interactive process, with many of the questions designed to identify and/or modify their successors. The first is—what information is needed?

Required here is an understanding of the way in which a product is to be made and used and the potential health and safety risks associated with exposure of humans who will be associated with these processes. Such an understanding is the basis of a hazard and toxicity profile. Once such a profile

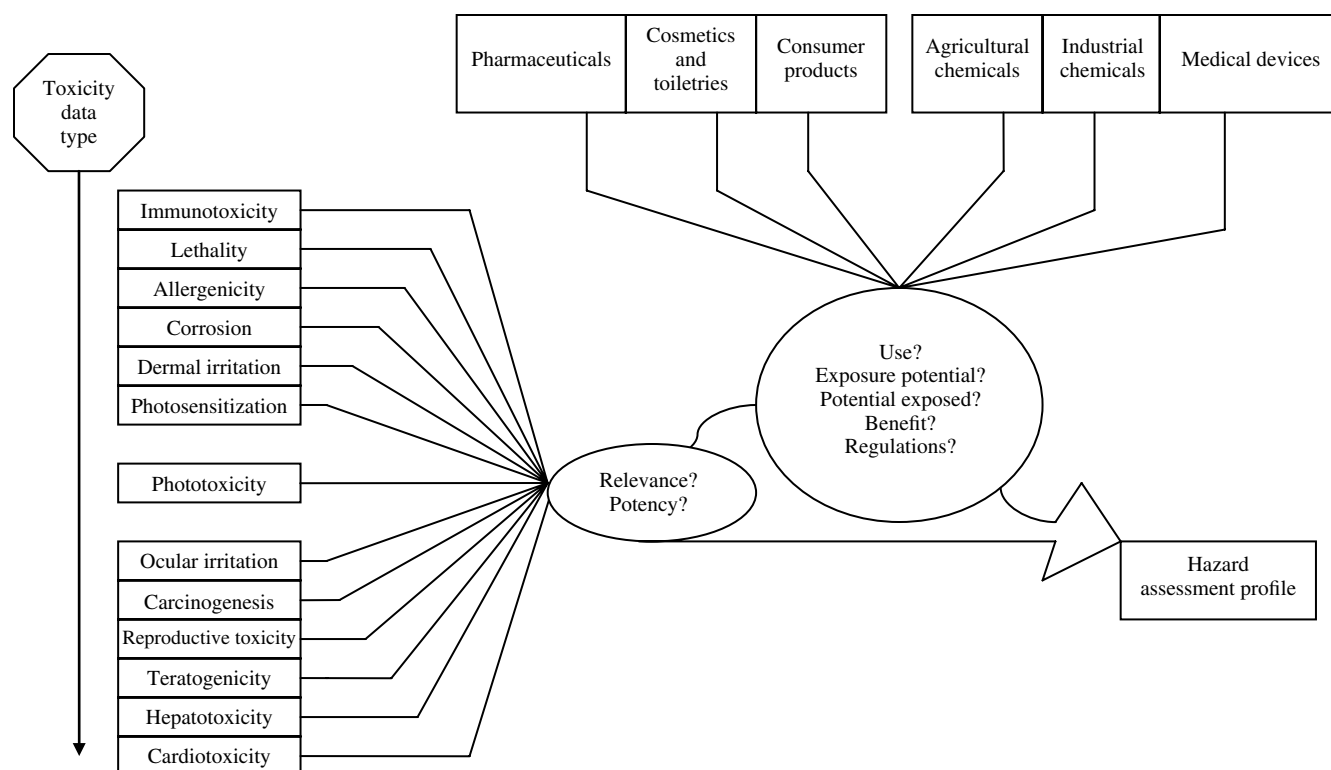


FIGURE 27.1 Development of a hazard assessment profile.

has been established (as illustrated in Figure 27.1), the available literature is searched to determine what is already known.

Taking into consideration this literature information and the previously defined exposure profile, a tier approach (Table 27.1) has traditionally been used to generate a list of tests or studies to be performed. What goes into a tier system is determined by regulatory requirements imposed by government agencies as well as the philosophy of the parent organization, economics, and available technology. How such tests are actually performed is determined on one of two bases. The first (and most common) is the menu approach: selecting a series of standard design tests as “modules” of data. The second is an interactive/iterative approach, where strategies are developed and studies are designed, based on both needs and what has been learned to date about the product. This process has been previously examined in some detail. Our interest here, however, is in the specific portion of the process involved in generating data—the test systems.

27.2 IN VITRO TESTING IN PHARMACEUTICAL SAFETY ASSESSMENT

The preclinical assessment of the safety of potential new pharmaceuticals and new devices represents a special case of the general practice of toxicology (Meyer, 1989; Gad, 1999,

2015), possessing its own peculiarities and special considerations and differing in several ways from the practice of toxicology in other fields—for some significant reasons. Because of the economics involved and the essential close interactions with other activities (e.g., clinical trials, chemical process optimization, formulation development, regulatory reviews), the development and execution of a crisp and flexible, yet scientifically sound, program are a prerequisite for success. The ultimate aim of preclinical safety and biocompatibility assessment also makes them different. A good safety assessment program seeks to efficiently and effectively move safe potential therapeutic agents or devices into the clinical evaluation, then to registration, and finally to market and to support them through this process. This requires the quick identification of those agents that are not safe so that efforts (and limited resources) are not wasted on them.

Pharmaceuticals are intended to have human exposure. Furthermore, pharmaceuticals are intended to have biological effects on the people that receive them. Frequently, the interpretation of results and the formulation of decisions about the continued development and eventual use of a drug are based on an understanding of both the potential adverse effects of the agent and its likely benefits, as well as the dose separation between these two. This makes a clear understanding of dose–response relationship critical, so that the actual risk/benefit ratio can be identified. It is also essential

TABLE 27.1 The Usual Way of Characterizing the Toxicity of a Compound or Product Is to Develop Information in a Tier Approach Manner

| Testing Tier | Tier Testing | | |
|--------------|--|---|--|
| | Mammalian Toxicology | Genetic Toxicology | Remarks |
| 0 | Literature review | Literature review | Upon initial identification of a problem, database of existing information and particulars of use of materials are established |
| 1 | Cytotoxicity screens GPMT or LLNA Acute systemic toxicity Receptor binding, that is, hERG | Ames test <i>In vitro</i> SCE <i>In vitro</i> cytogenetics Comet assay | R&D material and low-volume chemicals with severely limited exposure |
| 2 | Subacute studies ADME Primary dermal irritation Safety pharmacology Eye irritation | <i>In vivo</i> SCE <i>In vivo</i> cytogenetics | Medium-volume materials and/or those with a significant chance of human exposure |
| 3 | Subchronic studies Reproduction Developmental toxicity Chronic studies Mechanistic studies | — | Any materials with a high volume or a potential for widespread or long-term human exposure or one that gives indications of specific long-term effects |

More information is required (a higher tier level is attained) as the volume of production and potential for exposure increase. A common scheme is shown.

that the pharmacokinetics be understood and that “doses” (plasma tissue levels) at target organ sites (both therapeutic and toxicological) be known (Scheuplein et al., 1990). Integral pharmacokinetics are essential to such a safety program, especially now that there is wider recognition of the existence and importance of subpopulations with different metabolic competencies. As we have come to understand that pharmacogenetics underlie many of the subpopulation effects we see in both the safety and efficacy of drugs, we have also come to recognize that *in vitro* methods also offer some of the best and most efficient means of understanding the basis for these differences and for identifying members of specific subpopulations.

The development and safety evaluation of pharmaceuticals and medical devices have many aspects broadly or tightly specified by regulatory agencies (Gad, 2001). An extensive set of defined safety evaluations is required before a product is ever approved for market. For pharmaceuticals, regulatory agencies have increasingly come to require not only the establishment of a “clean dose” in two species with adequate safety factors to cover potential differences between species but also an elucidation of the mechanisms underlying those adverse effects that are seen at higher doses and are not well understood. These regulatory requirements are compelling to the pharmaceutical toxicologist (Traina, 1983). There is not, however, a set menu of what must be done. Rather, much (particularly in terms of the timing of testing) is open to professional judgment. Devices have tended to be more set piece in their testing approach but are beginning to likewise require more mechanistic understanding to allow for competitive positioning in the marketplace.

The discovery, development, and registration of a pharmaceutical or biologic are an immensely expensive operation and represent a rather unique challenge. For every 9000–10000 compounds specifically synthesized or isolated as potential therapeutics, one (on average) will actually reach the market (though of every 250 actually entering development, two will likely reach the marketplace). The overall cost for each successful compound is currently estimated to be between 1 and 1.3 billion dollars (though those figures are, of course, burdened with the cost of all the unsuccessful compounds), with each successive stage in the development process being more expensive. This dynamic makes it of great interest to identify as early as possible those agents that are likely not to go the entire distance, allowing a concentration of effort on the compounds that have the highest probability of reaching the market (and of possessing therapeutic utility) to do so.

Compounds “drop out” of the process primarily for three reasons: (i) toxicity or (lack of) tolerance, (ii) (lack of) efficacy, and (iii) (lack of) bioavailability of the therapeutic active moiety in humans. Early identification of “losers” in each of these three categories is thus extremely important (Fishlock, 1990), forming the basis for the use of screening in pharmaceutical discovery and development. How much and which resources to invest in screening and each successive step in support of the development of a potential drug are matters of strategy and phasing that are detailed elsewhere (Gad, 2000). A range of test systems is available to be used in screening and in the definitive testing that follows for selected promising compounds. Table 27.2 presents a summary of the levels of available model systems. Those test systems

TABLE 27.2 Levels of Models for Safety Assessment and Toxicological Research

| Level/Model | Advantages | Disadvantages |
|---|---|--|
| <i>In vivo</i> (intact higher organism) | Full range of organismic responses similar to target species | Cost Ethical/animal welfare concerns Species-to-species variability |
| Lower organisms (earthworms and fish) | Range of integrated organismic responses | Frequently lack responses typical of higher organisms Animal welfare concerns |
| Isolated organisms | Intact yet isolated tissue and vascular system Controlled environment and exposure conditions | Donor organism still required Time-consuming and expensive No intact organismic responses Limited duration of viability |
| Cultured cells | No intact animals directly involved Ability to carefully manipulate system | Instability of system Limited enzymatic capabilities and viability of system |
| Chemical/biochemical systems | Low cost Ability to study a wide range of variables No donor organism problems Low cost Long-term stability of preparation Ability to study a wide range of variables Specificity of response | No (or limited) integrated multicell and/or organismic responses No <i>de facto</i> correlation to <i>in vivo</i> systems Limited to investigation of a single defined mechanism |
| Genomics and proteomics | Speed and broad scope | Much effort is still required to correlate to intact organism effects |
| Computer simulations | No animal welfare concerns Speed and low per-evaluation cost | May not have predictive value beyond a narrow range of structures Expensive to establish |

that involve *in vitro* methods are now providing new tools for use in both early screening and in understanding the mechanisms of observed toxicity in preclinical and clinical studies (Gad, 1988b). Devices are generally less complicated in design and in their testing procedures and have a much lower rate of failure in the qualification and approval stages that precede going to market. The trend in devices, however, is for regulatory authorities to require more testing, to be more critical of results, and to take longer in the review and approval process.

The entire safety assessment process that supports new product research and development is a multistage effort in which none of the individual steps are overwhelmingly complex, but for which the integration of the whole process involves fitting together a large and complex pattern of pieces. This chapter proposes an approach in which integration of *in vitro* test systems calls for a modification of the approach to the general safety assessment problem. This modification can be addressed by starting with the current general case and progressing to a means of changing the process in an iterative fashion as new tools become available. Particularly with an understanding of mechanisms of toxicity becoming increasingly important in both candidate drug selection and the design and evaluation of the relevance of findings, the integration of *in vitro* methodologies particularly into the pharmaceutical safety assessment process has

become essential. Determining what information is needed calls for an understanding of the way in which the device or pharmaceutical is to be made and used, as well as an understanding of the potential health and safety risks associated with exposure of humans who will be either using the drug/device or associated with the processes involved in making it. This is on the basis of a hazard and toxicity profile. Once such a profile is established, the available literature is searched to determine what is already known. Much of the necessary information for support of safety claims in registration of a new drug/device is regulatorily mandated. This is not the case at all, however, for those safety studies done (i) to select candidate products or materials for development, or (ii) to design pivotal safety studies to support registration, or (iii) to pursue mechanistic questions about materials and products in development.

Taking into consideration this literature information and the previously defined exposure profile, investigators have traditionally used a tier approach to generate a list of tests or studies to be performed based on regulatory requirements. What goes into a tier system is determined by (i) regulatory requirements imposed by government agencies, (ii) the philosophy of the parent organization, (iii) economics, and (iv) available technology. How such tests are actually performed is determined on one of the two bases. The first (and most common) is the menu approach, which involves selecting a

series of standard design tests as “modules” of data. This assumes that all drugs or devices are alike except for route and duration of administration. The second is an interactive/iterative approach, where strategies are developed and studies are designed based both on needs and on what has been learned to date about the product.

27.3 DEFINING TESTING OBJECTIVE

The initial and most important aspect of a product safety evaluation program is the series of steps that leads to an actual statement of the problem or of the objectives of testing and research programs. This definition of objectives is essential and, as proposed here, consists of five steps: (i) defining product or material use, (ii) estimating or quantitating exposure potential, (iii) identifying potential hazards, (iv) gathering baseline data, and (v) designing and defining the actual research program to answer outstanding questions.

27.3.1 Objectives behind Data Generation and Utilization

To understand how product safety and toxicity data are used and how the data generation process might be changed to better meet the product safety assessment needs of society, it is essential to understand that different regulatory organizations have different answers to these questions. The ultimate solution is in the form of a multidimensional matrix, with the three major dimensions of the matrix being (i) the toxicity/biocompatibility data type (lethality, sensitization, corrosion, irritation, photosensitization, phototoxicity, etc.), (ii) exposure characteristics (extent, population size, population characteristics, etc.), and (iii) the stage in the research and development process we are dealing with.

What is called for is a careful zero-based consideration of what the optimum product safety assessment strategy for a particular development problem should be. Before formulating such a strategy and deciding what mix of tests should be used, it is first necessary to decide criteria for what would constitute an ideal (or at least acceptable) test system.

The ideal test should have an end-point measurement that provides data such that dose–response relationships can be obtained where possible or necessary (and such are almost always necessary). Furthermore, any criterion of effect must be sufficiently accurate in the sense that it can be used to reliably resolve the relative toxicity of two compounds that produce distinct (in terms of hazard to humans) yet similar responses. In general, it may not be sufficient to classify compounds into generic toxicity categories, such as “intermediate” toxicity, since a candidate chemical that falls in a given category yet is borderline to the next more severe toxicity category should be treated with more concern than a second candidate that falls at the less toxic extreme of the same category. Therefore it is useful for a test system to be

able to rank compounds with potentially similar uses accurately within any common toxicity category.

The end-point measurement of the “ideal” test system must be objective, so that a given compound will give similar results when tested using the standard test protocol in different laboratories. If it is not possible to obtain reproducible results in a given laboratory over time or between various laboratories, then the historical database against which new compounds are evaluated will be time and laboratory dependent. Along these lines, it is important for the test protocol to incorporate internal standards to serve as quality controls (QCs). Thus, test data could be represented utilizing a reference scale based on the test system response to the internal controls. Such normalization, if properly documented, could reduce between-test variability.

The test results from any given compound should be reproducible both intrinsically (within the same laboratory over time) and extrinsically (between laboratories). If these conditions are not satisfied, then there will be significant limitations on the application of the test system because it could potentially produce conflicting results at different times and places. Such a possibility would significantly reduce confidence in the outcome of any single assay or assay set. From a regulatory point of view, this possibility would be highly undesirable (and perhaps indefensible). Alternatives to current *in vivo* test systems basically should be designed to evaluate the subject toxic response in a manner as closely predictive of that occurring in humans as possible while also reducing animal use and avoiding inhumane treatments where possible.

From a practical point of view, several additional features of the “ideal” test should be satisfied. The test should be rapid so that the turnaround time for a given compound is reasonable. Obviously, the speed of the test and the ability to conduct tests on several candidate drugs or materials simultaneously will determine the overall productivity. The test should be inexpensive, so that it is economically competitive with current testing practices (in the pharmaceutical industry, any reduction in critical path time for decisions has great economic value, so speed is generally preferable to lower cost—within limits). And finally, the technology should be easily transferred from one laboratory to another without excessive capital investment specific to test implementation. Although some of these practical considerations may appear to present formidable limitations for a given test system at the present time, the possibility of future developments in testing technology could overcome these obstacles.

27.4 TEST SYSTEMS: CHARACTERISTICS, DEVELOPMENT, AND SELECTION

Any useful test system must be sufficiently sensitive to ensure that the incidence of false negatives is low. Clearly a high incidence of false negatives is intolerable. In such a

situation, a large number of dangerous chemical agents would be carried through extensive additional testing only for it to be found that they possess undesirable toxicological properties after the expenditure of significant time and money. On the other hand, a test system that is overly sensitive will give rise to a high incidence of false positives, which will have the deleterious consequence of rejecting potentially beneficial chemicals. The "ideal" test will fall somewhere between these two extremes and thus provide adequate protection without unnecessarily stifling development.

The "ideal" test should have an end-point measurement that provides data such that dose-response relationships can be obtained. Furthermore, any criterion of effect must be sufficiently accurate in the sense that it can be used reliably to resolve the relative toxicity of two test chemicals that produce distinct (in terms of hazard to humans) yet similar responses. In general, it may not be sufficient to classify test chemicals into generic toxicity categories. For instance, if a test chemical falls into an "intermediate" toxicity category, yet is borderline to the next more severe toxicity category, it should be treated with more concern than a second test chemical that falls at the less toxic extreme of the same category. Therefore, it is essential for a test system to be able to both place test chemicals in an established toxicity category and rank materials relative to others in the category.

The end-point measurement of the "ideal" test system must be objective. This is important to ensure that a given test chemical will give similar results when tested using the standard test protocol in different laboratories. If it is not possible to obtain reproducible results in a given laboratory over time or between various laboratories, then the historical database against which new test chemicals are evaluated will be time/laboratory dependent. If this condition is the case, then there will be significant limitations on the application of the test system since it could potentially produce conflicting results. From a regulatory point of view, this possibility would be highly undesirable. Along these lines, it is important for the test protocol to incorporate internal standards to serve as QCs. Thus, test data could be represented utilizing a reference scale based on the test system response to the internal controls. Such normalization, if properly documented, could reduce interest variability.

From a practical point of view, there are several additional features of the "ideal" test which should be satisfied. Alternatives to current *in vivo* test systems basically should be designed to evaluate the observed toxic response in a manner as closely predictive of the outcome of interest in man as possible. In addition, the test should be fast enough to ensure that the turnaround time for a given test chemical is reasonable for the intended purpose, very rapid for a screen, and timely for a definitive test. Obviously the speed of the test and the ability to conduct tests on several

chemicals simultaneously will determine the overall productivity. The test should be inexpensive, so that it is economically competitive with current testing practices. And finally, the technology should be easily transferred from one laboratory to another without excessive capital investment (relative to the value of the test performed) for test implementation.

It should be kept in mind that although some of these practical considerations may appear to present formidable limitations for any given test system at the present time, the possibility of future developments in testing technology could overcome these obstacles. In reality, these practical considerations are grounds for consideration of multiple new candidate tests on the basis of competitive performance. The most predictive test system in the universe of possibilities will never gain wide acceptance if it takes years to produce an answer or costs substantially more than other test systems that are only marginally less predictive.

The point is that these characteristics of the "ideal" test system provide a general framework for evaluation of alternative test systems in general. No test system is likely to be "ideal." Therefore, it will be necessary to weigh the strengths and weaknesses of each proposed test system in order to reach a conclusion on how "good" a particular test is.

In both theory and practice, *in vivo* and *in vitro* tests have potential advantages. Tables 27.3 and 27.4 summarize their advantages. How, then, might the proper tests be selected, especially in the case of the choice of staying with an existing test system or adopting a new one? The next section will present the basis for selection of specific tests.

27.5 IN VITRO MODELS

In vitro models, at least as screening tests, have been in use in toxicology since the early 1980s. The last 20+ years have brought a great upsurge in interest in such models. This increased interest is due to economic and animal welfare pressures and technological improvements.

Criteria against which an *in vitro* model should be evaluated for its suitability in replacing (partially or entirely) an accepted *in vivo* model are incorporated in the process detailed in Table 27.5, which presents the proposed steps for taking a new *in vitro* testing technology from being a research construct to a validated and accepted test system.

There are substantial potential advantages in using an *in vitro* system in toxicological testing which include isolation of test cells or organ fragments from homeostatic and hormonal control and accurate dosing and quantification of results. It should be noted that, in addition to the potential advantages, *in vitro* systems per se also have a number of limitations which can contribute to them not being acceptable models. Findings from an *in vitro* system that either limit their use in predicting *in vivo* events or make them

TABLE 27.3 Rationale for Using *In Vivo* Test Systems

1. Provides evaluation of actions/effects on intact animal organ–tissue interactions
2. Either neat chemicals or complete formulated products (complex mixtures) can be evaluated
3. Either concentrated or diluted products can be tested
4. Yields data on the recovery and healing processes
5. Required statutory tests for agencies under such laws as the Federal Hazardous Substances Act (unless data are already available); Toxic Substances Control Act; Federal Insecticide, Fungicide, and Rodenticide Act (FIFRA); and Organisation for Economic Co-operation and Development (OECD) and Food and Drug Administration laws
6. Quantitative and qualitative tests with scoring system generally capable of ranking materials as to relative hazards
7. Amenable to modifications to meet the requirements of special situations (such as multiple dosing or exposure schedules)
8. Extensive available database and cross-reference capability for evaluation of relevance to human situation
9. The ease of performance and relative low capital costs in many cases
10. Tests are generally both conservative and broad in scope, providing for maximum protection by erring on the side of overprediction of hazard to man
11. Tests can be either single end point (such as lethality and corrosion) or shotgun (also called multiple end point, including such test systems as a 13-week oral toxicity study)

TABLE 27.4 Limitations of *In Vivo* Testing Systems Which Serve as a Basis for Seeking *In Vivo* Alternatives for Toxicity Tests

1. Complications and potential confounding or masking findings of *in vivo* systems
2. *In vivo* systems may only assess short-term site of application or immediate structural alterations produced by agents. Specific *in vivo* tests may only be intended to evaluate acute local effects (i.e., this may be a purposeful test system limitation)
3. Technician training and monitoring are critical (particularly in view of the subjective nature of evaluation)
4. *In vivo* tests in animals do not perfectly predict results in humans if the objective is to exclude or identify severe-acting agents
5. Structural and biochemical differences between test animals and humans make extrapolation from one to the other difficult
6. Lack of standardization of *in vivo* systems
7. Variable correlation with human results
8. Large biological variability between experimental units (i.e., individual animals)
9. Large, diverse, and fragmented databases which are not readily comparable

totally unsuitable for the task include wide differences in the doses needed to produce effects or differences in the effects elicited. Some reasons for such findings are detailed in Table 27.6.

Tissue culture has the immediate potential to be used in two very different ways by industry. First, it has been used to examine a particular aspect of the toxicity of a compound in relation to its toxicity *in vivo* (i.e., mechanistic or explanatory studies). Second, it has been used as a form of rapid screening to compare the toxicity of a group of compounds for a particular form of response. Indeed, the pharmaceutical industry has used *in vitro* test systems in these two ways for years in the search for new potential drug entities.

The theory and use of screens in toxicology have previously been reviewed (Gad, 1988a, 1989a, b). Mechanistic and explanatory studies are generally called for when a traditional test system gives a result that is either unclear or is one for which the relevance to the real-life human exposure is doubted. *In vitro* systems are particularly attractive for such cases because they can focus on very defined single aspects of a problem or pathogenic response, free of the confounding influence of the multiple responses of an intact higher-level organism. Note, however, that the first one must know the nature (indeed the existence) of the questions to be addressed. It is then important to devise a suitable model system which is related to the mode of toxicity of the compound.

There is currently much controversy over the use of *in vitro* test systems—will they find acceptance as “definitive tests systems” or only be used as preliminary screens for such final tests? Or, in the end, not be used at all? Almost certainly, all three of these cases will be true to some extent. Depending on how the data generated are to be used, the division between the first two is ill-defined at best.

Before trying to definitely answer these questions in a global sense, each of the end points for which *in vitro* systems are being considered should be overviewed and considered against the factors outlined to this point.

27.6 LETHALITY

Many of the end points of interest in toxicology present a fundamental limitation to the development and use of an *in vitro* or nonmammalian system in place of established *in vivo* methods. While cytotoxicity is a component mechanism in many of these toxic responses, disruption or diminution of the integrated function of multiple cells and systems is just as important.

The evaluation of lethality (symbolized in the public mind by the LD₅₀ test) would seem to offer a unique opportunity for the development and use of alternatives. Approaches to alternatives for lethality testing include no living materials at all (the structure–activity relationship (SAR) or computer model approaches), those that use no intact higher organisms (but rather cultured cells or bacteria) and those that use lower forms of animal life (e.g., invertebrates and fish). Each of these approaches presents a different approach to the

TABLE 27.5 Multistage Scheme for the Development, Validation, and Transfer of *In Vitro* Test System Technology in Toxicology

Stage I: Statement of Test Objective

- A. Identify existing test system and its strengths and weaknesses
- B. Clearly state objectives for alternative test system
- C. Identify potential alternative test system

Stage II: Define Developmental Test Design

- A. Identify relevant variables
- B. Evaluate effects of variables on test system
- C. Optimize test performance
- D. Understand what the test does in a functional sense
 - 1. Is it a simulation of an *in vivo* event?
 - 2. Is this simply a response to the presence of the agent?
 - 3. Is the measured response a functional step or link in the *in vivo* event of interest?
 - 4. Is the measured response a functional step or link in the *in vivo* event of interest or some intermediate stage?
 - 5. Is this an effect on some structure or function analogous to the *in vivo* structure or function?

Stage III: Evaluate Performance of Optimum Test

- A. Develop library of known positive- and negative-response materials of diverse structure and a range of response potencies (i.e., if the end point is irritation, then materials should range from nonirritating to severely irritating)
- B. Use optimum test design to evaluate the library of “knowns” under “blind” conditions
- C. Compare correlation of test results with those of other test systems and with real case of interest—results in humans

Stage IV: Technology Transfer

- A. Present and publish results through professional media (at society meetings and in peer-reviewed journals)
- B. Provide hands-on training to personnel from other facilities and facilitate internal evaluations of test methods

Stage V: Validation

- A. Arrange for test of coded samples in multiple laboratories (i.e., interlaboratory validation)
- B. Compare, present, and publish results
- C. Gain regulatory acceptance

Stage VI: Continue to Refine and Evaluate Test System Performance and Utilization

- A. Continually strive for an understanding of why the test “works” and its relevance to effects in man
- B. Remain skeptical. Why should any one of us be the one to make the big breakthrough? Clearly, there is some basic flaw in the design or conduct of the study which has given rise to these promising results. Doubt, check, and question; let your most severe critic review the data; go to a national meeting and give a presentation; then go back home and doubt, check, and question some more!

Stage VII: Regulatory Acceptance and Adaptation

- A. Accepted use will first be side by side with existing system, which it will come to supplant
 - B. It will be modified in use as practical factors reveal means of optimization
-

TABLE 27.6 Possible Interpretations When *In Vitro* Data Do Not Predict Results of *In Vivo* Studies

-
1. Chemical is not absorbed at all or is poorly absorbed in *in vivo* studies
 2. Chemical is well absorbed but is subject to “first-pass effect” in the liver
 3. Chemical is distributed so that less (or more) reaches the receptors than would be predicted on the basis of its absorption
 4. Chemical is rapidly metabolized to an active or inactive metabolite that has a different profile of activity and/or different duration of action from that of the parent drug
 5. Chemical is rapidly eliminated (e.g., through secretory mechanisms)
 6. Species of the two test systems used are different
 7. Experimental conditions of the *in vitro* and *in vivo* experiments differed and may have led to different effects from those expected. These conditions include factors such as temperature or age, sex, and strain of animal
 8. Effects elicited *in vitro* and *in vivo* by the particular test substance in question differ in their characteristics
 9. Tests used to measure responses may differ greatly for *in vitro* and *in vivo* studies, and the types of data obtained may not be comparable
 10. The *in vitro* study did not use adequate controls (e.g., pH, vehicle used, volume of test agent given, and samples taken from sham-operated animals), resulting in “artifacts” of method rather than results
 11. *In vitro* data cannot predict the volume of distribution in central or in peripheral compartments
 12. *In vitro* data cannot predict the rate constants for chemical movement between compartments
 13. *In vitro* data cannot predict the rate constants of chemical elimination
 14. *In vitro* data cannot predict whether linear or nonlinear kinetics will occur with specific dose of a chemical *in vivo*
 15. Pharmacokinetic parameters (e.g., bioavailability, peak plasma concentration, and half-life) cannot be predicted solely on the basis of *in vitro* studies
 16. *In vivo* effects of chemical are due to an alteration in the higher-order integration of an intact animal system, which cannot be reflected in a less complex system
-

objective of predicting acute lethality in humans or, rarely, economic animals and will be examined in turn.

There are systems that do not directly use any living organisms but, rather, seek to predict the lethality (in particular, the LD_{50}) of a chemical on the basis of what is known about structurally related chemicals. Such SAR systems have improved markedly over the last 10 years (Enslein et al., 1983b; Lander et al., 1984) but are still limited. Accurate predictions are usually possible only for those classes of structures where data have previously been generated on several members of the classes. For new structural classes, the value of such predictions is minimal. Accordingly, this approach is valuable when working with analogs in a series but not for novel structures. It is also a strong argument for getting as many data as possible into the published literature.

A more extensive and once promising approach has been the use of various cultured cell systems. Kurack et al. (1986), for example, have developed and suggested a system based on cultured mammalian hepatocytes. The system does metabolize materials in a manner like mammalian target species and has shown promise in a limited battery of chemicals. Such mammalian cell culture and bacterial screening systems have significant weaknesses for assessing the lethality of many classes of chemicals, since they lack any of the integrative functions of a larger organism. Thus, they would miss all agents that act by disrupting functions such as the organophosphate pesticides, most other neurologically mediated lethal agents, and agents that act by modifying hormonal or immune systems.

Clive et al. (1979) have reported on the correlation of the LC_{50} of a variety of chemicals in mouse lymphoma cell cultures with their oral LD_{50} in mice, as shown in Figure 27.2.

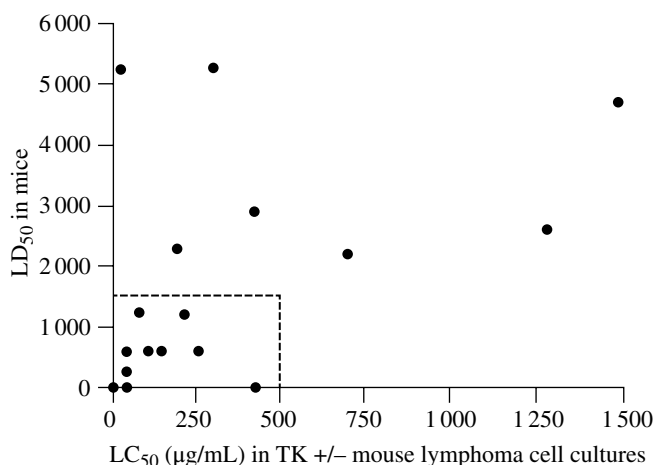


FIGURE 27.2 Graph showing a comparison of the lethality of a group of 18 drugs of diverse structure in *in vivo* (mouse) and *in vitro* (cultured mouse lymphoma cells) test systems. Correlation of these LD_{50}/LC_{50} values is very poor, though extreme high- and low-scale values seem to be more closely associated in the two systems.

No linear correlation is present, but highly cytotoxic substances (in this group) are significantly more toxic orally. Given the impression of some LD_{50} values, due to such factors as steepness of slope of the lethality curve, the lack of linear correlation should be no surprise. Most recently, Ekwall et al. (1989) have reported on the MEIC program system, which utilizes a battery of five cellular systems. For a group of ten chemicals, the system provided good correlation with, or predictive power of, rat LD_{50} .

Recently Parce et al. (1989) reported on a biosensor technique in which cultured cells are confined to a flow chamber through which a sensor measures the rate of production of acidic metabolites. It is proposed to use this as a functional measure of cytotoxicity and as a screening technique for a number of uses, including *in vivo* lethality.

Three lower species of intact animals have been proposed for use in screening or testing of the lethal effects of chemicals. First, some researchers have shown a good correlation between the LD_{50} of the same chemicals in rats. This correlation is nonlinear but still suggests that more toxic materials could be at least initially identified and classified in some form of screening system based on *Daphnia*. A broader range of chemical structures will need to be evaluated, however, and some additional laboratories will need to confirm the finding. It must also be kept in mind that the metabolic systems and many of the other factors involved in species differences (as presented in Gad and Chengelis (1997), Gad (2015)) contribute to a nonlinear correlation and may also make the confidence in prediction of human effects in cases somewhat limited.

Earthworms have been one of the more common species used to test chemicals for potential hazardous impact on the environment. The 48 h contrast test has proved to be a fast and resource-effective way of assessing acute toxicity of chemicals in earthworms and is outlined in Table 27.7.

TABLE 27.7 Earthworm 48 h Contact Test: Acute Lethality

1. Place filter paper of known size (9 cm or 12×6.7 cm) in a petri dish or standard scintillation vial
2. Dilute test article in acetone or some other volatile solvent
3. Slowly and evenly deposit known amounts of test-article solution onto filter paper
4. Dry thoroughly with air or nitrogen gentle stream
5. Add 1.0 mL of distilled water to filter paper
6. Add worm (*Lumbricus rubellus*). Use 400–500 mg body weight range
7. Then replicate vials per concentration
8. Store/incubate in the absence of light at $15\text{--}20^\circ\text{C}$ for 48 h
9. Examine for lethality (swollen, lack of movement upon warming up to room temperature, and lack of response to tactile stimulation)
10. Express dose as $\mu\text{g cm}^{-2}$ and mortality as usual. Calculate LD_{50} using standard techniques
11. Always include negative and positive (benchmark) controls

The standardized method, approved by the EEC, is discussed by Neuhauser et al. (1986). This test is for environmental impact assessment where cross-laboratory comparisons are important. If, however, one wishes to adopt this technology for the purpose of screening new chemicals or releasing batches of antibiotics, then variants of this method may be acceptable, as internal consistency is more important than interlaboratory comparisons. There are two important considerations. First, because of seasonal variation in the quality of earthworms obtained from suppliers, positive controls or comparator chemicals should be included on every assay run. Second, distilled water must be used, as worms are quite sensitive to contaminants that may occur in chlorinated water. The filter paper should completely cover the sides of the vessel; otherwise the worms will simply crawl up the sides to escape the adverse stimulus the chemical contact may provide.

Using these techniques, Roberts and Dorough (1985) and Neuhauser et al. (1986) have compared acute toxicity in a variety of organic chemicals in several earthworm species. While there are some obvious differences between worm species, in general the rank order of toxicity is about the same. *Lumbricus rubellus* tends to be the most sensitive species. All earthworms are very sensitive to carbofuran under the conditions of this test. Neuhauser et al. (1985a, b) have proposed a toxicity-rating scheme based on acute lethality in the earthworms which is similar to the more familiar scheme based on acute lethality in rodents (Table 27.8). Roberts and Dorough (1985) and Neuhauser et al. (1986) have published extensive compilations of acute lethality in worms and compared these with acute lethality in rats and mice. A selection of these is shown in Table 27.9. Applying the rating scheme of Neuhauser, most chemicals receive about the same toxicity rating based on results in *Eisenia fetida* and mice. This may suggest that replacing the LD₅₀ with the LC₅₀ for rating toxicity (e.g., for transportation permits) deserves serious consideration.

The main advantages of the 40h contact test are the savings of time and money. The cost savings fall into three categories. First, earthworms are cheap. One hundred *L. rubellus* will cost about US \$2.00. The one hundred mice they could replace in screens and QC testing, for example, would cost \$300–\$400 (£150–£200) at an exchange rate of \$2.00=£1.00. Second, earthworms require no vivarium space, and their use could reduce the number of rodents used, resulting in a net decrease in vivarium use. Third, adapting the 48h contact test would require little capital investment other than a dedicated under-the-counter refrigerator set at 15–20°C. Otherwise, the assay can be easily performed in a standard biochemistry laboratory. With regard to time savings, the standard lethality test with rodents requires 7–14 days of postdosing observations. The 48h contact test is completed in 48h. Not only is the turnaround

TABLE 27.8 Earthworm Toxicity—Toxicity Rating

| Rating | Designation | Rat LD ₅₀ (mg kg ⁻¹) | <i>Eisenia fetida</i> LC ₅₀ (µg cm ⁻²) |
|--------|---------------------|--|--|
| 1 | Super toxic | <5 | <1.0 |
| 2 | Extremely toxic | 5–50 | 1.0–10 |
| 3 | Very toxic | 50–500 | 10–100 |
| 4 | Moderately toxic | 500–5000 | 100–1000 |
| 5 | Relatively nontoxic | >5000 | >1000 |

From Neuhauser et al. (1985a, b).

TABLE 27.9 Earthworm Acute Lethality: Comparative Values

| Chemical | <i>Eisenia fetida</i> (LC ₅₀) | Mouse (LD ₅₀) |
|-----------------------|---|---------------------------|
| 2,4-Dinitrophenol | 0.6 (1) ^a | 45 (2) |
| Carbaryl | 14 (3) | 438 (3) |
| Benzene | 75 (3) | 4,700 (4) |
| 1,1,1-Trichloroethane | 83 (3) | 11,240 (5) |
| Dimethyl phthalate | 550 (4) | 7,200 (5) |

time faster but also the amount of time that technical personnel will have to spend observing animals and recording observations will be reduced. An incidental advantage of earthworms is that they are cold-blooded vertebrates and thus are exempt from current animal welfare laws.

There are two main disadvantages to the use of earthworms in acute toxicity testing. First, there are a limited number of end points. Other than death and a few behavioral abnormalities (Stenersen 1979; Drewes et al., 1984), the test does not yield much qualitative information. Second, there probably is some institutional bias. Because the test is basically low technology (no tissue culture) and uses a non-mammalian model, it may be easy to dismiss the utility of the test.

Finally, the use of smaller species of fish as a surrogate for man has gained some supporters. Currently, the zebra fish has shown significance as a surrogate model for mammalian species toxicity (Hill et al., 2005; Chiu et al., 2008; Tanguay and Reimers, 2008) and as a screen for carcinogens (Stern and Zon, 2003; Berghmans et al., 2005). There is certainly no reason why they could not be used for screening water-soluble compounds for extreme acute toxicity.

Although the intact organisms would seem to be the most utilitarian on the face of it, they still will not totally replace mammalian systems, owing to the need to be concerned about those systems that are significantly different in the higher organisms. Still, it would appear that for those compounds for which human exposure is not intentional, testing in an intact lower organism system (or perhaps even in a cell culture system) should be sufficient to identify agents of significant concern. In these cases, lethality testing in intact mammals is probably unwarranted.

TABLE 27.10 Rationales for Seeking *In Vitro* Alternatives for Eye Irritancy Tests

1. Avoid whole-animal and organ *in vivo* evaluation
2. Strict Draize scale testing in the rabbit assesses only three eye structures (conjunctiva, cornea, and iris), and traditional rabbit eye irritancy tests do not assess cataracts, pain, discomfort, or clouding of the lens
3. *In vivo* tests assess only inflammation and immediate structural alterations produced by irritants (not sensitizers, photoirritants, or photoallergens). Note, however, that the test was (and generally is) intended to evaluate any pain or discomfort
4. Technician training and monitoring are critical (particularly in view of the subjective nature of evaluation)
5. Rabbit eye tests do not perfectly predict results in humans, if the objective is either the total exclusion of irritants or the identification of truly severe irritants on an absolute basis (i.e., without false positives or negatives). Some (such as Reinhardt et al., 1985) have claimed that these tests are too sensitive for such uses
6. There are structural and biochemical differences between rabbit and human eyes which make extrapolation from one to the other difficult. For example, Bowman's membrane is present and well developed in man (8–12 µm thick) but not in the rabbit, possibly giving the cornea greater protection
7. Lack of standardization
8. Variable correlation with human results
9. Large biological variability between experimental units
10. Large, diverse, and fragmented databases which are not readily comparable

27.6.1 Ocular Irritation

Testing for potential to cause irritation or damage to the eyes remains the most active area for the development (and validation) of alternatives and the most sensitive area of animal testing in biomedical research. This has been true since the beginning of the 1980s. Table 27.10 presents an overview of the reasons for pursuing such alternatives. The major reason, of course, has been the pressure from public opinion.

Indeed, many of the *in vitro* tests now being evaluated for other end points (such as skin irritation and lethality) are adaptations of test systems first developed for eye irritation uses. A detailed review of the underlying theory of each test system is beyond the scope of this chapter. Frazier et al. (1987) and Gad (1993) performed such reviews, and Table 27.11 presents an updated version of the list of test systems overviewed in those volumes.

There are six major categories of approach to *in vitro* eye irritation tests. Because of the complex nature of the eye, the different cell types involved, and interactions between them, it is likely that a successful replacement for existing *in vivo* systems (such as the rabbit) would require some form of battery of such test systems. Many individual systems, however, might constitute effective screens in defined situations. The first five of these aim at assessing portions of the irritation response, including alterations in tissue morphology, toxicity to individual component cells or tissue physiology, inflammation or immune modulation, and alterations in repair and/or recovery processes. These methods have the limitation that they assume that one of the component parts can or will predict effects in the complete organ system. While each component may serve well to predict the effects of a set of chemical structures which determine part of the ocular irritation response, a valid assessment across a broad range of structures will require the use of a collection or battery of such tests.

The sixth category contains tests that have little or no empirical basis, such as computer-assisted SAR models. These approaches can only be assessed in terms of how well

or poorly they perform. Table 27.12 presents an overview of all six categories and some of the component tests within them, updated from the assessment by Frazier et al. (1987), along with references for each test.

Given that there are now some 80 or more potential *in vitro* alternatives, the key points along the route to the eventual objective of replacing the *in vivo* test systems are thus: (i) How do we select the best candidates from this pool? (ii) How do we want to use the resulting system (as a screen or test)? (iii) How do we gain regulatory and user acceptance of the appropriate test systems?

There have been some large-scale validations of some of these tests (IRAG, 1993). Most of the individual investigators have performed smaller “validations” as part of their development of the test system, and in a number of cases, trade associations have sponsored comparative and/or multi-laboratory validations. At least for screening, several systems should be appropriate for use and, in fact, are used now by several commercial organizations. But the Interagency Regulatory Alternatives Group (1993) and the Interagency Coordinating Committee on the Validation of Alternative Methods (1997) have coordinated and reported on large-scale evaluations. In terms of use within defined chemical structural classes, use of *in vitro* systems for testing of chemicals for nonhuman exposure should supplant traditional *in vivo* systems once validated on a broad scale by multiple laboratories. Broad use of single tests based on single end points (such as cytotoxicity) is not likely to be successful, as demonstrated by such efforts as those of Kennah et al. (1989). In 2015, the FDA announced that it would no longer require a rabbit eye test but rather wanted the alternative bovine corneal opacity and permeability (BCOP) test.

27.6.2 Dermal Irritation

Extensive progress has been made in devising alternative (*in vitro*) systems for evaluating the dermal irritation potential of chemicals since this author first reviewed the field

TABLE 27.11 *In Vitro* Alternatives for Eye Irritation Tests

| |
|---|
| Morphology |
| Enucleated superfused rabbit eye system (Burton et al., 1981) |
| Balb/c 3T3 cells/morphological assays (HTD) (Borenfreund and Puermer, 1984) |
| Cell Toxicity |
| Adhesion/cell proliferation |
| BHK cells/growth inhibition (Reinhardt et al., 1985) |
| BHK cells/colony formation efficiency (Reinhardt et al., 1985) |
| BHK cells/cell detachment (Reinhardt et al., 1987) |
| SIRC cells/colony-forming assay (North-Root et al., 1982) |
| Balb/c 3T3 cells/total protein (Shopsis and Eng, 1985) |
| BCL/D1 cells/total protein (Balls and Horner, 1985) |
| Primary rabbit corneal cells/colony-forming assay (Watanabe et al., 1988) |
| Membrane integrity |
| LS cells/dual dye staining (Scaife, 1982) |
| Thymocytes/dual fluorescent dye staining (Aeschbacher et al., 1986) |
| LS cells/dual dye staining (Kemp et al., 1983) |
| RCE-SIRC-P815-YAC-1/Cr release (Shaddock et al., 1985) |
| L929 cells/cell viability (Simons, 1981) |
| Bovine red blood cell/hemolysis (Shaddock et al., 1985) |
| Mouse L929 fibroblasts/erythrocin C staining (Frazier, 1988) |
| Rabbit corneal epithelial and endothelial cells/membrane leakage (Meyer and McCulley, 1988) |
| Agarose diffusion (Barnard, 1989) |
| Corneal protein profiles (Eurell and Meachum, 1994) |
| Cell metabolism |
| Rabbit corneal cell cultures/plasminogen activator (Chan, 1985) |
| LS cells/ATP assay (Kemp et al., 1985) |
| Balb/c 3T3 cells/neutral red uptake (Borenfreund and Puermer, 1984) |
| Balb/c 3T3 cells/uridine uptake inhibition assay (Shopsis and Sathe, 1984) |
| HeLa cells/metabolic inhibition test (MIT-24) (Selling and Ekwall, 1985) |
| MDCK cells/dye diffusion (Tchao, 1988) |
| Cell and Tissue Physiology |
| Epidermal slice/electrical conductivity (Oliver and Pemberton, 1985) |
| Rabbit ileum/contraction inhibition (Muir et al., 1983) |
| Bovine cornea/corneal opacity (Muir, 1984) |
| Proposed mouse eye/permeability test (Maurice and Singh, 1986) |
| Inflammation/Immunity |
| Chlorioallantonic membrane (CAM) |
| CAM (Leighton et al., 1983) |
| HET-CAM (Luepke, 1985) |
| Bovine corneal cup model/leukocyte chemotactic factors (Elgebaly et al., 1985) |
| Rat peritoneal cells/histamine release (Jacaruso et al., 1985) |
| Rat peritoneal mast cells/serotonin release (Dubin et al., 1984) |
| Rat vaginal explant/prostaglandin release (Dubin et al., 1984) |
| Bovine eye cup/histamine (Hm) and leukotriene C4 (Lt/C4) release (Benassi et al., 1986) |
| Recovery/Repair |
| Rabbit corneal epithelial cells/wound healing (Jumblatt and Neufeld, 1985) |
| Other |
| EYTEX assay (Gordon et al., 1986; Soto et al., 1988) |
| Computer-based structure-activity relationship (SAR) (Enslein, 1984) |
| <i>Tetrahymena</i> /motility (Silverman, 1983) |

(Gad and Chengelis, 1997). Table 27.12 overviews 20 proposed systems which now constitute five very different approaches.

The first approach (I) uses patches of excised human or animal skin maintained in some modification of a glass diffusion cell which maintains the moisture, temperature, oxygenation, and electrolyte balance of the skin section. In this approach, after the skin section has been allowed to equilibrate for some time, the material of concern is placed on the exterior surface and wetted (if not a liquid). Irritation is evaluated either by swelling of the skin (a crude and relatively insensitive method for mild and moderate irritants), by evaluation of inhibition of uptake of radiolabeled nutrients, or by measurement of leakage of enzymes through damaged membranes.

The second set of approaches (II) utilizes a form of surrogate skin culture comprising a mix of skin cells which closely mirror key aspects of the architecture and function of the intact organ. These systems seemingly offer a real potential advantage, but, to date, the "damage markers" employed (or proposed) as predictors of dermal irritation have been limited to cytotoxicity.

The third set of approaches (III) is to use some form of cultured cell (either primary or transformed), with primary human epidermal keratinocytes (HEKs) preferred. The cell cultures are exposed to the material of interest, and then either cytotoxicity, release of inflammation markers, or decrease of some indicator of functionality (lipid metabolism, membrane permeability, or cell detachment) is measured.

The fourth group (IV) contains two miscellaneous approaches—the use of a membrane from the hen's egg with a morphological evaluation of damage being the predictor end point (Lei et al., 1986) and the SKINTEX system, which utilizes the coagulation of a mixture of soluble proteins to predict dermal response.

Finally, in group V there are two SAR models which use mathematical extensions of past animal results correlated with structure to predict the effects of new structures.

Many of these test systems are in the process of evaluation of their performance against various small groups of compounds for which the dermal irritation potential is known. Evaluation by multiple laboratories of a wider range of structures will be essential before any of these systems can be generally utilized.

27.6.3 Irritation of Parenterally Administered Pharmaceuticals

Intramuscular (IM) and intravenous (IV) injection of parenteral formulations of pharmaceuticals can produce a range of discomfort resulting in pain, irritation, and/or damage to the muscular or vascular tissue. These are normally evaluated for prospective formulations before use in humans by evaluation in intact animal models—usually the rabbit (Gad and Chengelis, 1997).

Currently, a protocol utilizing a cultured rat skeletal muscle cell line (the L6) as a model is in an interlaboratory

TABLE 27.12 *In Vitro* Dermal Irritation Test Systems

| System | End Point | Validation Data? ^a | References |
|---|---|-------------------------------|---------------------------------|
| I. | | | |
| Excised patch of perfused skin | Swelling | No | Dannenberget al. (1987) |
| Mouse skin organ culture | Inhibition of incorporation of [³ H]-thymidine and [¹⁴ C]-leucine labels | No | Kao et al. (1982) |
| Mouse skin organ culture | Leakage of LDH and GOT | Yes | Bartnik et al. (1989) |
| II. | | | |
| TESTSKIN—cultured surrogate skin patch | Morphological evaluation | No | Bell et al. (1988) |
| Cultured surrogate skin patch | Cytotoxicity | No | Naughton et al. (1989) |
| III. | | | |
| Human epidermal keratinocytes (HEKs) | Release of labeled arachidonic acid | Yes | DeLeo et al. (1988) |
| Human polymorphonuclear cells | Migration and histamine release | Yes (surfactants) | Frosch and Czarnetzki (1987) |
| Fibroblasts | Acid | | Lamont et al. (1989) |
| HEKs | Cytotoxicity | Yes | Gales et al. (1989) |
| HEKs | Cytotoxicity (MTT) | Yes | Swisher et al. (1988) |
| HEKs, dermal fibroblasts | Cytotoxicity | Yes | Babich et al. (1989) |
| HEKs | Inflammation mediator release | No | Boyce et al. (1988) |
| Cultured Chinese hamster ovary (CHO) cells | Increases in β -hexosaminidase levels in media | No | Lei et al. (1986) |
| Cultured C ₃ H10T _{1/2} and HEK cells | Lipid metabolism inhibition | No | DeLeo et al. (1987) |
| Cultured cells—BHK21/C13 | Cell detachment | Yes | Reinhardt et al. (1987) |
| BHK21/C13 primary rat thymocytes | Growth inhibition | | |
| | Increased membrane permeability | | |
| Rat peritoneal mast cells | Inflammation mediator release | Yes (surfactants) | Prottey and Ferguson (1976) |
| IV. | | | |
| Hen's egg | Morphological evaluation | | Reinhardt et al. (1987) |
| SKINTEX—protein mixture | Protein coagulation | Yes | Gordon et al. (1989) |
| V. | | | |
| Structure–activity relationship (SAR) model | NA | Yes | Enslein et al. (1987) |
| SAR model | NA | No | Firestone and Guy (1986) |

NA, not available.

^aEvaluated by comparison of predictive accuracy for a range of compounds compared with animal test results. Not validated in the sense used in this chapter.

validation program among more than ten pharmaceutical company laboratories. This methodology (Young et al., 1986) measures creatine kinase levels in media after exposure of the cells to the formulation of interest and predicts *in vivo* intramuscular damage based on this end point. It is reported to give excellent rank-correlated results across a range of antibiotics (Williams et al., 1987) and in a recent multilaboratory evaluation a broader structural range of compounds (PMA/Drusafe *In Vitro* Task Force, 1994).

Another proposed *in vitro* assay for muscle irritancy for injectable formulations is the red blood cell hemolysis assay (Brown et al., 1989). Water-soluble formulations are gently mixed at a 1:2 ratio with freshly collected human blood for 5 s and then mixed with a 5% w/v dextrose solution and centrifuged for 5 min. The percentage of red blood cell survival is then determined by measuring differential absorbance at 540 nm, and this is compared with values for known irritants and nonirritants. Against a very small group of

compounds (four), this is reported to be an accurate predictor of muscle irritation.

There is no current candidate alternative for the venous irritation test, but the *in vitro* alternative for pyrogenicity testing—the *Limulus* test—is one of the success stories for the alternatives movement. It has totally replaced the classical intact rabbit test in both research and product release testing. The test is based on the jelling or color development of a pyrogenic preparation in the presence of the lysate of the amoebocytes of the horseshoe crab (*Limulus polyphemus*). It is simpler, more rapid, and of greater sensitivity than the rabbit test it replaced (Cooper, 1975).

27.6.4 Sensitization and Photosensitization

There are actually several approaches available for the *in vitro* evaluation of materials for sensitizing potential. These use cultured cells from various sources and, as end

TABLE 27.13 Requested Reference Compounds for Skin Sensitization Studies (US Consumer Product Safety Commission)

| | |
|---|--|
| Hydroxylamine sulfate | Penicillin G |
| Ethyl amino benzoate | <i>p</i> -Phenylenediamine |
| Iodochlorohydroxyquinoline (clioquinol, chionoform) | Epoxy systems (ethylenediamine, diethylenetriamine, and diglycidyl ethers) |
| Nickel sulfate | Toluene-2,4-diisocyanate |
| Monomethyl methacrylate | Oil of Bergamot |
| Mercaptobenzothiazole | |

points, look at either biochemical factors (such as production of MIF—migration inhibition factor) or cellular events (such as cell migration or cell “transformation”).

While on the surface, the local lymph node assay (LLNA), previously discussed in Chapter 11, was a significant step toward replacing traditional guinea pig-based sensitization assays, it has both limitations on performance and does not eliminate the use of animals. Arancioglu et al. (2015) have proposed an *ex vivo* alternative that measures three additional end points (lymphocyte proliferation, ear swelling, and cytokine profiles) to improve test performance.

Milner (1970) reported that lymphocytes from guinea pigs sensitized to dinitrofluorobenzene (DNFB) would transform in culture, as measured by the incorporation of tritiated thymidine, when exposed to epidermal proteins conjugated with DNFB. This work was later extended to guinea pigs sensitized to *p*-phenylenediamine. He also reported (Milner, 1971) that his method was capable of detecting allergic contact hypersensitivity to DNFB in humans, using human lymphocytes from sensitized donors and human epidermal extracts conjugated with DNFB.

Miller and Levis (1973) reported the *in vitro* detection of allergic contact hypersensitivity to DNCB conjugated to leukocyte and erythrocyte cellular membranes. This indicated that the reaction was not specifically directed toward epidermal cell conjugates. Thulin and Zacharian (1972) extended others' earlier work on MIF-induced migration of human peripheral blood lymphocytes to a test for delayed contact hypersensitivity. Burka et al. reported in 1981 (Burka et al., 1981) on an assay system based on isolated guinea pig trachea. No further mention of this has been found in the literature. None of these approaches have yet been developed as an *in vitro* predictive test, but work is progressing. Milner published a review of the history and state of this field in 1983 which still provides an accurate and timely overview (Milner, 1983).

Any alternative (*in vitro* or *in vivo*) test for sensitization will need to be evaluated against a battery of “known” sensitizing compounds. The Consumer Product Safety Commission in 1977 proposed such a battery, which is shown in Table 27.13. This has not yet been done for any of the proposed systems. Owing to the complexity of the system involved, it is unlikely that a suitable *in vitro* replacement system will be available soon.

Gad (1988b) published comparative data on multiple animal and human test system data for some 72 materials. Such a database should be considered for the development and evaluation of new test systems.

Troese et al. (2012), Natsch et al. (2013), Nukada et al. (2013), and Bauch et al. (2012) have proposed a scheme to combine multiple *in vitro* methods to serve as an effective means of evaluating for skin sensitization, and these are being utilized as screening tools by pharmaceutical companies.

27.6.5 Phototoxicity and Photosensitization

The Daniel test for phototoxicity (also called photoirritant contact dermatitis) utilizes the yeast *Candida albicans* as a test species and has been in use for more than 20 years (Daniel, 1965). The measured end point is simply cell death. The test is simple to perform and cheap but does not reliably predict the phototoxicity of all classes of compounds (e.g., sulfanilamide). Test systems utilizing bacteria have been suggested as alternatives over the last 10 years (Harter et al., 1976; Ashwood-Smith et al., 1980) for use in predicting the same end point.

Most recently, ICI has conducted studies on an *in vitro* phototoxicity assay which involves using three cultured cell lines: the A431 human epidermal cell line (a derived epidermal carcinoma), normal HEKs (a primary cell line derived from cosmetic surgery), and the 3T3 Swiss mouse fibroblast cell line. The protocol for this assay involves subculturing the particular cell type into microtiter tissue culture grade plates and incubating them over a period of 24 h. Following incubation, the cultures are exposed to the test compound at a concentration predetermined as nontoxic. After a 4 h exposure to the compound, the cell cultures are exposed to either UV A (320–400 nm) or UV A/B (280–400 nm) radiation for varying lengths of time. The degree of enhanced toxicity effected by either UV A or UV A/B radiation in the presence of the test compound relative to the control is assessed, using the 3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide (MTT) assay. MTT undergoes a reduction reaction which is specific to mitochondrial dehydrogenases in viable cells. Work on validation of this test using 30 compounds of known phototoxic potential has shown a high degree of correlation between *in vitro* and

in vivo results. Jackson and Goldner (1989) have described several other *in vitro* assay systems for this end point.

The area of development of *in vitro* photosensitization assays has been a very active one, as the review of McAuliffe et al. (1986) illustrates. Such tests have focused on being able to predict the photosensitizing potential of a compound and variously employed cultured mammalian cell lines, red blood cells, microorganisms, and biochemical reactions. McAuliffe's group has developed and proposed a test that measures the incorporation of tritiated thymidine into human peripheral blood mononuclear cells as a predictive test (Morison et al., 1982). They claim to have internally validated the test using a battery of known photosensitizers.

Bockstahler et al. (1982) have developed and proposed another *in vitro* test system which uses the responses of two *in vitro* mammalian virus–host cell systems to the photosensitizing chemicals proflavine sulfate and 8-methoxypsoralen (8-MOP) in the presence of light as a predictive system. They found that infectious simian virus 40 (SV40) could be induced from SV40-transformed hamster cells by treatment with proflavine plus visible light or 8-MOP plus near-UV radiation. The same photosensitizing treatments inactivated the capacity of monkey cells to support the growth of herpes simplex virus. SV40 induction and inactivation of host cell capacity for herpesvirus growth might be useful as screening systems for testing the photosensitizing potential of chemicals. Advantages (ease and speed of conduct) and disadvantages (use of potentially infective agent and the limited range of compounds evaluated to date) were found to be associated with both of these test systems.

27.6.6 Developmental Toxicity

The area of developmental toxicity actually is one of the earliest to have alternative models suggested for it and has one of the most extensive and oldest literatures. This is, of course, partly owing to such models originally being used to elucidate the essential mechanisms and process of embryogenesis.

Because of the complicated and multiphasic nature of the developmental process, it has not been proposed that any of these systems be definitive tests but rather that they serve as one form or another of a screen. As such, these test systems would either preclude or facilitate more effective full-scale evaluation in one or more of the traditional whole-animal test protocols.

The literature and field are much too extensive to review comprehensively here. There are a number of extensive review articles and books on the subject (Saxen and Saksela, 1971; Wilson, 1978; Clayton, 1979; Kochhar, 1981; Homburger and Goldberg, 1985; Faustman, 1988; Daston and D'Amato, 1989), which should be consulted by those with an in-depth interest.

The existing alternative test systems fall into six broad classes: (i) lower organisms, (ii) cell culture systems,

(iii) organ culture systems, (iv) submammalian embryos, (v) mammalian embryos, and (vi) others.

Table 27.14 provides an overview of the major representatives of these six groups, along with at least one basic reference to the actual techniques involved and the system components for each.

The comparative characteristics of these different classes of test systems are presented in Table 27.15. The key point is that these systems can be used for a wide range of purposes, only one of which is to screen compounds to determine the degree of concern for developmental toxicity.

The utility of these systems for screening is limited by the degree of dependability in predicting effects primarily in people and secondarily in the traditional whole-animal test systems. Determining the predictive performance of alternative test systems requires the evaluation of a number of compounds for which the “true” (human) effect is known. In 1983 a consensus workshop generated a so-called “gold standard” set of compounds of known activity (Smith et al., 1983). The composition of this list has been open to a fair degree of controversy over the years (Flint, 1989; Johnson, 1989; Johnson et al., 1989). However, an agreed-upon “gold standard” set of compounds of known activity is an essential starting point for the validation of any single test system or battery of test systems because of the multitude of mechanisms for developmental toxicity. It is unlikely that any one system will be able to stand in place of Segment II studies in two species, much less to accurately predict activity in humans. Their use as general screens or as test systems for compounds with little potential for extensive or intended human exposure will, however, probably be appropriate.

27.6.7 Target Organ Toxicity Models

This final model review section addresses perhaps the most exciting potential area for the use of *in vitro* models—as specific tools to evaluate and understand discrete target organ toxicities. Here the presumption is that there is reason to believe (or at least suspect) that some specific target organ (nervous system, lungs, kidney, liver, heart, etc.) is or may be the most sensitive site of adverse action of a systemically absorbed agent. From this starting point, a system that is representative of the target organ's *in vivo* response would be useful in at least two contexts. New developments range from specific cell-based systems to “organ-on-a-chip” technologies.

First, as with all the other end points addressed in this chapter, a target organ-predictive system could serve as a predictive system (in general, a screen) for effects in intact organisms, particularly man. As such, the ability to identify those agents with a high potential to cause damage in a specific target organ at physiological concentrations would be extremely valuable.

The second use is largely specific to this set of *in vitro* models. This is to serve as tools to investigate, identify,

TABLE 27.14 Alternative Developmental Toxicity Test Systems

| Category | Test System | Model | References |
|-------------------------|---|--|--|
| I: Lower organisms | Sea urchins | Organism | Kotzin and Baker (1972) |
| | <i>Drosophila</i> | Intact and embryonic cells | Abrahamson and Lewis (1971) |
| | Trout | (Fish species) | MacCrimmon and Kwain (1969) |
| | <i>Planaria</i> | Regeneration | Best et al. (1981) |
| | Brine shrimp | Disruption of elongation; DNA and protein levels in <i>Artemia nauplii</i> | Kerster and Schaeffer (1983), Sleet and Brendel (1985) |
| | Animal virus | Growth of poxvirus in culture | Keller and Smith (1982) |
| | Slime mold | <i>Dictyostelium discoideum</i> | Durston et al. (1985) |
| | Medaka | (Fish species) | Cameron et al. (1985) |
| | “Artificial embryo” | <i>Hydra attenuata</i> | Johnson et al. (1982) |
| | Protein synthesis of cultured cells | Pregnant mouse and chick lens epithelial cells | Clayton (1979) |
| II: Cell culture | Avian neural crest | Differentiation of cells | Sieber-Blum (1985) |
| | Neuroblastoma | Differentiation of cells | Mummery et al. (1984) |
| | Lectin-mediated attachment | Tumor cells | Braun and Horowicz (1983) |
| III: Organ culture | Frog limb | Regeneration | Bazzoli et al. (1977) |
| | Mouse embryo limb bud | Inhibition of incorporation of precursor and of DNA synthesis | Kochhar and Aydelotte (1974) |
| | Metanephric kidney organ cultures | From 11-day mouse embryos | Saxen and Saksela (1971) |
| IV: Submammalian embryo | Chick embryo | | Gebhart (1972) |
| | Frog embryo | <i>Xenopus laevis</i> | Davis et al. (1981) |
| V: Mammalian embryo | Rat embryo culture | Whole postimplantation embryos | Brown and Fabro (1981), Cockroft and Steele (1989) |
| | Chernoff | Mouse embryo short test | Chernoff and Kavlock (1980) |
| | “Micromass cultures” | Rat embryo midbrain and limb | Flint and Orton (1984) |
| VI: Other | Structure–activity relationships (SARs) | Mathematical correlations of activity with structural features | Enslein et al. (1983a), Gombar et al. (1991) |

TABLE 27.15 Developmental Toxicity Test System Considerations

| Possibility | <i>In Vivo</i> | Organ Culture | Cell Culture | Lower Organisms | Mammalian Embryo Culture | Submammalian Embryos | Other |
|---|----------------|---------------|--------------|-----------------|--------------------------|----------------------|-------|
| To study maternal and organ factors | Yes | No | No | No | No/yes | No/yes | NA |
| To study embryogenesis as a whole | Yes | No | No | No | Yes | Somewhat | NA |
| To eliminate maternal confounding factors (nutrition) | No | Yes | Yes | No | Yes | Yes | NA |
| To eliminate placental factors (barrier differences) | No | Yes | Yes | No | Yes | No | NA |
| To study single morphogenetic events | Difficult | Yes | No | Maybe | Yes | Yes | NA |
| To create controllable, reproducible conditions | Difficult | Yes | Yes | Yes | Yes | Yes | NA |
| For exact exposure and timing | Difficult | Yes | Yes | Yes | Yes | Yes | NA |
| For microsurgical manipulations | Difficult | Yes | No | Maybe | Yes | Yes | NA |
| For continuous registration of the effects | Difficult | Yes | Yes | No | Yes | Yes | NA |
| To collect large amounts of tissue for analysis | Yes | Difficult | Yes | No | Yes | No | NA |
| To use human embryonic tissue for testing | No | Yes | Yes | No | No | No | NA |
| Screening | Expensive | Yes | Yes | Yes | Yes | Yes | Yes |

NA, not available.

and/or verify the mechanisms of action for selective target organ toxicities. Such mechanistic understandings then allow for one to know whether such toxicities are relevant to man (or to condition of exposure to man), to develop means either to predict such responses while they are still

reversible or to develop the means to intervene in such toxicosis (i.e., first aid or therapy), and finally to potentially modify molecules of interest to avoid unwanted effects while maintaining desired properties (particularly important in drug design).

In the context of these two uses, the concept of a library of *in vitro* models (Gad, 1989c, 1996) becomes particularly attractive. If one could accumulate a collection of “validated,” operative methodologies that could be brought into use as needed (and put away, as it were, while not being used), this would represent an extremely valuable competitive tool. The question becomes one of selecting which systems/tools to put into the library and how to develop them to the point of common utility.

Additionally, one must consider what forms of markers are to be used to evaluate the effect of interest. Initially, such markers have been exclusively either morphological (in that there is a change in microscopic structure), observational (is the cell/preparation dead or alive or has some gross characteristic changed?), or functional (does the model still operate as it did before?). Recently it has become clear that more sensitive models do not just generate a single end point type of data but rather a multiple set of measures which in aggregate provide a much more powerful set of answers.

There are multiple approaches to *in vitro* target organ models.

The first and oldest is that of the isolated organ preparation. Perfused and superfused tissues and organs have been used in physiology and pharmacology since the late nineteenth century. There is a vast range of these available, and a number of them have been widely used in toxicology (Mehendale (1989) presents an excellent overview). Almost any end point can be evaluated in most target organs (the CNS being a notable exception), and these are closest to the *in vivo* situation and therefore generally the easiest to extrapolate or conceptualize from. Those things that can be measured or evaluated in the intact organism can largely also be evaluated in an isolated tissue or organ preparation. However the drawbacks or limitations of this approach are also compelling.

An intact animal generally produces one tissue preparation. Such a preparation is viable generally for a day or less before it degrades to the point of losing utility. As a result, such preparations are useful as screens only for agents that have rapidly reversible (generally pharmacological or biomechanical) mechanisms of action. They are superb for evaluating mechanisms of action at the organ level for agents that act rapidly—but not generally for cellular effects or for agents that act over a course of more than a day.

The second approach is to use tissue or organ culture. Such cultures are attractive, owing to maintaining the ability for multiple cell types to interact in at least a near-physiological manner. They are generally not as complex as perfused organs, but are stable over a longer period of time, increasing their utility as screens somewhat. They are truly a middle ground between perfused organs and cultured cells. Only for relatively simple organs (such as the skin and bone marrow) are good models which perform in a manner representative of the *in vitro* organ available.

The third and most common approach remains that of cultured cell models. These can be either primary or transformed (immortalized) cells, but the former have significant advantages in use as predictive target organ models. Such cell culture systems can be utilized to identify and evaluate interactions at the cellular, subcellular, and molecular level on an organ- and species-specific basis (Acosta et al., 1985). The advantages of cell culture are that single organisms can generate multiple cultures for use, that these cultures are stable and useful for protracted periods of time, and that effects can be studied very precisely at the cellular and molecular levels. The disadvantages are that isolated cells cannot mimic the interactive architecture of the intact organ and will respond over time in a manner that becomes decreasingly representative of what happens *in vivo*. An additional concern is that, with the exception of hepatocyte cultures, the influence of systemic metabolism is not factored in unless extra steps are taken. Stamatii et al. (1981) and Tyson and Stacey (1989) present some excellent reviews of the use of cell culture in toxicology. Any such cellular systems would be more likely to be accurate and sensitive predictors of adverse effects if their function and integrity were evaluated while they were operational. For example, cultured nerve cells should be excited while being exposed and evaluated.

Finally, we now have “organ-on-a-chip” systems, which combine multiple cell systems (from one or more organs) to provide a more relevant to actual physiological function test system (Kim et al., 2015).

A wide range of target organ-specific models have already been developed and used. Their incorporation into a library-type approach requires that they be evaluated for reproducibility of response, ease of use, and predictive characteristics under the intended conditions of use. These evaluations are probably at least somewhat specific to any individual situation. Tables 27.16, 27.17, 27.18, 27.19, 27.20, and 27.21 present overviews of representative systems for a range of target organs: respiratory, nervous, renal, cardiovascular, hepatic, pancreatic, gastrointestinal, and reticuloendothelial. These tables do not mention any of the new coculture systems in which hepatocytes are “joined up” in culture with a target cell type to produce a metabolically competent cellular system.

For more complex organ systems, great progress has been made in deriving an integrated approach to utilizing several types of *in vitro* model data. McKim (2012) has provided a leading edge example in preproposing such an approach for the leading cause of drug failure due to safety issues—liver-specific toxicity (particularly DLT—delayed liver toxicity) which starts with understanding the structure of the liver and the types (and functions) of different cell types in the organ.

Additionally, there are ongoing efforts to use human tissue as models to evaluate the safety and efficacy of new drugs (Coleman, 2011).

TABLE 27.16 Representative *In Vitro* Test Systems for Respiratory System Toxicity

| System | End Point | Evaluation | References |
|--|---|---|--|
| Isolated perfused rat and rabbit lungs (S) | Damage markers: exudate of hormones | Correlation with results <i>in vivo</i> | Anderson and Eling (1976), Roth (1980), Mehendale (1989) |
| Alveolar macrophages (S) | Cytotoxicity: as a predictor of fibrogenicity | Correlation with <i>in vivo</i> fibrogenicity across a broad range of compounds | Reiser and Last (1979) |
| Lung organ culture (M, S) | Morphological: structure and macromolecular composition | Proposed from prior experience in pharmacology | Placke and Fisher (1987) |
| Hamster lung culture (M) | Morphological: structure and cell death | Correlation of <i>in vivo</i> effects of cigarette smoke | Stammati et al. (1981) |

Letters in parentheses indicate primary employment of system: S=screening system; M=mechanistic tool.

TABLE 27.17 Representative *In Vitro* Test Systems for Neurotoxicity

| System | End Point | Evaluation | References |
|--|--|--|---|
| Perfused rat phrenic nerve—hemidiaphragm (M) | Functional: release of ACh, conduction velocities, and muscle response | Correlates with <i>in vivo</i> effects of trialkyltins | Bierkamper (1982) |
| Primary rat cerebral cells (S) | Observational: cell growth and differentiation | Cell diameter and outgrowth | Hooisma (1982) |
| Primary rat tissue culture (S) | Functional: receptor–ligand binding | Binding rates | Bondy (1982), Volpe et al. (1985) |
| Organotypic neural cultures (S) | Functional: electrophysiological and pharmacological properties | Correlation with <i>in vivo</i> results for a range of known active agents | Spencer et al. (1986), Kontur et al. (1987) |
| Isolated perfused brain (M) | Functional: biochemical and electrophysiological | Unknown | Mehendale (1989) |
| Cultured mouse otocyst (M) | Morphological | Unknown—a tool for potentially evaluating ototoxins | Harpur (1988) |

Letters in parentheses indicate primary employment of system: S=screening system; M=mechanistic tool.

TABLE 27.18 Representative *In Vitro* Test Systems for Renal Toxicity

| System | End Point | Evaluation | References |
|-----------------------------------|--|---|---|
| Rat proximal tubular cells (S) | Functional: 3- <i>O</i> -methylglucose uptake or organic ion transport | Correlation with effects of known nephrotoxins | Boogaard et al. (1989) |
| Rat cortical epithelial cells (S) | Functional: biochemical | Good correlation with <i>in vivo</i> findings for nephrotoxic metals and acetaminophen | Smith et al. (1986, 1987), Rylander et al. (1987) |
| Isolated perfused kidney (M) | Functional: biochemical and metabolic | Correlation with <i>in vivo</i> findings for some nephrotoxins | Mehendale (1989) |
| Renal slices (S, M) | Morphological Full range of functional (biochemical and metabolic) | Correlation with <i>in vivo</i> findings for a range of nephrotoxins. Still allows evaluation of a degree of cell-to-cell and nephron-to-nephron interactions | Smith et al. (1988) |

Letters in parentheses indicate primary employment of system: S=screening system; M=mechanistic tool.

TABLE 27.19 Representative *In Vitro* Test Systems for Cardiovascular Toxicity

| System | End Point | Evaluation | References |
|--|---|---|-------------------------|
| Coronary artery smooth muscle cells (S) | Morphological evaluation—vacuole formation | Correlates with <i>in vivo</i> results | Ruben et al. (1984) |
| Isolated perfused rabbit or rat heart (M, S) | Functional: operational, electrophysiological, biochemical, and metabolic | Long history of use in physiology and pharmacology | Mehendale (1989) |
| Isolated superfused atrial and heart preparations (S, M) | Functional: operational and biochemical | Correlation with <i>in vivo</i> findings for antioxidants | Gad et al. (1977, 1979) |

Letters in parentheses indicate primary employment of system: S=screening system; M=mechanistic tool.

TABLE 27.20 Representative *In Vitro* Test Systems for Hepatic Toxicity

| System | End Point | Evaluation | References |
|-----------------------------|---|---|---|
| Primary hepatocytes (S, M) | Multiple: <ul style="list-style-type: none"> • Biotransformation • Genotoxicity • Peroxisome proliferation • Biliary dysfunction • Membrane damage • Ion regulation • Energy regulation • Protein synthesis | NA | Tyson and Stacey (1989), ^a Stammati et al. (1981) |
| Hamster hepatocytes (S) | Functional: biochemical | Correlates with <i>in vivo</i> effects of acetaminophen | Harman and Fisher (1983) |
| Rat liver slices (S) | Functional: alterations in ion content, leakage of damage markers, and changes in biosynthetic capability Morphological: histopathological evaluation | Rank correlation with <i>in vivo</i> findings for a wide range of chemicals | Gandolfi et al. (1989), Adams (1995), Fisher et al. (1995) |
| Isolated perfused liver (M) | Functional: biochemical and metabolic | Correlation with <i>in vivo</i> findings for a wide range of chemicals | Mehendale (1989), Wyman et al. (1995) |

Letters in parentheses indicate primary employment of system: S=screening system; M=mechanistic tool. NA=not available.

^aTyson and Stacey estimated in 1989 that there were 800 unpublished studies of a toxicological nature on cultured hepatocytes.

TABLE 27.21 Representative *In Vitro* Test Systems for Other Target Organ Studies

| Organ | System | End Point | Evaluation | References |
|---------------------|---|---|---|------------------------|
| Pancreas | Isolated perfused intestines (M) | Functional: biochemical and metabolic | Correlation with <i>in vivo</i> findings for methylprednisolone | Mehendale (1989) |
| GI Tract | Isolated perfused intestines (M) | Functional: biochemical and metabolic | Limited | Mehendale (1989) |
| Reticuloendothelial | Erythrocytes (S) | Observational: cytotoxicity Functional: inhibition of colony formation | Correlation with hemolytic effects | Stammati et al. (1981) |
| Testicular | Sertoli and germ cell cultures (S) | Observational: cytotoxicity Functional: steroid and hormone production | Correlation with <i>in vivo</i> effects for phytholate esters and glycol ethers | Garside (1988) |
| Adrenal gland | Primary adrenocortical ADC cell cultures (S, M) | Functional: cortisol production | Correlation with <i>in vivo</i> effects for three known ADC toxicants | Wolfgang et al. (1994) |
| Thyroid | Cultured thyroid cells (S, M) | Functional: biochemical and metabolic | Correlation with <i>in vivo</i> findings for a wide range of agents with thyroid-specific toxicity; evaluation against "negative" compounds not significant | Brown and Fabro (1981) |

Letters in parentheses indicate primary employment of system: S=screening system; M=mechanistic tool.

27.7 *IN SILICO* METHODS

The concept that the biological activity of a compound is a direct function of its chemical structure is now at least a century old (Crum-Brown and Fraser, 1869). During most of this century, the development and use of SARs were

the domain of pharmacology and medicinal chemistry. These two fields are responsible for the beginnings of all the basic approaches in SAR work, usually with the effort being called drug design. An introductory medicinal chemistry text (such as Foye et al. (1995)) is strongly recommended as a starting place for SAR. Additionally, Burger's Medicinal

Chemistry (Abraham and Rotella, 2010), with its excellent overview of drug structures and activities, should enhance at least the initial stages of identifying the potential biological actions of *de novo* compounds (drug or otherwise) using pattern recognition approach.

SAR methods have become a legitimate and useful part of toxicology since the mid-1970s and an extremely important part in the twenty-first century as Cramer et al. (1978) published the first modern approach (intended primarily to prioritize structures for toxicological evaluation). These methods are various forms of mathematical or statistical models which seek to predict the adverse biological effects of chemicals based on their structure. The prediction may be either of a qualitative (carcinogen/noncarcinogen) or quantitative (LD_{50}) nature, with the second group usually being denoted as quantitative structure–activity relationship (QSAR) models. It should be obvious at the onset that the basic techniques utilized to construct such models are those which are termed modeling and extrapolation (now called “expert systems” based on of structures and substructures) or reduction of dimensionality methods (now called “statistical SAR systems”), using a set of structures with known activity and based on proximity of overall structure descriptors in a multidimensional space. Models may also be characterized as global (covering a wide range of structural features) or local (which cover only a narrow domain of structures but tend to provide much higher accuracy of prediction within that range).

For toxicology, SARs have to be a significant and important range of uses. These can all be generalized as identifying potentially toxic effects or restated as three main uses:

1. For the selection and design of toxicity tests to address end points of possible concern.
2. If a comprehensive or large testing program is to be conducted, SAR predictions can be used to prioritize the tests, so that outlined questions (the answers to which might preclude the need to do further testing) may be addressed first.
3. As an alternative to testing at all. Though in general it is not believed that the state of the art for SAR methods allows such usage in certain special cases (such as selecting which of the several alternative candidate compounds to develop further and then set).

Having already classified SAR methods into qualitative and quantitative, it should also be pointed out that both of these can be approached on two different levels. The first is on a local level, where prediction of activity (or lack of activity) is limited to other members of a congeneric series or structural near neighbors. The accuracy of predictions via this approach is generally greater but is of value only if one has sufficient information on some of the structures within a series of interest.

The ICH (ICH M7) and FDA now accept QSAR models in place of actual mutagenicity testing for assessing mutagenicity of impurities and of drug and device extractables and leachables.

27.8 THE FINAL FRONTIER AND BARRIER: REGULATORY ACCEPTANCE

While in Europe there are numerous validated (by ECVAM standards) *in vitro* alternatives for personal care products and industrial and agricultural chemicals (though only two—the LAL assay for pyrogenicity and the 3T3 cell assay for phototoxicity—have acceptance for use with pharmaceutical regulators) with regulatory acceptance (as summarized in Table 27.22), this is not the case in the United States. Indeed, only four alternatives have been approved by the ICVAM process to date (Gaul, 2008). Likewise, while most primary biocompatibility testing performed in accordance with ICH (for drugs) or ISO 10993 standards (for medical devices) is performed using *in vitro* models, no new ones have been added or served to replace existing *in vivo* methods. Indeed, such new tests that have been added are *in vivo*. Though there is considerable research continuing on the development (and putatively on the validation) of many more methods (Kuehn, 2008), progress in obtaining regulatory acceptance of such methods to take the place of animal test methods is very slow, and those methods that are accepted are limited in scope screens for eye and skin irritation pyrogenicity, phototoxicity, and genotoxicity.

Thus regulatory acceptance is, indeed, as it has been since 1990, the final frontier and true obstacle to any further significant reduction in animal use.

27.9 SUMMARY

The tools are currently at hand (or soon will be) to provide the practicing toxicologist with unique opportunities both for identifying potentially toxic compounds in a much more rapid and efficient manner than before and for teasing apart the mechanisms underlying such toxicities on an integrated basis (from the level of the molecule to that of the intact organism). The *in vitro* systems overviewed here, once understood (by investigators and regulators) in how they function and fail (just as *in vivo* systems have come to be understood), will allow this to happen while reducing the need to have recourse to intact mammalian test systems. However, the intact animal models—and, indeed, man for pharmaceuticals—will still be an essential element in the safety assessment armamentarium for the foreseeable future, though increasingly supplemented by alternatives that have had time and utilization enough to establish their dependability.

TABLE 27.22 Status of Nonanimal Methods That Have Regulatory Standing (Acceptance)

| Test Method | Test System | End Point | OECD/TG or Other Regulatory Comments |
|---|---|--------------------------------------|---|
| <i>In vitro</i> test methods for which there are OECD health effects test guidelines (including draft guidelines under review for acceptance) can be found at http://www.oecd.org/home/ | | | |
| Transcutaneous electrical resistance (TER) test | Monitors changes in the electrical resistance as a measure of loss of corneum integrity and barrier function; involves skin disks from euthanized rats | Skin corrosion (topical agents) | TG 430 |
| Human skin models (EpiDerm™ and EPISKIN™) | Reconstructed human epidermal equivalent (commercial system) used to assess cell viability, involving the MTT reduction test | Skin corrosion (topical agents) | TG 431 |
| 3T3 NRU phototoxicity test | BALB/c 3T3 (murine) cell line cytotoxicity based on neutral red uptake to measure cell viability; not direct replacement alternative, as there is no <i>in vivo</i> equivalent test | Phototoxicity | TG 432/FDA guidance |
| Corrositex™ membrane barrier test | An artificial barrier system coupled to a pH-based chemical detection system | Skin corrosion (topical agents) | Draft TG 435 |
| Bacterial reverse mutation test (Ames) | Revertant bacteria detected by their ability to grow in the absence of the amino acid | Genotoxicity | TG471/ICH, ISO, FDA guidance |
| <i>In vitro</i> mammalian chromosomal aberration test | Microscopic detection of chromosomal damage to cells in culture | Genotoxicity | TG473/ICH, ISO, FDA guidance |
| <i>In vitro</i> mammalian cell gene mutation test | Functional bioassays to monitor mutations in enzyme-encoding genes | Genotoxicity | TG476 |
| Sister chromatid exchange assay | Cells in culture are examined after two rounds of division by metaphase arrest and chromosomal preparation; chromatid exchange is monitored by microscopy | Genotoxicity | TG479/ICH, ISO, FDA guidance |
| Gene mutation assay in yeast | <i>Saccharomyces cerevisiae</i> exposed to the test substance are grown under different culture conditions used to monitor mutagenic potential (cf. Ames test) | Genotoxicity | TG 480 |
| Mitotic recombination assay in yeast | Crossover or gene conversion following exposure of yeast to the test substance; relies on different growth requirements of mutated and wild-type yeast strains | Genotoxicity | TG 481 |
| Unscheduled DNA synthesis in mammalian cells | Measures the DNA repair synthesis after deletions caused by the test substance based on the incorporation of radioactive nucleotides into the newly synthesized DNA | Genotoxicity | TG 482/ICH, ISO, and FDA guidance |
| Bovine corneal opacity and permeability (BCOP) | See eye irritation table | Ocular irritation | FDA guidance |
| QSAR mutagenicity models | One expert-/rule-based and one statistical-based system | Mutagenicity | ICH M7(R2) FDA, ICH countries |
| <i>In vitro</i> micronucleus test | Cell-based assay; supplement to TG 474 (<i>in vivo</i> micronucleus test); detection of chromosomal damage and formation of micronuclei in interphase | Genotoxicity | Draft TG 487/ICH guidance |
| Sex-linked recessive lethal test | <i>Drosophila</i> are exposed to the test substance. Germline transmission of mutations is monitored through two successive generations | Reproductive toxicity | TG 477 |
| <i>Validated methods that are yet to be introduced into regulatory use</i> | | | |
| EpiOcular™ | Human keratinocyte-derived model of the corneal epithelium barrier function | Eye irritation (topical application) | Retrospective (weight-of-evidence) validation (ECVAM) |
| <i>In vitro</i> micronucleus test | CHL/IU, CHO, SHE, or V79 cell lines are commonly used, with or without metabolic activation, to monitor damage and formation of micronuclei in interphase | Mutagenicity | Retrospective (weight-of-evidence) validation (ECVAM) |

TABLE 27.22 (Continued)

| Test Method | Test System | End Point | OECD/TG or Other Regulatory Comments |
|---|--|---------------------------|---|
| Embryonic stem cell test | 3T3 cell cytotoxicity and differentiation of embryonic stem murine cell lines used to examine teratogenic potential | Developmental toxicity | Endorsed as screening test (EU) |
| Postimplantation rat whole-embryo test | Morphological assessment of rat embryos | Developmental toxicity | Endorsed as screening test (EU) |
| Micromass test | Micromass cultures of rat limb are bud monitored for inhibition of cell proliferation and differentiation | Developmental toxicity | Endorsed as screening test (EU) |
| <i>Methods undergoing validation</i> | | | |
| EPISKIN | Reconstructed human skin system used with MTT assay to monitor barrier function | Skin irritation | Report stage in EU |
| EpiDerm | Similar to EPISKIN | Skin irritation | Report stage in EU |
| <i>Prevalidated methods</i> | | | |
| SkinEthic eye model | Epithelial corneal cell line used for cytotoxicity testing based on the MTT reduction assay | Eye irritation | Appraisal stage in EU |
| <i>Methods Undergoing Development, Prevalidation, or Evaluation</i> | | | |
| Tissue culture models | Neutral red release and silicon microphysiometry or fluorescein leakage bioassays with human keratinocytes and MDCK cells, respectively; red blood cell (RBC) hemolysis test | Eye irritation | Being reviewed by ICCVAM for possible retrospective (weight-of-evidence) validation |
| Organotypic models | Bovine corneal opacity and permeability (BCOP) assay, with postmortem corneas; hen's egg test on the chorioallantoic membrane (HET-CAM assay); isolated rabbit and chicken eye tests (IRE and ICE) | Eye irritation | Being reviewed by ICCVAM for possible retrospective (weight-of-evidence) validation |
| Modified Leydig cell line | Analysis of progesterone production as a measure of the test substance effects on steroid hormone production | One-/two-generation study | For use as part of test battery |
| Testis slices | Assessment of steroid production capacity of the Leydig cells upon exposure of <i>ex vivo</i> rat tissue to toxicants | One-/two-generation study | For use as part of test battery |
| Human adrenocortical carcinoma cell line | Assay to allow entire steroid pathway effects to be mapped | One-/two-generation study | For use as part of test battery |
| Placental microsomal aromatase assay | Monitors the ability of substances to affect steroid production; a subcellular microsomal assay is used industrially | One-/two-generation study | For use as part of test battery |

This is a comprehensive list of methods that have been validated or that are at various stages of development for toxicity testing (OECD, 2005; Kandorona and Letasiova, 2011).

REFERENCES

- Abraham DJ, Rotella DP. (2010) *Burgers Medicinal Chemistry*. John Wiley & Sons, Inc., Hoboken, NJ.
- Abrahamson S, Lewis EB. (1971) The detection of mutations in *Drosophila melanogaster*. In: *Chemical Mutagens. Principles and Methods of Their Detection* (Hollaender A, Ed.), Vol. 2. Plenum Press, New York, pp. 461–488.
- Acosta D, Sorensen EMB, Anuforo DC, Mitchell DB, Ramos K, Santone KS, Smith MA. (1985) An *in vitro* approach to the study of target organ toxicity of drugs and chemicals. *In Vitro Cell. Dev. Biol.* 21:495–504.
- Adams PE. (1995) *In Vitro* Methods to Study Hepatic Drug Metabolism, *Emphasis*, Corning Hazleton, Madison, WI 6(2).
- Aeschbacher M, Reinhardt CA, Zbinden G. (1986) A rapid cell membrane permeability test using fluorescent dyes and flow cytometry. *Cell Biol. Toxicol.* 2:247.
- Anderson MW, Eling TE. (1976) Studies on the uptake, metabolism, and release of endogenous and exogenous chemicals by the use of the isolated perfused lung. *Environ. Health Perspect.* 16:77–81.
- Arancioglu S, Ulker OC, Karakaya A. (2015) Utilization of the *ex vivo* LLNA: BrdU-ELISA to distinguish the sensitizers from irritants in respect of 3 end points—Lymphocyte proliferation, ear swelling, and cytokine profiles. *Int. J. Toxicol.* 34(1):24–30.
- Ashwood-Smith MJ, Poulton GA, Barker M, Midenberger M. (1980) 5-Methoxypsoralen, an ingredient in several suntan preparations, has lethal mutagenic and clastogenic properties. *Nature.* 285:407–409.

- Babich H, Martin-Alguacil N, Borenfreund E. (1989) Comparisons of the cytotoxicities of dermatotoxicants to human keratinocytes and fibroblasts *in vitro*. In: *In Vitro Toxicology: New Directions* (Goldberg AM, Ed.). Mary Ann Liebert, New York, pp. 153–167.
- Balls M, Horner SA. (1985) The FRAME interlaboratory program on *in vitro* cytotoxicology. *Food Chem. Toxicol.* 23:205–213.
- Barnard ND. (1989) A Draize Alternative. *The Animal's Agenda*, 6:45.
- Bartnik FG, Pittermann WF, Mendorf N, Tillmann U, Kunstler K. (1989) *Skin organ culture for the study of skin irritancy*. Fifth International Congress of Toxicology, July 16–21, 1989, Brighton, England.
- Bauch C, Kolle SN, Ramirez T, Eltze T, Fabian E, Mehling A, Teubner W, van Ravenzwaay B, Landsiedel R. (2012) Putting the parts together, combining *in vitro* methods to test for skin sensitizing potentials. *Regul. Toxicol. Pharmacol.* 63:489–504.
- Bazzoli AS, Manson J, Scott WJ, Wilson JG. (1977) The effects of thalidomide and two analogues on the regenerating forelimb of the newt. *J. Embryol. Exp. Morphol.* 41:125–135.
- Bell E, Parenteau NL, Haimes HB, Gay RJ, Kemp PD, Fofonoff TW, Mason VS, Kagan DT, Swiderek M. (1988) Testskin: a hybrid organism covered by a living human skin equivalent designed for toxicity and other testing. In: *Progress in In Vitro Toxicology* (Goldberg AM, Ed.). Mary Ann Liebert, New York, pp. 15–25.
- Benassi CA, Angi MR, Salvaloi L, Bettero A. (1986) Histamine and leukotriene C4 release from isolated bovine sclerachoroid complex: a new *in vitro* ocular irritation test. *Chim. Agg.* 16:631–634.
- Berghmans S, Jette C, Langenau D, Hsu K, Stewart R, Look T, Kanki JP. (2005) Making waves in cancer research: new models in the zebrafish. *Biotechniques*. 39(2):227–237.
- Best JB, Morita M, Ragin J, Best J Jr. (1981) Acute toxic responses of the freshwater planarian, *Dugesia dorotocephala*, to methylmercury. *Bull. Environ. Contam. Toxicol.* 27:49–54.
- Bierkamper GG. (1982) *In vitro* assessment of neuromuscular toxicity. *Neurobehav. Toxicol. Teratol.* 4:597–604.
- Birnbaum LS, Stokes WS. (2010) Safety testing moving towards alternative methods. *Environ. Health Perspect.* 118:A12–A13.
- Bockstahler LE, Coohill TP, Lytle CD, Moore SP, Cantwell JM, Schmidt BJ. (1982) Tumor virus induction and host cell capacity inactivation: possible *in vitro* test for photosensitizing chemicals. *J. Natl. Cancer Inst.* 69:183–187.
- Bondy SC. (1982) Neurotransmitter binding interactions as a screen for neurotoxicity. In: *Mechanisms of Actions of Neurotoxic Substances* (Prasad KN and Vernadakis A, Eds.). Raven Press, New York, pp. 25–50.
- Boo JD, Knight A. (2009) Replacing reducing and refining procedures in animals research. *J. Oral Tissue Eng.* 6:215–222.
- Boogaard PJ, Mulder GJ, Nagelkerke JF. (1989) Isolated proximal tubular cells from rat kidney as an *in vitro* model for studies on nephrotoxicity. *Toxicol. Appl. Pharmacol.* 101:135–157.
- Borenfreund E, Puerer JA. (1984) A simple quantitative procedure using monolayer cultures for cytotoxicity assays (HTD/NR-NE). *J. Tissue Cult. Methods*. 9:7–10.
- Boyce ST, Hansbrough JF, Norris DA. (1988) Cellular responses of cultured human epidermal keratinocytes as models of toxicity to human skin. In: *Progress in In Vitro Toxicology* (Goldberg AM, Ed.). Mary Ann Liebert, New York, pp. 27–37.
- Braun AG, Horowicz PB. (1983) Lectin mediated attachment assays for teratogens. Results with 32 pesticides. *J. Toxicol. Environ. Health.* 11(2):275–286.
- Brown NA, Fabro S. (1981) Quantitation of rat embryonic development *in vitro*: a morphological scoring system. *Teratology*. 24:65–78.
- Brown S, Templeton L, Prater DA, Potter CJ. (1989) Use of an *in vitro* haemolysis test to predict tissue irritancy in an intramuscular formulation. *J. Parenter. Sci. Technol.* 43:117–120.
- Burka JF, Ali M, McDonald JWD, Paterson NAM. (1981) Immunological and non-immunological synthesis and release of prostaglandins and thromboxanes from isolated guinea pig trachea. *Prostaglandins*. 2:683–690.
- Burton ABG, York M, Lawrence RS. (1981) The *in vitro* assessment of severe eye irritants. *Food Cosmet. Toxicol.* 19:471–480.
- Cameron IL, Lawrence WC, Lum JB. (1985) Medaka eggs as a model system for screening potential teratogens. *Prog. Clin. Biol. Res.* 163C:239–243.
- Chan KY. (1985) An *in vitro* alternative to the Draize test. In: *In Vitro Toxicology. Alternative Methods in Toxicology* (Goldberg AM, Ed.), Vol. 3. Mary Ann Liebert, New York, pp. 405–422.
- Chernoff N, Kavlock RJ. (1980) A potential *in vivo* screen for the determination of teratogenic effects in mammals. *Teratology*. 21:33A–34A.
- Chiu LL, Cunningham LL, Raible DW, Rubel EW, Ou HC. (2008) Using the zebrafish lateral line to screen for ototoxicity. *J. Assoc. Res. Otolaryngol.* 9(2):178–190.
- Clayton RM. (1979) Protein synthesis of cultured cells. In: *Alternatives in Drug Research* (Rowan AN and Stratmann CJ, Eds.). Macmillan Press, London, pp. 153–173.
- Clive D, Johnson K, Spector J, Batson A, Brown M. (1979) Validation and characterization of the L5178Y/TK mouse lymphoma mutagen assay system. *Mutat. Res.* 59:61–108.
- Cockroft DL, Steele CE. (1989) Postimplantation embryo culture and its application to problems in teratology. In: *In Vitro Methods in Toxicology* (Atterwill CK and Steele CE, Eds.). Cambridge University Press, New York, pp. 365–389.
- Coleman RA. (2011) Human tissue in the evaluation of safety and efficacy of new medicines: a viable alternatives to animals models. *ISRN Pharm.* 2011:8.
- Cooper JF. (1975) Principles and applications of the Limulus test for pyrogen in parenteral drugs. *Bull. Parenter. Drug Assoc.* 3:122–130.
- Cramer GM, Ford RA, Hall RL. (1978) Estimation of toxic hazard—a decision tree approach. *Food Cosmet. Toxicol.* 16:255–276.
- Crum-Brown A, Fraser TR. (1869) On the connecting between chemical constitution and physiologic action. Part II. *Trans. R. Soc. Edinb.* 25:693–739.
- Daniel F. (1965) A simple microbiological method for demonstrating phototoxic compounds. *J. Invest. Dermatol.* 44:259–263.

- Dannenber AM, Moore KG, Schofield BH, Higuchi K, Kajiki A, Au K, Pula PJ, Bassett DP. (1987) Two new *in vitro* methods for evaluating toxicity in skin (employing short-term organ culture). In: *Alternative Methods in Toxicology* (Goldberg AM, Ed.), Vol. 5. Mary Ann Liebert, New York, pp. 115–128.
- Daston GP, D'Amato RA. (1989) *In vitro* techniques in teratology. In: *Benchmarks: Alternative Methods in Toxicology* (Mehlman M, Ed.). Princeton Scientific, Princeton, NJ, pp. 79–109.
- Davis KR, Schultz TW, Dumont JN. (1981) Toxic and teratogenic effects of selected aromatic amines on embryos of the amphibian *Xenopus laevis*. *Arch. Environ. Contam. Toxicol.* 10:371–391.
- DeLeo V, Midlarsky L, Harber LC, Kong BM, Salva SD. (1987) Surfactant-induced cutaneous primary irritancy: an *in vitro* model. In: *Alternative Methods in Toxicology* (Goldberg AM, Ed.), Vol. 5. Mary Ann Liebert, New York, pp. 129–138.
- DeLeo V, Hong J, Scheide S, Kong B, DeSalva S, Bagley D. (1988) Surfactant-induced cutaneous primary irritancy: an *in vitro* model-assay system development. In: *Progress in In Vitro Toxicology* (Goldberg AM, Ed.). Mary Ann Liebert, New York, pp. 39–43.
- Drewes C, Vining E, Callahan C. (1984) Non-invasive electrophysiological monitoring: a sensitive method for detecting sublethal neurotoxicity in earthworms. *Environ. Toxicol. Chem.* 3:559–607.
- Dubin NH, De Blasi MC, Thomas CL, Wolff MC. (1984) Development of an *in vitro* test for cytotoxicity in vaginal tissue: effect of ethanol on prostanoid release. In: *Acute Toxicity Testing: Alternative Approaches* (Goldberg AM, Ed.), Vol. 2. Mary Ann Liebert, New York, pp. 127–138.
- Durston A, Van de Wiel F, Mummery C, de Loat S. (1985) *Dictyostelium discoideum* as a test system for screening for teratogens. *Teratology.* 32:21A.
- Ekwall B, Bondesson I, Castell JV, Gomez-Lechon MJ, Hellberg S, Hagberg J, Jover R, Ponsoda X, Romert L, Stenberg K, Watum E. (1989) Cytotoxicity evaluation of the first ten MEIC chemicals: acute lethal toxicity in man predicted by cytotoxicity in five cellular assays and by oral LD50 tests in rodents. *Anim. Tech. Lab. Assoc.* 17:83–100.
- Elgebaly SA, Nabawi K, Herkbert N, O'Rourke J, Kruetzer DL. (1985) Characterization of neutrophil and monocyte specific chemotactic factors derived from the cornea in response to injury. *Invest. Ophthalmol. Vis. Sci.* 26:320.
- Enslein K. (1984) Estimation of toxicology end points by structure-activity relationships. *Pharmacol. Rev.* 36:131–134.
- Enslein K, Lander TR, Strange JL. (1983a) Teratogenesis: a statistical structure-activity model. *Teratog. Carcinog. Mutagen.* 3:289–309.
- Enslein K, Lander TR, Tomb ME, Craig PN. (1983b) *A Predictive Model for Estimating Rat Oral LD50 Values*. Princeton Scientific, Princeton, NJ.
- Enslein K, Borgstedt HH, Blake BW, Hart JB. (1987) Prediction of rabbit skin irritation severity by structure-activity relationships. *In Vitro Toxicol.* 1:129–147.
- Eurell TE, Meachum SH. (1994) *In vitro* evaluation of ocular irritants using tissue isoelectric focusing protein profiles from human, rabbit, and bovine corneal specimens. *Toxicol. Methods.* 4:66–75.
- Faustman EM. (1988) Short-term tests for teratogens. *Mutat. Res.* 205:355–384.
- Firestone BA, Guy RH. (1986) Approaches to the prediction of dermal absorption and potential cutaneous toxicity. In: *In Vitro Toxicology, Alternative Methods in Toxicology* (Goldberg AM, Ed.), Vol. 3. Mary Ann Liebert, New York, pp. 526–536.
- Fisher RL, Hasal SJ, Sanuik JT, Gandolfi AJ, Brendel K. (1995) Determination of optimal incubation media and suitable slice diameters in precision-cut tissue slice culture, part 2. *Toxicol. Methods.* 5:115–130.
- Fishlock D. (1990) Survival of the Fittest Drugs. *Financial Times*, April 24, 1990:16–22.
- Flint OP. (1989) Reply to letter to the editor. *Toxicol. Appl. Pharmacol.* 99:176–180.
- Flint OP, Orton TC. (1984) An *in vitro* assay for teratogens with cultures of rat embryo midbrain and limb bud cells. *Toxicol. Appl. Pharmacol.* 76:383–395.
- Foye WO, Lenke TL, Williams DA. (1995) *Principles of Medicinal Chemistry*. Lippincott Williams and Wilkins, Philadelphia, PA.
- Frazier JM. (1988) Update: a critical evaluation of alternatives to acute ocular irritancy testing. In: *Progress in In Vitro Toxicology* (Goldberg AM, Ed.). Mary Ann Liebert, New York, pp. 67–75.
- Frazier JM. (1992) *In Vitro Toxicity Testing*. Marcel Dekker, New York.
- Frazier JM, Gad SC, Goldberg AM, McCulley JP. (1987) A critical evaluation of alternatives to acute ocular irritation testing. In: *Alternative Methods in Toxicology* (Goldberg AM, Ed.), Vol. 4. Mary Ann Liebert, New York.
- Frosch PJ, Czarnetzki BM. (1987) Surfactants cause *in vitro* chemotaxis and chemokinesis of human neutrophils. *J. Invest. Dermatol.* 88(Suppl. 3):52.
- Gad SC. (1988a) A scheme for ranking of relative potencies of dermal sensitizers based on data from several test systems. *J. Appl. Toxicol.* 8:461–467.
- Gad SC. (1988b) A tier testing strategy incorporating *in vitro* testing methods for pharmaceutical safety assessment. *Hum. Innov. Altern. Anim. Exp.* 3:75–79.
- Gad SC. (1989a) Principles of screening in toxicology: with special emphasis on applications to neurotoxicology. *J. Am. Coll. Toxicol.* 8:21–27.
- Gad SC. (1989b) Statistical analysis of screening studies in toxicology: with special emphasis on neurotoxicology. *J. Am. Coll. Toxicol.* 8:171–183.
- Gad SC. (1989c) A tier testing strategy incorporating *in vitro* testing methods for pharmaceutical safety assessment. *Hum. Innov. Altern. Anim. Exp.* 3:75–79.
- Gad SC. (1990a) Industrial applications for *in vitro* toxicity testing methods: a tier testing strategy for product safety assessment. In: *In Vitro Toxicity Testing* (Frazier J, Ed.). Marcel Dekker, New York.
- Gad SC. (1990b) Recent developments in replacing, reducing and refining animal use in toxicologic research and testing. *Fundam. Appl. Toxicol.* 15(1):8–16.
- Gad SC. (1993) *In Vitro Toxicology*. Raven Press, New York.

- Gad SC. (1996) Preclinical toxicity testing in the development of new therapeutic agents. *Scand. J. Lab. Anim. Sci.* 23:299–314.
- Gad SC. (1999) *Product Safety Evaluation Handbook*, 2nd Ed. Marcel Dekker, New York.
- Gad SC. (2000) *In Vitro Toxicology*, 2nd Ed. Taylor & Francis, New York.
- Gad SC. (2001) Alternatives to in vivo studies in toxicology. In: *General & Applied Toxicology* (Ballantyne B, Marvis T and Turner P, Eds.), 2nd Ed. Macmillan, New York, pp. 401–424.
- Gad SC. (2009) *Drug Safety Evaluation*, 2nd Ed. John Wiley & Sons, Inc., Hoboken, NJ.
- Gad SC. (2015) *Animal Models in Toxicology*, 3rd Ed. Marcel Dekker, New York.
- Gad SC, Chengelis CP. (1997) *Acute Toxicology*, 2nd Ed. Academic Press, San Diego, CA.
- Gad SC, Leslie SW, Brown RG, Smith RV. (1977) Inhibitory effects of dithiothreitol and sodium bisulfite on isolated rat ileum and atrium. *Life Sci.* 20:657–664.
- Gad SC, Leslie SW, Acosta D. (1979) Inhibitory actions of butylated hydroxytoluene (BHT) on isolated rat ileal, atrial and perfused heart preparations. *Toxicol. Appl. Pharmacol.* 48:45–52.
- Gales YA, Gross CL, Karebs RC, Smith WJ. (1989) Flow cytometric analysis of toxicity by alkylating agents in human epidermal keratinocytes. In: *In Vitro Toxicology: New Directions* (Goldberg AM, Ed.). Mary Ann Liebert, New York, pp. 169–174.
- Gandolfi AJ, Brendel K, Tisher R, Azri S, Hanan G, Waters SJ, Hanzlick RP, Thomas CM. (1989) *Utilization of precision-cut liver slices to profile and rank-order potential hepatotoxins*. PMA-Drusafe East Spring Meeting, May 2, 1989, San Francisco.
- Garside DA. (1988) Use of *in vitro* techniques to investigate the action of testicular toxicants. In: *In Vitro Methods in Toxicology* (Atterwill CK and Steele CE, Eds.). Cambridge University Press, New York, pp. 411–423.
- Gaul GM. (2008) In US, Few Alternatives to Testing in Animals. *Washington Post*, April 12, 2008:A01.
- Gebhart DOE. (1972) The use of the chick embryo in applied teratology. In: *Advances in Teratology* (Woollam DHM, Ed.), Vol. 5. Academic Press, London, pp. 97–111.
- Gombar VK, Borgstedt HH, Enslein K, Hart JB, Blake BW. (1991) A QSAR model of teratogenesis. *Quant. Struct.-Act. Relat.* 10:306–332.
- Gordon VC, Kelly CP, Bergman HC. (1989) *SKINTEX, an in vitro method for determining dermal irritation*. Fifth International Congress of Toxicology, July 16–21, 1989, Brighton, England.
- Gordon VC, Mirhaslehi S, Wei R. (1986) Evaluation of the corroded method to determine the corrosivity potential of surfactants, surfactant-based formulations, chemicals and mixtures. In: *Advances in Animal Alternatives for Safety and Efficacy*. Mary Ann Liebert, Philadelphia, PA, pp. 309–358.
- Gribaldo L. (2007) Overview of alternative methodologies in toxicology. *Curr. Protoc. Toxicol.* 34:20.1.1–20.1.6.
- Harman AW, Fisher LJ. (1983) Hamster hepatocytes in culture as a model for acetaminophen toxicity: studies with inhibitors of drug metabolism. *Toxicol. Appl. Pharmacol.* 71:330–341.
- Harpur ES. (1988) Ototoxicity. In: *In Vitro Methods in Toxicology* (Atterwill CK and Steele CE, Eds.). Cambridge University Press, New York, pp. 37–58.
- Harter ML, Felkner IC, Song PS. (1976) Near-UV effects of 5,7-dimethoxycoumarin in *Bacillus subtilis*. *Photochem. Photobiol.* 24:491–493.
- Hill AJ, Teraka H, Heideman W, Peterson RE. (2005) Zebrafish as a model vertebrate for investigating chemical toxicity. *Toxicol. Sci.* 86(1):6–19.
- Homburger F and Goldberg AM. (1985) *In vitro embryotoxicity and teratogenicity tests*. Joint American-Swiss Seminar, November 1984, Zürich.
- Hooisma J. (1982) Tissue culture and neurotoxicology. *Neurobehav. Toxicol. Teratol.* 4:617–622.
- ICCVAM. (1997) *Validation and Regulatory Acceptance of Toxicological Test Methods*. NIEHS, Research Triangle Park, NC.
- Indans I. (2002) The use and interpretation of in vitro data in regulatory toxicology: cosmetics, toiletries, and household products. *Toxicol. Lett.* 127:177–182.
- Interagency Regulatory Alternatives Group (IRAG). (1993) *Washington Group Reports*. Washington, DC.
- Jacaruso RB, Barlett MA, Carson S, Trombetta LD. (1985) Release of histamine from rat peritoneal cells *in vitro* as an index of irritational potential. *J. Toxicol. Cutan. Ocul. Toxicol.* 4:39–48.
- Jackson EM, Goldner R. (1989) *Irritant Contact Dermatitis*. Marcel Dekker, New York.
- Johnson EM, Gorman RM, Gabel BEC, George ME. (1982) The *Hydra attenuata* system for detection of teratogenic hazards. *Teratog. Carcinog. Mutagen.* 2:263–276.
- Johnson EM, Newman LM, Fu L. (1989) Letter to the editor. *Toxicol. Appl. Pharmacol.* 99:173–176.
- Johnson EM. (1989) Problems in validation of *in vitro* developmental toxicity assays. *Fundam. Appl. Toxicol.* 13:863–867.
- Jumblatt MM, Neufeld AH. (1985) A tissue culture model of the human corneal epithelium. In: *In Vitro Toxicology, Alternative Methods in Toxicology* (Goldberg AM, Ed.), Vol. 3. Mary Ann Liebert, New York, pp. 391–404.
- Kandorona H, Letasiova S. (2011) Alternative methods in toxicology: pre-validation and validated methods. *Interdiscip. Toxicol.* 4:107–113.
- Kao J, Hall J, Holland JM. (1982) Quantitation of cutaneous toxicity: an *in vitro* approach using skin organ culture. *Toxicol. Appl. Pharmacol.* 68:206–217.
- Keller SJ, Smith M. (1982) Animal virus screens for potential teratogens: poxvirus morphogenesis. *Teratog. Carcinog. Mutagen.* 2:361–374.
- Kemp RV, Meredith RWJ, Gamble S, Frost M. (1983) A rapid cell culture technique for assaying to toxicity of detergent based products *in vitro* as a possible screen for high irritants *in vivo*. *Cytobios.* 36:153–159.
- Kemp RV, Meredith RWJ, Gamble S. (1985) Toxicity of commercial products on cells in suspension: a possible screen for the Draize eye irritation test. *Food Chem. Toxicol.* 23:267–270.
- Kennah HE, Albulescu D, Hignet S, Barrow CS. (1989) A critical evaluation of predicting ocular irritancy potential from an *in vitro* cytotoxicity assay. *Fundam. Appl. Toxicol.* 12:281–290.

- Kerster HW, Schaeffer DJ. (1983) Brine shrimp (*Artemia salina*) nauplii as a teratogen test system. *Ecotoxicol. Environ. Saf.* 7:342–349.
- Kim J, Lee H, Selimović Š, Gauvin R, Bae H. (2015) Organ-on-a-chip: development and clinical prospects toward toxicity assessment with an emphasis on bone marrow. *Drug Saf.* 38:409–418.
- Kochhar DM, Aydelotte MB. (1974) Susceptible stages and abnormal morphogenesis in the developing mouse limb, analyzed in organ culture after transplacental exposure to vitamin A (retinoic acid). *J. Embryol. Exp. Morphol.* 31:721–734.
- Kochhar DM. (1981) Embryo explants and organ cultures in screening of chemicals for teratogenic effects. In: *Developmental Toxicology* (Kimmel CA and Buelke-Sam J, Eds.). Raven Press, New York, pp. 303–319.
- Kontur PJ, Hoffman PC, Heller A. (1987) Neurotoxic effects of methamphetamine assessed in three-dimensional reaggregate tissue cultures. *Dev. Brain Res.* 31:7–14.
- Kotzin BL, Baker RF. (1972) Selective inhibition of genetic transcription in sea urchin embryos. *J. Cell Biol.* 55:74–81.
- Kuehn BM. (2008) Health environmental agencies team up to develop new plan for toxicity testing. *JAMA.* 299:1415–1416.
- Kurack G, Vossen P, Deboysier D, Goethals F, Roberfud M. (1986) An *in vitro* model for acute toxicity screening using hepatocytes freshly isolated from adult mammals. In: *In Vitro Toxicology* (Goldberg AM, Ed.). Mary Ann Liebert, New York.
- Lamont GS, Bagley DM, Kong BM, DeSalva SJ. (1989) Developing an alternative to the Draize skin test: comparison of human skin cell responses to irritants *in vitro*. In: *In Vitro Toxicology: New Directions* (Goldberg AM, Ed.). Mary Ann Liebert, New York, pp. 175–181.
- Lander T, Enslein K, Craig P, Tomb N. (1984) Validation of a structure-activity model of rat oral LD50. In: *Acute Toxicity Testing: Alternative Approaches* (Goldberg AM, Ed.). Mary Ann Liebert, New York, pp. 183–184.
- Lei H, Carroll K, Au L, Krag SS. (1986) An *in vitro* screen for potential inflammatory agents using cultured fibroblasts. In: *In Vitro Toxicology, Alternative Methods in Toxicology* (Goldberg AM, Ed.), Vol. 3. Mary Ann Liebert, New York, pp. 74–85.
- Leighton J, Nassauer J, Tchao R, Verdon J. (1983) Development of a procedure using the chick egg as an alternative to the Draize test. In: *Product Safety Evaluation, Alternative Methods in Toxicology* (Goldberg AM, Ed.), Vol. 1. Mary Ann Liebert, New York, pp. 165–177.
- Lijinsky W. (1988) Importance of animal experiments in carcinogenesis research. *Environ. Mol. Mutagen.* 11:307–314.
- Luepke NP. (1985) Hen's egg chorioallantoic membrane test for irritation potential. *Food Chem. Toxicol.* 23:287–291.
- MacCrimmon HR, Kwain WH. (1969) Influences of light on early development and meristic characters in the rainbow trout (*Salmo gairdneri* Richardson). *Can. J. Zool.* 47:631–637.
- Maurice D, Singh T. (1986) A permeability test for acute corneal toxicity. *Toxicol. Lett.* 31:125–130.
- May J, Xu J, Morse R, Ovent N, Donaldson C. (2009) Toxicity testing: the search for an *in vitro* alternative to animal testing. *Br. J. Biomed. Sci.* 66:160–165.
- McAuliffe DJ, Hasan T, Parrish JA, Kochevar IE. (1986) Determination of photosensitivity by an *in vitro* assay as an alternative to animal testing. In: *In Vitro Toxicology* (Goldberg AM, Ed.). Mary Ann Liebert, New York, pp. 30–41.
- McKim JM. (2012) *In vitro* approaches for determining liver-specific toxicity of new drug candidates. In: *New Horizons in Predictive Toxicology: Current Status and Application* (Wilson AGE, Ed.). Royal Society of Chemistry, London, pp. 157–220.
- Mehendale HM. (1989) Application of isolated organ techniques in toxicology. In: *Principles and Methods of Toxicology* (Hayes AW, Ed.). Raven Press, New York, pp. 699–740.
- Meyer DR, McCulley JP. (1988) Acute and protracted injury to cornea epithelium as an indication of the biocompatibility of various pharmaceutical vehicles. In: *Progress in In Vitro Toxicology* (Goldberg AM, Ed.). Mary Ann Liebert, New York, pp. 215–235.
- Meyer DS. (1989) Safety evaluation of new drugs. In: *Modern Drug Research* (Martin YC, Kutter E and Austel V, Eds.). Marcel Dekker, New York, pp. 355–399.
- Miller AE Jr, Levis WR. (1973) Studies on the contact sensitization of man with simple chemicals. I. Specific lymphocyte transformation in response to dinitrochlorobenzene sensitization. *J. Invest. Dermatol.* 61:261–269.
- Milner JE. (1970) *In vitro* lymphocyte responses in contact hypersensitivity. *J. Invest. Dermatol.* 55:34–38.
- Milner JE. (1971) *In vitro* lymphocyte responses in contact hypersensitivity II. *J. Invest. Dermatol.* 56:349–352.
- Milner JE. (1983) *In vitro* tests for delayed skin hypersensitivity: lymphokine production in allergic contact dermatitis. In: *Dermatotoxicology* (Marzulli FN and Maibach HD, Eds.). Hemisphere Publishing, New York, pp. 185–192.
- Morison WL, McAuliffe DJ, Parrish JA, Bloch KJ. (1982) *In vitro* assay for phototoxic chemicals. *J. Invest. Dermatol.* 78:460–463.
- Muir CK, Flower C, Van Abbe NJ. (1983) A novel approach to the search for *in vitro* alternatives to *in vivo* eye irritancy testing. *Toxicol. Lett.* 18:1–5.
- Muir CK. (1984) A simple method to assess surfactant-induced bovine corneal opacity *in vitro*: preliminary findings. *Toxicol. Lett.* 23:199–203.
- Mummery CL, van den Brink CE, van der Saag PT, de Loat SW. (1984) A short-term screening test for teratogens using differentiating neuroblastoma cells *in vitro*. *Teratology.* 29:271–279.
- Natsch A, Ryan CA, Foertsch L, Emter R, Jaworska J, Gerberick F, Kern P. (2013) A dataset on 145 chemicals tested in alternative assays for skin sensitization undergoing prevalidation. *J. Appl. Toxicol.* 33:1337–1352.
- Naughton GK, Jacob L, Naughton BA. (1989) A physiological skin model for *in vitro* toxicity studies. In: *Progress in In Vitro Toxicology* (Goldberg AM, Ed.). Mary Ann Liebert, New York, pp. 183–189.
- Neubert D. (1982) The use of culture techniques in studies on prenatal toxicity. *Pharmacol. Ther.* 18:397–434.
- Neuhauser E, Durkin P, Malecki M, Antara M. (1985a) Comparative toxicity of ten organic chemicals to four earthworm species. *Comp. Biochem. Physiol.* 83C:197–200.

- Neuhauser E, Loehr C, Malecki M. (1986) Contact and artificial soil tests using earthworms to evaluate the impact of wastes in soil. In: *Hazardous and Industrial Solid Waste Testing: Fourth Symposium*, ASTM STP 886 (Petros J, Lacy W and Conway RC, Eds.). American Society for Testing and Materials, Philadelphia, PA, pp. 192–202.
- Neuhauser E, Loehr C, Malecki M, Milligan D, Durkin P. (1985b) The toxicity of selected organic chemicals to the earthworm *Eisenia fetida*. *J. Environ. Qual.* 14:383–388.
- North-Root H, Yackovich F, Demetruilas J, Gucula M, Heinze JE. (1982) Evaluation of an *in vitro* cell toxicity test using rabbit corneal cells to predict the eye irritation potential of surfactants. *Toxicol. Lett.* 14:207–212.
- Nukada Y, Miyazawa M, Kazutoshi S, Sakaguchi H, Nishiyama N. (2013) Data integration of non-animal tests for the development of a test battery to predict the skin sensitizing potential and potency of chemicals. *Toxicol. In Vitro.* 27:609–618.
- OECD (2005) *Environmental, Health and Safety News*, No. 17, April 2005.
- Oliver GJA, Pemberton NA. (1985) An *in vitro* epidermal slice technique for identifying chemicals with potential for severe cutaneous effects. *Food Chem. Toxicol.* 23:229–232.
- Parce JW, Owicki JC, Kercso DM, Sigal GB, Wada HG, Muir VC, Bousse LJ, Ross KL, Sikic BI, McConnell HM. (1989) Detection of cell-affecting agents with a silicon biosensor. *Science*. 246(4927):243–247.
- Placke ME, Fisher GL. (1987) Adult peripheral lung organ culture—a model for respiratory tract toxicology. *Toxicol. Appl. Pharmacol.* 90:284–298.
- PMA/Drusafe *In Vitro* Task Force. (1994) A collaborative evaluation of an *in vitro* muscle irritation assay. *Toxicol. Methods*. 4:215–223.
- Prottey C, Ferguson TFM. (1976) The effect of surfactants upon rat peritoneal mast cells *in vitro*. *Food Chem. Toxicol.* 14:425.
- Reinhardt CA, Aeschbacher M, Bracker M, Spengler J. (1987) Validation of three cell toxicity tests and the hen's egg test with guinea pig eye and human skin irritation data. In: *In Vitro Toxicology—Approaches to Validation. Alternative Methods in Toxicology* (Goldberg AM, Ed.), Vol. 5. Mary Ann Liebert, New York, pp. 463–470.
- Reinhardt CA, Pelli DA, Zbinden G. (1985) Interpretation of cell toxicity data for the estimation of potential irritation. *Food Chem. Toxicol.* 23:247–252.
- Reiser KM, Last JA. (1979) Silicosis and fibrogenesis: fact and artifact. *Toxicology*. 13:51–72.
- Roberts R, Dorough H. (1985) Hazards of chemicals to earthworms. *Environ. Toxicol. Chem.* 4:307–323.
- Rofe PC. (1971) Tissue culture and toxicology. *Food Cosmet. Toxicol.* 9:685–696.
- Roth JA. (1980) Use of perfused lung in biochemical toxicology. *Rev. Biochem. Toxicol.* 1:287–309.
- Ruben Z, Fuller GC, Knodle SG. (1984) Disubutamide-induced cytoplasmic vacuoles in cultured dog coronary artery muscle cells. *Arch. Toxicol.* 55:206–212.
- Russell WMS, Burch RL. (1959) *The Principles of Humane Experimental Technique*. Methuen, London.
- Rylander LA, Phelps JS, Gandolfi AJ, Brendel K. (1987) *In vitro* nephrotoxicity: response of isolated renal tubules to cadmium chloride and dichlorovinyl cysteine. *In Vitro Toxicol.* 1:111–127.
- Salem H. (1995) *Animal Test Alternatives*. Marcel Dekker, New York.
- Salem H, Katz SA. (1998) *Advances in Animal Alternatives for Safety and Efficacy Testing*. Taylor & Francis, Philadelphia, PA.
- Saxen L, Saksela E. (1971) Transmission and spread of embryonic induction. Exclusion of an assimilatory transmission mechanism in kidney tubule induction. *Exp. Cell Res.* 66:369–377.
- Scaife MC. (1982) An investigation of detergent action on *in vitro* and possible correlations with *in vivo* data. *Int. J. Cosmet. Sci.* 4:179–193.
- Scheuplein RJ, Schoal SE, Brown RN. (1990) Role of pharmacokinetics in safety evaluation and regulatory considerations. *Annu. Rev. Pharmacol. Toxicol.* 30:1970218.
- Selling J, Ekwall B. (1985) Screening for eye irritancy using cultured Hela cells. *Xenobiotica*. 15:713–717.
- Shadduck JA, Everitt J, Bay P. (1985) Use of *in vitro* cytotoxicity to rank ocular irritation of six surfactants. In: *In Vitro Toxicology: Alternative Methods in Toxicology* (Goldberg AM, Ed.), Vol. 3. Mary Ann Liebert, New York, pp. 641–649.
- Shopsis C, Eng B. (1985) Uridine uptake and cell growth cytotoxicity tests: comparison, applications and mechanistic studies. *J. Cell Biol.* 101:87a.
- Shopsis C, Sathe S. (1984) Uridine uptake inhibition as a cytotoxicity test: correlation with the Draize test. *Toxicology*. 29:195–206.
- Sieber-Blum MF. (1985) *Differentiation of Avian Neural Crest Cells In Vitro (Quail, Chick, Rodent)*. Crisp Data Base HD15311-04. National Institutes of Health (NIH), Bethesda, MD.
- Silverman J. (1983) Preliminary findings on the use of protozoa (*Tetrahymena thermophila*) as models of ocular irritation testing in rabbits. *Lab. Anim. Sci.* 33:56–59.
- Simons PJ. (1981) An alternative to the Draize test. In: *The Use of Alternatives in Drug Research* (Rowan AN and Stratmann CJ, Eds.). Macmillan Press, London.
- Singer P. (1975) *Animal Liberation: A New Ethic for Our Treatment of Animals*. Random House, New York.
- Sleet RB, Brendel K. (1985) Homogenous populations of *Artemia nauplii* and their potential use for *in vitro* testing in developmental toxicology. *Teratog. Carinog. Mutagen.* 5(1):41–54.
- Smith AM, Acosta D, Bruckner JV. (1986) Development of a primary culture system of rat kidney cortical cells to evaluate the nephrotoxicity of xenobiotics. *Food Chem. Toxicol.* 24:551–556.
- Smith MA, Acosta D, Bruckner JV. (1987) Cephaloridine toxicity in primary cultures of rat renal epithelial cells. *In Vitro Toxicol.* 1:23–29.
- Smith MA, Hewitt WR, Hook J. (1988) *In Vitro* methods in renal toxicology. In: *In Vitro Methods in Toxicology* (Aterwill CK and Steele CE, Eds.). Cambridge University Press, New York, pp. 13–36.
- Smith MK, Kimmel GL, Korchhar DM, Shepard TH, Spielberg SP, Wilson JC. (1983) A selection of candidate compounds for *in vitro* teratogenesis test validation. *Teratog. Carcinog. Mutagen.* 3:461–480.
- Soto RJ, Servi MJ, Gordon VC. (1988) Evaluation of an alternative method of ocular irritation. In: *Progress in In Vitro Toxicology* (Goldberg AM, Ed.). Mary Ann Liebert, New York, pp. 289–296.
- Spencer PS, Crain SM, Bornstein MB, Peterson ER, van de Water T. (1986) Chemical neurotoxicity: detection and analysis in organotypic cultures of sensory and motor systems. *Food Chem. Toxicol.* 24:539–544.

- Stammati AP, Silano V, Zucco F. (1981) Toxicology investigations with cell culture systems. *Toxicology*. 20:91–153.
- Stenersen J. (1979) Action of pesticides on earthworms. Part I: toxicity of cholinesterase-inhibiting insecticides to earthworms as evaluated by laboratory tests. *Pestic. Sci.* 10:66–74.
- Stern HM, Zon LI. (2003) Cancer genetics and drug discovery in the zebrafish. *Nat. Rev. Cancer*. 3(7):533–539.
- Swisher DA, Prevo ME, Ledger PW. (1988) The MTT *in vitro* cytotoxicity test: correlation with cutaneous irritancy in two animal models. In: *Progress in In Vitro Toxicology* (Goldberg AM, Ed.). Mary Ann Liebert, New York, pp. 265–269.
- Tanguay RL, Reimers MJ. (2008) Analysis of ethanol developmental toxicity in zebrafish. *Methods Mol. Biol.* 447:63–74.
- Tchao R. (1988) Trans-epithelial permeability of fluorescein *in vitro* as an assay to determine eye irritants. In: *Progress in In Vitro Toxicology* (Goldberg AM, Ed.). Mary Ann Liebert, New York, pp. 271–284.
- Thulin H, Zacharian H. (1972) The leukocyte migration test in chromium hypersensitivity. *J. Invest. Dermatol.* 58:55–58.
- Traina VM. (1983) The role of toxicology in drug research and development. *Med. Res. Rev.* 3:43–72.
- Troese M, Pratt LF, Cerven DR, DeGeorge G. (2012) IVSA: *in vitro* sensitization assay. ACT Annual Meeting, November 4–7, 2012, Champions Gate, FL.
- Tyson CA, Stacey NH. (1989) *In vitro* screens from CNS, liver and kidney for systemic toxicity. In: *Benchmarks: Alternative Methods in Toxicology* (Mehlman M, Ed.). Princeton Scientific, Princeton, NJ, pp. 111–136.
- Volpe LS, Biagioni TM, Marquis JK. (1985) *In vitro* modulation of bovine caudata muscarinic receptor number by organophosphates and carbamates. *Toxicol. Appl. Pharmacol.* 78:226–234.
- Watanabe M, Watanabe K, Suzuki K, Nikaido O, Sugahara T, Ishii I, Konishi H. (1988) *In vitro* cytotoxicity test using primary cells derived from rabbit eye is useful as an alternative for Draize testing. In: *Progress in In Vitro Toxicology* (Goldberg AM, Ed.). Mary Ann Liebert, New York, pp. 285–290.
- Williams GM, Dunkel VC, Ray VA. (Eds.) (1983) Cellular systems for toxicity testing. *Ann. N. Y. Acad. Sci.* 407:.
- Williams PD, Masters BG, Evans LD, Laska DA, Hattendorf GH. (1987) An *in vitro* model for assessing muscle irritation due to parenteral antibiotics. *Fundam. Appl. Toxicol.* 9:10–17.
- Wilson JG. (1978) Review of *in vitro* systems with potential for use in teratogenicity screening. *J. Environ. Pathol. Toxicol.* 2:149–167.
- Wolfgang GHI, Verneti LA, MacDonald JR. (1994) Isolation and use of primary adrenocortical cells from guinea pigs, dogs and monkeys for *in vitro* toxicity studies. *Toxicol. Methods*. 4:149–160.
- Wyman J, Stokes JS, Goehring M, Buring M, Moore T. (1995) Data collection interface for isolated perfused rat liver: recording oxygen consumption, perfusion pressure and pH. *Toxicol. Methods*. 5:1–14.
- Young MF, Trombetta LD, Sophia JV. (1986) Correlative *in vitro* and *in vivo* study of skeletal muscle irritancy. *Toxicologist*. 6:1225.
- Zbinden G. (1987) *Predictive Value of Animal Studies in Toxicology*. Centre for Medicines Research, Carshalton.

FURTHER READING

- American Medical Association (1989) Public Support for Animals in Research. *American Medical News*, June 9, 1989, pp. 12–16.
- Baldessari D, Mione M. (2008) How to create the vascular tree? (Latest) help from the zebrafish. *Pharmacol. Ther.* 118(2):206–230.
- Basketter DA. (2008) Nonanimal alternatives for skin sensitization: a step forward? *Toxicol. Sci.* 102:1–2.
- Blagg J. (2006) Structure-activity relationships for *in vitro* and *in vivo* toxicity. *Annu. Rep. Med. Chem.* 41:353–368.
- Brouwer KR, Hsiao P, Brouwer KLR. (2015) Intracellular Drug Concentrations: A Critical Consideration for *In Vitro* Assays. *AAPS Newsmagazine*, April 4, 2015:16–19.
- Brown CG. (1988) Application of thyroid cell culture to the study of thyrotoxicity. In: *In Vitro Methods in Toxicology* (Atterwill CK and Steele CE, Eds.). Cambridge University Press, New York, pp. 165–188.
- Chasin M, Scott C, Shaw C, Persico F. (1979) A new assay for the measurement of mediator release from rat peritoneal mast cells. *Int. Arch. Allergy Appl. Immunol.* 58:1–10.
- Cowley G, Hager M, Drew L, Namuth T, Wright L, Murr A, Abbot N, Robins, K. (1988) The Battle Over Animal Rights. *Newsweek*, December 26, 1988, 113(26).
- Doehmer J. (2006) Predicting drug metabolism dependent toxicity for humans with a genetically engineered cell battery. *Altern. Lab. Anim.* 34:561–573.
- Doucet O, Lanvin M, Thillou C, Pupat C, Merlin B, Zartrow L. (2006) Reconstituted human corneal epithelium: a new alternative to the Draize eye test for the assessment of the eye irritation potential of chemicals and cosmetics. *Toxicol. In Vitro*. 20:499–512.
- Enslein K, Blake VW, Tuzzeo TM, Borgstedt HH, Hart JB, Salem H. (1988) Estimation of rabbit eye irritation scores by structure-activity equations. *In Vitro Toxicol.* 2:1–14.
- Faulkner L, Martinsson K, Santoyo-Castelazo A, Cederbrant K, Schuppe-Koistinen I, Powell H, Tugwood J, Naisbitt DJ, Park BK. (2012) The development of *in vitro* culture methods to characterize primary T-cell responses to drugs. *Toxicol. Sci.* 127(1):150–158.
- Flamand N, Marrot L, Belaidi JP, Bowrouf L, Dourille E, Feister M, Meunier JR. (2006) Development of genotoxicity test procedures with chiohim, a reconstructed human skin model: towards new tools for *in vitro* risk assessment of dermally applied compounds. *Mutat. Res.* 606:39–51.
- Flint OP. (1989b) Reply to letter to the editor. *Toxicol. Appl. Pharmacol.* 99:176–180.
- Franchini JL, Propat JT, Coner GR, Yost MJ. (2007) Novel tissue engineered tubular heart tissue for *in vitro* pharmaceutical toxicity testing. *Microsc. Microanal.* 13:267–271.
- Gad SC. (1988a) An approach to the design and analysis of screening data in toxicology. *J. Am. Coll. Toxicol.* 7(2):127–138.
- Gad SC. (1988b) Defining product safety information and testing requirements. In: *Handbook of Product Safety Evaluation* (Gad SC, Ed.). Marcel Dekker, New York, pp. 1–22.
- Gad SC. (1988c) Principles of screening in toxicology: with special emphasis on applications to neurotoxicology. *J. Am. Coll. Toxicol.* 8:21–27.

- Gad SC. (1992) Industrial application for *in vitro* toxicity testing methods: a tier testing strategy for product safety assessment. In: *In Vitro Toxicity Testing* (Frazier J, Ed.). Marcel Dekker, New York, pp. 253–279.
- Gad SC. (1993) In vivo and in vitro tox. In: *Biotechnology and Safety Assessment* (Thomas JA and Myers LA, Eds.). Raven Press, New York, pp. 97–129.
- Gad SC. (1995a) Challenges and needs in the safety assessment of new therapeutics. *Toxicol. Methods*. 5:viii–x.
- Gad SC. (1995b) Integration of alternative methods into the safety evaluation process in the United States. *Comments Toxicol*. 5:301–314.
- Gad SC. (2000) *In Vitro Toxicology*, 2nd Ed. Taylor & Francis, Philadelphia, PA.
- Gad SC. (2001) *Regulatory Toxicology*, 2nd Ed. Taylor & Francis, Philadelphia, PA.
- Gad SC, Dunn BJ, Dobbs DW, Walsh RD. (1986) Development and validation of an alternative dermal sensitization test: the mouse ear swelling test (MEST). *Toxicol. Appl. Pharmacol*. 84:93–114.
- Garthoff B. (2005) Alternatives to animal experimentation, the regulatory background. *Toxicol. Appl. Pharmacol*. 207:388–392.
- Gupta K, Rispin A, Stitzel K, Cocake S, Harbell J. (2005) Ensuring quality of in vitro alternative test methods: issues and answers. *Regul. Toxicol. Pharmacol*. 43:219–224.
- ICH (2014) *M7(R2) Assessment and Control of DNA Reactive (Mutagenic) Impurities in Pharmaceuticals to Limit Potential Carcinogenic Risk*. June 23, 2014. International Conference on Harmonisation, Geneva, Switzerland.
- Jeong JY, Kwon HB, Ahn JC, Kang D, Kwon SH, Park JA, Kim KW. (2008) Functional and developmental analysis of the blood–brain barrier in zebrafish. *Brain Res. Bull*. 75(5):619–628.
- Jester JV. (2006) Extent of corneal injury as a biomarker for hazard assessment and the development of alternative models to the Draize rabbit eye test. *Cutan. Ocul. Toxicol*. 25:41–54.
- Klausner M, Neal PJ, Cannon CL, Kabilus J, Curren RD, Harbell JW. (1998) Intra- and inter-laboratory reproducibility of epidermis, and in vitro model for dermal irritancy testing. In: *Advances in Animal Alternatives for Safety and Efficacy Testing* (Salem H and Katz SA, Eds.). Taylor & Francis, Philadelphia, PA, pp. 347–358.
- Knauer K, Lampert C, Gonzalez-Valero J. (2007) Comparison of in vitro and in vivo acute fish toxicity in relation to toxicant mode of action. *Chemosphere*. 68:1435–1441.
- Kneuer C, Labona C, Honsiha W. (2007) Prediction of acute toxicity in HPCT-IE3 hepatocytoma cells with liver like transport activities. *Altern. Lab. Anim*. 35:411–420.
- Loueskarni K, Sihvonen K, Kuittensen M, Sommes V. (2006) In vitro tests within the REACH information strategies. *Altern. Lab. Anim*. 34:377–386.
- Maldenhauser F. (2003) Using in vitro prediction models instead of the rabbit eye irritation test to classify and label new chemicals: a post for data analysis of the international EC/HO validation study. *Altern. Lab. Anim*. 31:31–46.
- Mueller SO, Tusch F, Kling M. (2007) Alternatives in pharmaceutical toxicology: global and focused approaches—two case studies. *ALTEX*. 24:117–124.
- Persini A, Bonomi A. (2007) CFU-GM assay for evaluation of drug myelotoxic activity. *Curr. Protoc. Toxicol*. 20.2.1–20.2.18.
- Rogiers V, Danniels M. (2006) Good science must be the key factor in the development and use of alternative methods for safety assessment of cosmetics. *ALTEX*. 23:346–352.
- Shah PP, Myers MC, Beavers MP, Purvis JE, Jing H, Grieser HJ, Sharlow ER, Napper AD, Huryn DM, Cooperman BS, Smith AB, Diamond SL. (2008) Kinetic characterization and molecular docking of a novel, potent, and selective slow-binding inhibitor of human Cathepsin L. *Mol. Pharmacol*. 74(1):34–41.
- Stigson M, Kultima K, Jergil M, Scholz B, Alm H, Gastophon AL, Dencher L. (2007) Molecular targets and early response biomarkers for the prediction of developmental toxicity in vitro. *Altern. Lab. Anim*. 35:561–573.
- Tornier C, Rosdy M, Maibach HI. (2006) In vitro skin irritation testing on reconstituted human epidermis: reproducibility for 50 chemically tested with two protocols. *Toxicol. In Vitro*. 20:401–416.
- Tweats DJ, Scott AD, Westmoreland C, Carmichael PL. (2007) Determination of genetic toxicity and potential carcinogenicity in vitro challenges for the seventh amendment to the European cosmetics guideline. *Mutagenesis*. 22:5–13.
- Ubels TL, Clousing DP. (2005) In vitro alternatives to the use of animals in ocular toxicity testing. *Occul. Surf*. 3:126–142.
- Uyeki EM, Ashkar AE, Shoeman DW, Bisel JU. (1977) Acute toxicity of benzene inhalation of hemopoietic precursor cells. *Toxicol. Appl. Pharmacol*. 40:49–57.
- Verma K, Ugraval N, Meisa RB, Faroog M, Haas RK. (2008) Phototoxicity assessment of drugs and cosmetic products using *E. coli*. *In Vitro Toxicol*. 22:249–253.
- Verwei M, Van Burgsteden JA, Krul CA, Van De Sandt JJ, Freidig AP. (2007) Prediction of *in vivo* embryo toxic effect levels with a combination of in vitro studies and PBPK modeling. *Toxicol. Lett*. 165:79–87.
- Wilhelm KP, Maibach HI. (2008) OECD guidelines for testing chemicals. In: *Dermatotoxicology* (Marzulli FN and Maibach HI, Eds.), 7th Ed. CRC Press, Boca Raton, FL, pp. 303–305.
- Williams A, Sarkar S, Cuddon P, Tfofi EK, Saiki S, Siddiqi FH, Jahreiss L, Fleming A, Pask D, Goldsmith P, O’Kane CJ, Floto RA, Rubinsztein DC. (2008) Novel targets for Huntington’s disease in an mTOR-independent autophagy pathway. *Nat. Chem. Biol*. 4(5):295–305.
- Zhang C, Cao J, Kenyon JR, Panzica-Kelly JM, Gong L, Augustine-Rauch K. (2012) Development of a streamlined rat whole embryo culture assay for classifying teratogenic potential of pharmaceutical compounds. *Toxicol. Sci*. 127(2):535–546.

EVALUATION OF HUMAN TOLERANCE AND SAFETY IN CLINICAL TRIALS: PHASE I AND BEYOND

28.1 THE PHARMACEUTICAL CLINICAL DEVELOPMENT PROCESS AND SAFETY

As introduced at the beginning of this volume, the process for developing a new therapeutic is a long (6–16 years from drug inception to market approval) and costly (in reality, from \$100 million to \$1.3 billion, depending on how one allocates costs) process, when successful. And most therapeutics fail, with an actual decline in success rates over the last decade. It is instructive to consider that most failures are due to a lack of demonstration of efficacy in trials and all the ways that can happen. Downing et al. (2014) have provided an excellent review of what types of clinical evidence has supported recent (2005–2012) Food and Drug Administration (FDA) approvals. It is shaped by medical needs, regulatory requirements, economics, our understanding of sciences and diseases, limitations of available technology, and sometimes politics. All of these interact to shape a process which serves to iteratively reduce risks (both economic and human safety), with the probability of failure being reduced in a stepwise fashion (Matoren, 1984; PhRMA, 2016; Gad, 2009). Figure 28.1 briefly summarizes this process.

Much has changed in the clinical trials world since the last edition of this book:

- As the pharmaceutical industry and market have become more global and the need for suitable clinical trial subjects has increased, many more trials are multinational. While providing the opportunity to increase the number and diversity of patients, this has also led to significant difficulties in controlling variability in study conduct and results between sites. Table 28.1 summarizes factors contributing to this variability.
- Nations have acted to increase both access to clinical trial participation (in the United States, all treatment trials must be regulated with clinicaltrials.gov and described on this site—see Table 28.2 for the yearly, resulting number of registered trials through 2013) and outcomes (see Mello et al., 2013 for an excellent review of progress and the issues).
- Regulations now require much more formal and rigorous approaches to collecting evaluating and reporting post market clinical safety data. Such pharmacovigilance requirements and approaches to meeting them are described in a subsequent chapter.
- The successful integration of genomic data to identify at-risk populations in clinical trials and develop specific comparison biomarkers to sensitively identify desired therapeutic effects of biological therapeutic agents has vastly improved some specificity and safety (see Wilke et al., 2007 for approaches to identifying genetic risk factors for serious ADRs) for new drugs, and in the utilization of pharmacogenomics, challenges remain (Table 28.3).
- As of the time of this writing, revision of “The Common Rule” regulating all biomedical and behavioral research involving human subjects in the United States (covering not just FDA and other parts of the Department of Health and Human Services (HHS), but also 15 other US federal departments and agencies) is under revision (Hudson and Collins, 2015). Table 28.4 summarizes aspects of the current Common Rule and proposed changes.

From our perspective (i.e., a safety assessment perspective), the purpose of all nonclinical (animal and *in vitro*) testing is

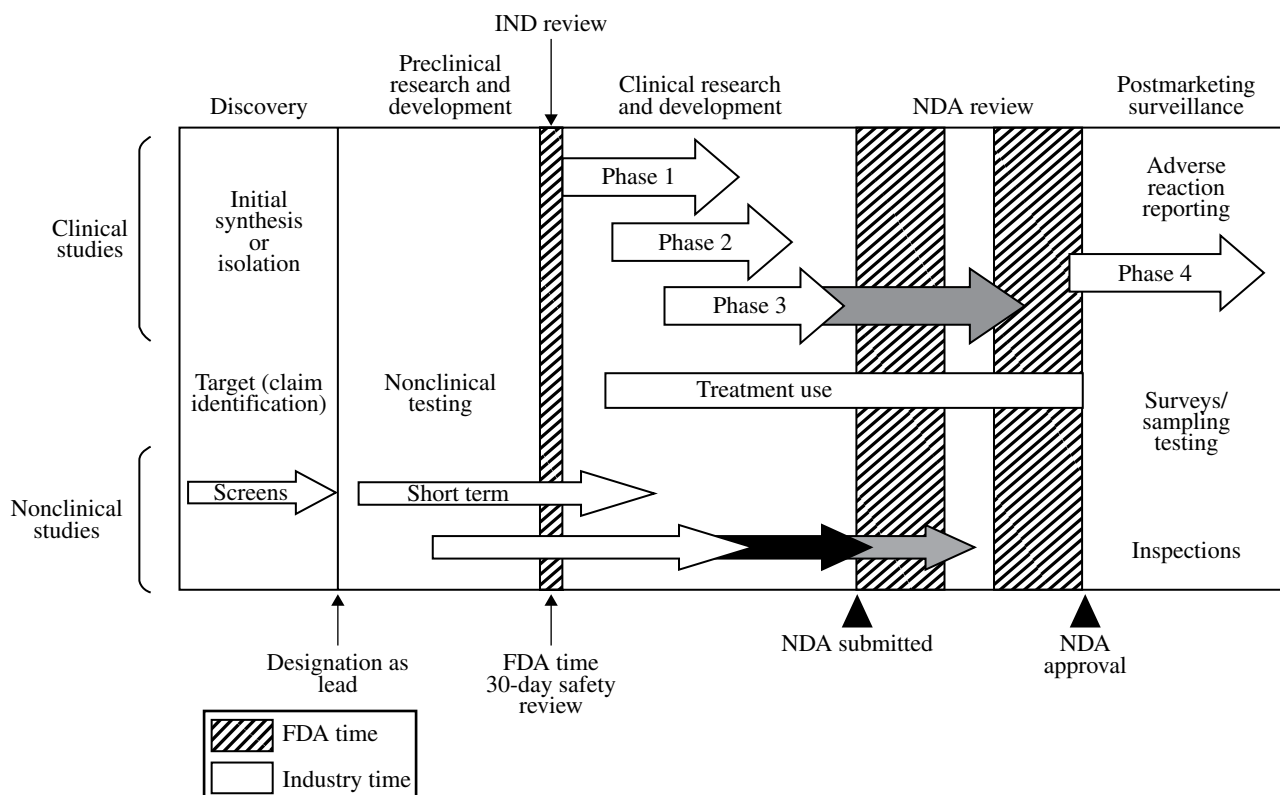


FIGURE 28.1 Pharmaceutical development process viewed as four stages (discovery, preclinical development, clinical development, and NDA review) as well as important postmarket surveillance phase.

TABLE 28.1 Sources of Variability in Drug Response That May Cause Toxicity or Lack of Efficacy

| Source of Variability | Mechanism | Drug Therapy | Description of Problem |
|---------------------------------------|---|--------------------------|--|
| <i>Biology</i> | | | |
| Genomics | Individual patients' genomic makeup influences PK and PD, affecting drug concentration profile at the target site and the likelihood and magnitude of desired and adverse effects | Trastuzumab | PD: only effective in patients overexpressing HER2 receptor on tumor cells |
| | | Abacavir | PD: high risk of severe hypersensitivity reaction in patients with <i>HLA-B*5701</i> allele |
| | | Codeine | PK: lack of analgesia in carriers of nonfunctional <i>CYP2D6</i> alleles, toxicity with multiple <i>CYP2D6</i> gene copies |
| Other intrinsic and extrinsic factors | Comorbidity, baseline severity of disease, other altered physiological states, or external factors influencing PK and/or PD | Insulin | PD: glucose control and risk of hypoglycemia affected by stress or physical activity |
| | | Several drugs | PK: increased toxicity due to increased absorption with concomitant consumption of grapefruit juice |
| <i>Behavior</i> | | | |
| Prescribing and drug handling | Inappropriate or off-label prescribing, coprescribing with an interacting drug, continued prescribing to nonresponders or medication errors | Cerivastatin | Rhabdomyolysis due to too high starting doses |
| | | Gemfibrozil | Interactions in drug label ignored |
| | | Mibefradil | Toxicity due to drug–drug interactions, label often ignored |
| Patient adherence | Poor adherence to prescribed drug regime, discontinuation (nonpersistence), “drug holidays,” or inadvertent overdosing | Antihypersensitive drugs | Nonadherence or nonpersistence perceived as treatment-resistant hypertension |
| | | Anti-infective therapies | Drug holidays leading to the development of resistance |

Note: The overall variability and the extent of the efficacy–effectiveness gap are driven by the interplay between biological and behavioral factors. *CYP2D6*, cytochrome P450 2D6; *HLA*, human leukocyte antigen; PD, pharmacodynamics; PK, pharmacokinetics.

TABLE 28.2 Number of New Phase 2, 2/3, and 3 Clinical Trials within ClinicalTrials.gov by Year, *n* (%)

| Receipt Year | Total | Phase 2 | Phase 2/3 | Phase 3 |
|--------------|-------|-------------|-----------|-------------|
| 1999 | 202 | 110 (54.5) | 3 (1.5) | 89 (44.1) |
| 2000 | 77 | 43 (55.8) | 1 (1.3) | 33 (42.9) |
| 2001 | 76 | 41 (53.9) | 1 (1.3) | 34 (44.7) |
| 2002 | 439 | 210 (47.8) | 18 (4.1) | 211 (48.1) |
| 2003 | 348 | 197 (56.6) | 12 (3.4) | 139 (39.9) |
| 2004 | 551 | 213 (38.7) | 10 (1.8) | 328 (59.5) |
| 2005 | 2873 | 1014 (35.3) | 92 (3.2) | 1767 (61.5) |
| 2006 | 2376 | 920 (38.7) | 88 (3.7) | 1368 (57.6) |
| 2007 | 2448 | 1165 (47.6) | 88 (3.6) | 1195 (48.8) |
| 2008 | 3027 | 1437 (47.5) | 85 (2.8) | 1505 (49.7) |
| 2009 | 2396 | 1221 (51.0) | 84 (3.5) | 1091 (45.5) |
| 2010 | 2182 | 1028 (47.1) | 63 (2.9) | 1091 (50.0) |
| 2011 | 2179 | 1011 (46.4) | 80 (3.7) | 1088 (49.9) |
| 2012 | 2071 | 961 (46.4) | 70 (3.4) | 1040 (50.2) |
| 2013 | 1823 | 816 (44.8) | 57 (3.1) | 950 (52.1) |

TABLE 28.3 Challenges in Pharmacogenomic Profiling in Drug Discovery*Clinical Trial Design*

- There remains no clear universal consensus within the scientific community as to the best design of pharmacogenomic trials
- More research is needed on the consequences of variable degrees of penetrance for drug response and adverse events

Stratification of Research Subjects

- Possible loss of benefits from trial participation
- Possible unfair representation of certain groups or populations in trials
- Possible confounding of statistical analysis and clinically relevant interpretation of results
- Possible sampling biases

Pharmacoeconomics

- Narrowing of drug markets
- Possible “orphan drug” syndrome
- Possible “orphan population” syndrome
- Questionable whether it is a good investment for venture capitalists

New Social Risks

- New “hidden” disease categories and new labels
- Possible “shifting of blame”
- Possible overemphasis on “pharmacogenomic cures”

TABLE 28.4 Major Proposed Changes in the Common Rule Notice of Proposed Rulemaking (NPRM)^a

| Objective | Current Common Rule | NPRM |
|---|---|--|
| Enhancing respect, strengthening informed consent | Does not apply to research with de-identified specimens Generally does not allow broad consent Provides basic elements of informed consent | Requires consent for all biospecimens, with a few exceptions Allows broad consent for future research use of biospecimens and data Adds new elements of informed consent; streamlines consent forms to required elements |
| Enhancing safeguards | Requires IRBs to ensure confidentiality but provides no specific measures | Specifies privacy and security safeguards for research involving biospecimens and identifiable information |
| Streamlining review | Allows use of single or local IRBs for multisite studies Allows expedited review of minimal-risk research Defines minimal risk | Requires use of single IRB for multisite studies in the United States, unless otherwise required by law Requires justification for continuing review of studies eligible for expedited review Guidance will be published on determining minimal-risk research |
| Calibrating oversight | Requires IRB review of grant or contract proposals Does not apply to research not funded by Common Rule agencies Exempts certain types of low-risk research but does not list examples of exemptions for nonresearch activities | Does not require IRB review of grant or contract proposals Extends rule to unregulated, privately funded clinical trials conducted at institutions funded for human subjects research Sets out excluded activities; clarifies entirely exempt activities and creates new category of partially exempt categories |

Source: Adapted from Hudson and Collins (2015).

^aThis table provides an overview of the major changes; it is not comprehensive. IRB denotes institutional review board. The following agencies have signed on the NPRM for the Common Rule: Department of Homeland Security, Department of Agriculture, Department of Energy, National Aeronautics and Space Administration, Department of Commerce, Social Security Administration, Agency for International Development, Department of Justice, Department of Labor, Department of Defense, Department of Education, Department of Veterans Affairs, Environmental Protection Agency, Department of Health and Human Services, National Science Foundation, Department of Transportation, and Department of Housing and Urban Development.

to reduce the risks and probability of adverse events (AEs) in human populations. But between initial nonclinical testing (and concurrent with additional animal testing) and a drug reaching the marketplace, the potential for having adverse effects in the general patient population it is intended for is further guarded against by a scheme of increasingly more powerful human (or “clinical”) trials (Piantadosi, 1997). The final purpose of the clinical trial process (see Table 28.5) is to show a therapeutic effective and sufficiently safe (compared to the benefit to the intended patient population) and to optimize these two factors. How safety is evaluated in trials is the subject of this chapter. As has been discussed previously, “safety” is not a fixed thing but rather a relative term in balance with the needs of the patients and the benefit derived. For rapidly life-threatening diseases (cancers and ALS are more common cases, but fatal orphan and rare diseases also enter here), short (out to 2–5-year) maintenance of life and of a reasonable quality of life is an acceptable level of safety if life span and degradation of quality of life by the drug can be maintained. This is demonstrated in the conduct of cancer trials and made sharply clear comparing the conduct of standard nonorphan drugs versus drugs for orphan (rare disease) cancer drugs (Kesselheim et al., 2011).

Cancer trials themselves have evolved markedly in recent years, with adaptive trial methods (Harning, 2007; Meille et al., 2008) largely displacing the older rigid (and slower)

designs. Still, most cancer trials (especially if performed only in a single country) take years because of both the duration of a course of treatment and slow rates of patient recruitment.

Also seeking to allow more rapid and efficient (less costly) first clinical evolution (though more pharmacokinetics and “proof of concept” than safety) are exploratory Investigational New Drug (IND)/CTA and microdosing approaches (Robinson, 2008) (see Table 28.6).

At the same time, as discussed in the earlier chapter on biotechnology products (Chapter 19), understanding of potential safety concerns with immunomodulatory monoclonal antibodies has led to changes in thinking as to dose selection in initial clinical trials of such agents and in the duration and end points of interest to provide improved characterization of the potential clinical safety of such agents (Muller and Brennan, 2009), basing consideration on target (and off-target) receptor occupancy (Ehlert, 2015) and not just plasma levels (Figure 28.2).

The most common “unexpected” (from nonclinical trial results) safety findings in first-in-human trials continue to be those involving the skin (dermatitis of one form or another) and liver (Kaplowitz, 2001). But safety reasons for drug withdrawal have a different order of incidence:

- Hepatotoxicity (or delayed liver toxicity (DLT))
- Nephrotoxicity
- Immunotoxicity (primarily for monoclonal antibodies)
- Cardiotoxicity

TABLE 28.5 Key Features of Clinical Trial Design

Phase I

- Designed to provide information about pharmacokinetics and pharmacodynamics
- The first time that human participants are involved; focused on safety
- Not always randomized
- Useful in identifying minimal and maximal dosages

Phase II

- Focus on drug efficacy, safety, and determining appropriate range of drug doses in patients with a disease or condition of interest
- Can be randomized

Phase III

- Large population sample
- Usually a comparison of new therapeutic intervention with standard treatment or placebo
- Generally a randomized and blinded study
- Often the final stage of testing before new drug approval can be granted

Phase IV

- Large-scale, long-term postmarketing studies
- Focused on identifying morbidity, mortality, and adverse events
- Might identify new indications

Genomics and biomarkers seek to optimize the process and at least for oncology have demonstrated significant benefits, especially as we come to identify genetic associations with leading safety issues (e.g., liver toxicity; see Table 28.7).

Except for those cases where there is substantial potential to save or extend lives (such as anticancer ALS and anti-AIDS drugs) or where the intended target diseases are chronic and severe (Parkinson’s or MS) or the routes of administration are invasive (intrathecal), the initial evaluations in humans are performed in “normal” healthy volunteers with the primary objective being limited to defining the limits of tolerance (safety) of the potential drug and its pharmacokinetic (PK) characteristics. These trials may also seek to detect limited (usually surrogate—i.e., indirect) indicators of efficacy but are severely limited in doing so (Biomarkers Definitions Working Group, 2001). Later trials look at the drug’s actions on carefully defined groups of patients.

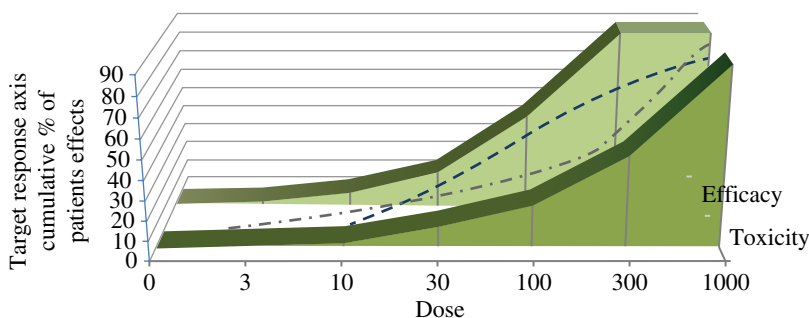
Some should be said about the special classes of studies (for cancer, HIV, ALS and other life-threatening diseases) for which first-in-humans are not performed in normal healthy volunteer, but rather in patient. Most frequently, these are patients that have failed other available forms of therapy. Even the design of such trials is different, through

TABLE 28.6 Exploratory Clinical Study Approaches

| Clinical Dosing Regimen | Recommendations for Nonclinical Safety Studies |
|--|---|
| Microdose approach 1 Single or multiple doses with total dose $\leq 100\mu\text{g}$ And total dose $\leq 1/100\text{th}$ of both NOAEL and PAD | Single-dose toxicity study (intended route or i.v.) One species, usually rodent Maximum dose of 1000-fold the clinical dose on a mg kg^{-1} or mg m^{-2} basis for i.v. and oral administration, respectively No safety pharmacology or genotoxicity studies |
| Microdose approach 2 Multiple doses (maximum of 5) with total dose $\leq 500\mu\text{g}$ And each dose $\leq 100\mu\text{g}$ And each dose $\leq 1/100\text{th}$ of both NOAEL and PAD | Seven-day toxicity study (intended route or i.v.) One species, usually rodent Maximum dose of 1000-fold the clinical dose on a mg kg^{-1} or mg m^{-2} basis for i.v. and oral administration, respectively No safety pharmacology or genotoxicity studies |
| Single-dose studies Dose up to therapeutic range Maximum dose yields up to $1/2$ exposure at NOAEL in most sensitive species | Single-dose toxicity studies Both rodent and nonrodent Maximum dose of MTD, MFD, or limited dose Safety pharmacology and Ames assay |
| Dosing for 3–14 days Dose up to therapeutic range Not to exceed nonrodent duration Maximum dose yields up to exposure at nonrodent NOAEL or $1/2$ exposure at rodent NOAEL | Two-week toxicity study in rodent Maximum dose of MTD, MFD, or limited dose Confirmatory study in nonrodent ($n=3$) of minimum 3 days Dose providing exposure at rodent NOAEL Safety pharmacology and genotoxicity studies |
| Dosing up to 14 days Dose up to therapeutic range Maximum dose yields up to exposure at NOAEL in most sensitive species, or, if no toxicity is seen in either species, $1/10\text{th}$ the high-dose exposure (using the lower between species) | Two-week toxicity studies in both rodent and nonrodent Maximum dose based on exposure multiples Safety pharmacology and genotoxicity studies |

Note: The five examples of exploratory clinical study approaches provided in ICH M3(R2) and the corresponding recommendations for nonclinical studies are summarized.

i.v., intravenous; MFD, maximum feasible dose; MTD, maximum tolerated dose; NOAEL, no-observed-adverse-effect level; PAD, pharmacologically active dose.

**FIGURE 28.2** Traditional view based on administered dose or plasma concentration metric (AUC or C_{max}).

the monitored safety factors and pharmacokinetics are the same. As the current design for oncology is proving to have exceptional failure rates, new “adaptive” study designs are actively being pursued (Meille et al., 2008; Green and Beneditti, 2012).

With the number of drugs withdrawn from the marketplace since 1990 (or, perhaps, the degree of media coverage of such withdrawals), public concern with the workings of the drug safety evaluation aspects of the development process has risen sharply (Ganter, 1999; Wechsler, 2001).

TABLE 28.7 Selected Reports of Genetic Associations with Drug-Induced Liver Injury

| Drug | Gene(s) | Drug Class | Form of Toxicity | Cases/Controls | References |
|---------------------------------|--|---|--|----------------|-------------------------|
| Ximelagatran | DRB1*07 DQA1*02 UGT1A16 | Oral thrombin inhibitor | Elevation in transaminase | 74/130 | Kindmark et al. (2008) |
| Tolcapone | | Catechol- <i>O</i> -methyltransferase inhibitor | Asymptomatic liver transaminase elevation | 135/274 | Acuña et al. (2002) |
| Diclofenac | UGT2B7 CYP2C8 ABCC2 | NSAID | Range from acute liver failure to nonspecific symptoms with transaminase elevation | 24/48 | Daly et al. (2007) |
| Tranilast ^a | UGT1A1 | TGT- α antagonist | Unconjugated hyperbilirubinemia | 127/99 | Danoff et al. (2004) |
| Isoniazid | CYP2E1 NAT2 | Antibiotic | Elevation in serum transaminases | 49/269 | Huang et al. (2003) |
| Isoniazid | GSTM1 NAT2 | Antibiotic | Icteric hepatitis (serum bilirubin >3.0 mg dL ⁻¹) | 37/33 | Roy et al. (2001) |
| Isoniazid | DRB1*03 | Antibiotic | Elevation in serum bilirubin or transaminases | 22/134 | Sharma et al. (2002) |
| Rifampin | DQA1*0102 | | | | |
| Ethambutol | DQB1*0201 | | | | |
| Pyrazinamide | | | | | |
| Streptomycin | | | | | |
| Amoxicillin/ clavulanic acid | DRB1*1501DRB5* 0101DQA1*0102D QB1*0602 | Antibiotic/penicillin analog | Jaundice and elevation in serum bilirubin | 22/134 | O'Donohue et al. (2000) |
| Tacrine | GST T1 GST M1 | Parasympathomimetic | Elevation in serum transaminases | 52/89 | Simon et al. (2000) |
| Troglitazone | GSTT1 GSTM1 | Thiazolidinedione | Elevation in serum transaminases | 25/85 | Watanabe et al. (2003) |

Source: Adapted from Wilke et al. (2007, pp. 904–916).

NSAID, nonsteroidal anti-inflammatory drug; TGF- α , transforming growth factor α .

^aThis study examined the effect of tranilast on evoked Gilbert's syndrome, not hepatotoxicity.

It is currently estimated that in the United States, adverse drug reactions (ADRs) rank between the fourth and sixth leading cause of death (Eikelboom et al., 2001), though the majority of these remain misdosing or misuse. While improvements in the nonclinical of drug safety assessment (as covered in the first 19 chapters of this book) are possible and even likely, clearly the clinical aspects (the subject of this chapter) are likely to be where most improvement is likely. This will come from both better selection of subjects for inclusion in trials and a better understanding of individual or subpopulation differences in human responses to drugs.

While there is much press about the concern that the “increased pace of drug approval” has caused the release onto the market of less safe drugs (Willman, 2000), the causes are more mundane and of much longer standing. Additionally, the actual flow of new drugs entering the market is reduced. An important reason for the high incidence of serious and fatal ADRs is that the existing drug development paradigms do not generate adequate information on the mechanistic sources of marked variability in pharmacokinetics and pharmacodynamics of new therapeutic candidates, precluding treatments from being tailored for individual patients (Ozdemir et al., 2001).

Pharmacogenetics is the study of the hereditary basis of person-to-person variations in drug response. The focus of pharmacogenetic investigations has traditionally been unusual and extreme drug responses resulting from a single gene effect. The Human Genome Project and advancements

in molecular genetics have provided an unprecedented opportunity to study all genes in the human genome, including genes for drug metabolism, drug targets, and post-receptor second messenger mechanisms, in relation to variability in drug safety and efficacy. In addition to sequence variations in the genome, high-throughput and genome-wide transcript profiling for differentially regulated mRNA species before and during drug treatment will serve as important tools to uncover novel mechanisms of drug action. Pharmacogenetic-guided drug discovery and development represent a departure for the conventional approach which markets drugs for broad patient populations, rather than smaller groups of patients in whom drugs may work more optimally.

Pharmacogenetics provides a rational framework to minimize the uncertainty in outcome of drug therapy and clinical trials and thereby should significantly reduce the risk of drug toxicity. The reader is referred to the Internet sources in Table 28.8 for more details on pharmacogenetics and drug development. Potential improvements in patient inclusion criteria will be addressed later in this chapter.

28.1.1 Pharmacokinetics

Current costs and time pressures for developing a new therapeutic motivate companies to make the best possible decision as to whether to continue or abandon the development of a new drug based on a likelihood matrix for three factors: does it work (efficacy), is it acceptably tolerated

TABLE 28.8 Educational Internet Sources on Pharmacogenetics and Drug Development

| Source | Focus | Web Address |
|---|---------------------------------------|---|
| Affymetrix | DNA microarray technology | www.affymetrix.com |
| Celera Genomics | Human genome sequencing and variation | www.celera.com |
| Center for Drug Development Science | Drug development | www.dml.georgetown.edu/depts/pharmacology/cdds/index.html |
| Center for Ecogenetics and Environmental Health | Gene–environment interactions | depts.washington.edu/ceeh |
| Cold Spring Harbor Laboratory | Genetics education | www.cshl.org |
| Food and Drug Administration | Drug development and regulation | www.fda.gov |
| Genaissance Pharmaceuticals | Human genetic variation | www.genaissance.com |
| Human Genome Project | Human genetic variation | https://www.genome.gov/10001772/all-about-the-human-genome-project-hgp/ |
| Karolinska Institutet | Genetics of drug metabolism | http://www.cypalleles.ki.se/ |
| National Institutes of Health | Glossary of genetic terms | https://www.genome.gov/glossary/ |
| Nature Genetics | Genomics | http://www.nature.com/genomics |
| Pharsight Corporation | Drug development | www.pharsight.com |
| SNP Consortium | Human genetic variation | |
| Stanford University | Genome resources | www-genome.stanford.edu/index.html |
| Whitehead Institute | Genome resources | www-genome.wi.mit.edu |

SNPs, single nucleotide polymorphisms.

at therapeutic doses (safety), and is it possible to deliver those therapeutic doses to the target sites/organs economically via the desired route (bioavailability)? These likelihoods change most rapidly in early phase development, and most compounds abandoned during clinical development are abandoned in this phase, with the most common reason being unsuitable pharmacokinetics (is it absorbed, and does it stay at therapeutic levels for an optimal or near optimal period?) (Rolan, 1997). Hence the assessment of pharmacokinetics in early phase drug development is strategically important, but as discussed poorly targeted, it is receptor selectivity and occupancy at therapeutic targets that is critical but rarely addressed—not plasma levels, but their surrogate. Some of the drug development issues which are likely to be answered at least in part by a thoughtful interpretation of PK data include the following:

1. Is the compound adequately absorbed to be likely to have a therapeutic effect?
2. Is the compound absorbed with a speed consistent with the desired clinical response?
3. Does the compound stay in the body long enough to be consistent with the desired duration of action?
4. Is the within- or between-subject variability acceptable given the likely therapeutic index of the compound?
5. Is there evidence of a formulation problem?
6. Is there a dose range which produces plasma (or tissue) concentrations which are likely to be associated with a desired clinical response, or which gives rise to safety concerns?
7. Is there a relationship between plasma concentrations and a relevant measure of drug effect?
8. Are metabolites produced which may confound the therapeutic response or safety profile?
9. From the absorption, metabolisms, and excretion profile, are there subsets of the target population which may behave differently from expected?
10. Considering the above issues, what is a suitable dosing regimen for clinical efficacy trials?

As most drugs are preferably given orally, absorption which is complete, consistent, and predictable is desirable. Although it may be possible from solubility, lipophilicity, pK_a , molecular size, and animal data to make some prediction about likely absorption, only a study in humans will give quantitative data as the mechanisms of drug absorption are complex and still incompletely understood (Washington et al., 2001). It may be helpful here to distinguish between the terms “absorption” and “bioavailability.”

“Absorption” refers to the fraction of the administered dose which is taken into the body. If a drug is taken up into intestinal cells but then extensively metabolized, it is still

regarded as having been absorbed. However, for drug to be “bioavailable,” unchanged drug must reach the systemic circulation. Hence a drug with a very high first-pass metabolism might be well absorbed but poorly bioavailable. Although in therapeutic terms poor absorption and poor bioavailability pose similar problems, it is important to distinguish between them, because there are likely to be different possible solutions. Poor absorption might be approached by reformulation, change in the route of administration, or the development of a prodrug; extensive presystemic metabolism might only be avoided by change in the route of administration or chemical modification. Poor absorption is still frequently encountered in modern drug development, because the rational drug discovery process often puts more emphasis on potency and selectivity (because these programs are run by biochemists and pharmacologists) than factors likely to be associated with good absorption. This can result in lead compounds which perform very well *in vitro* but which may present major bioavailability and/or formulation problems (see discussion of this by Rolan et al., 1994) *in vivo* and in the clinic.

Quantitative assessment of the extent of absorption (absolute bioavailability) is most rigorously obtained by comparison of the areas under the plasma concentration–time curves (after adjusting for dose) following intravenous (IV) and oral administration. However, even after oral administration alone, some idea of absorption or bioavailability can be obtained in the following ways:

1. If a drug is not substantially metabolized, urinary excretion of unchanged drug may be a useful measure of absorption and bioavailability.
2. If a drug is substantially metabolized but it is reasonable to assume that metabolites are not produced in the gut lumen, urinary recovery of drug and metabolites might be a useful measure of absorption.
3. If the “apparent” plasma clearance (dose/area under the plasma concentration–time curve; equivalent to true clearance/fraction of dose absorbed) gives an implausibly high value of clearance (e.g., greater than hepatic and renal plasma flow), it is likely the bioavailability is low. However, this could be due to presystemic metabolism in addition to low absorption.
4. If there is a very large within- or between-subject variability in “apparent” clearance, this might indicate variable absorption or bioavailability, which in turn is often seen when absorption or bioavailability is low.

Determining whether absorption is related to the formulation or to an intrinsic property of the molecule can be obtained by comparing absorption from a solid formulation and an oral solution, ideally with an IV solution as a reference.

Some idea of the rate of absorption can be obtained from examination of the plasma concentration–time profile. It should be remembered, however, that the time to maximum plasma concentration (t_{\max}) is not when absorption is complete but when the rates of drug absorption and elimination are equal. Thus two drugs with the same absorption rate will differ in t_{\max} if elimination rates differ. Assessment of the rate of absorption can also be confounded by complex or slow drug distribution. For example, the calcium channel blocker amlodipine has a much later t_{\max} than other similar drugs. This is not due to slow absorption but due to partitioning in the liver membrane with slow redistribution. A quantitative assessment of the rate of absorption can be obtained by deconvolution of plasma profiles following IV and oral administration.

28.1.1.1 *Relating the Time Course of Plasma Concentrations to the Time Course of Effect*

A critical decision to be made after the first human study is whether the compound's speed of onset and duration of action are likely to be consistent with the desired clinical response. Speed of onset is clearly of interest for treatments which are taken intermittently for symptom relief, for example, acute treatments for migraine, analgesics, or antihistamines for hay fever. Duration of action is particularly important when the therapeutic effect needs to be sustained continuously, for example, anticonvulsants. The first information on the probable time course of action often comes from the plasma PK profile. However it has become increasingly evident that the profile alone may be misleading, with the concentration–time and the effect–time curves being substantially different. Some reasons for this, with examples, include the following:

1. The effect may be delayed with respect to plasma concentration because of slow uptake into the target tissue from the plasma. A well-known example is digoxin, where there is a delay of several hours between peak plasma concentration and peak effect.
2. The effect may wane faster than the plasma elimination curve due to tolerance, for example, benzodiazepines and nitrates.
3. The effect may persist despite apparent elimination from plasma. This can occur with an irreversible effect of the drug (e.g., acetylation of platelet cyclooxygenase by aspirin). Another reason is very tight binding of the drug near the receptor (e.g., salmeterol) or concentration and trapping in the target tissue (omeprazole).
4. The formation of active metabolites may also contribute to a delay in onset and/or prolongation of action.

Some of these mechanisms may become apparent during animal pharmacology studies, but the clinical pharmacologist

must always be aware of the possible discrepancy between concentration and effect–time curves. Clearly, if a relevant drug effect can also be measured in early human studies, establishing a relationship between plasma concentration and effect may be possible. If the desired clinical effect can be measured directly (e.g., blood pressure for an antihypertensive drug), the PK profile may not contribute greatly to the assessment of time course of action, but these circumstances are the exception rather than the rule. Because of the many causes of discrepancies between the time course of drug concentrations and effect, and often the difficulty in measuring the clinical effects directly, a potentially useful approach comes from the use of surrogate markers of drug effect (discussed elsewhere in this book) combined with PK/pharmacodynamic (PD) modeling to explore the relationships between dose, plasma concentrations, and effects.

28.1.2 *Safety of Clinical Trial Subjects*

While there continues to be increased interest in and concern about the safety of marketed (“approved”) drugs (e.g., Ioannides and Lau (2001) have published a study showing that adverse safety findings is frequently low compared to actual numbers), as of this writing the public's confidence in the safety of participants in trials is at a low level (Lee et al., 2006). Certainly, the FIAU tragedy (Meinert, 1996) and the case of the death of a healthy volunteer in a Johns Hopkins trial brought this issue to the forefront of the public mind. Amid estimates of as many as 5000 subject deaths per year in federally funded clinical trials (out of seven million individuals enrolled in such trials) (Davis, 1998; Wilson, 1998; Association of American Universities, 2000; Henney, 2000; Shamoo, 2000), the current guidelines and procedures should be clearly understood and carefully adhered to, but continue to change. As a starting place, Table 28.9 presents a glossary of key terms employed in this discussion.

There are international regulations which govern (on a country-by-country basis) the conduct of clinical trials, with many national governments expecting researchers to follow specific guidelines, such as the International Conference on Harmonization (ICH) (1996 and 2006) Guideline for Good Clinical Practice. Regulations and guidelines are generally based on the principles of the Nuremberg Convention. The Nuremberg Code was written in 1946 in an effort to prevent recurrence of the human experimentation atrocities of World War II. This document states that all research in humans should be done with the well-being of the subject of primary concern (O'Donnell, 2005; Schmidt, 2001).

The 1964 Declaration of Helsinki includes significant detail about clinical trial practices and the rights of potential subjects to be informed about risks, benefits, and alternative therapies (World Medical Association, 2008). It has been amended several times, most recently in 2000, when the use of placebos in trials employing patients was

TABLE 28.9 Key Terms

| | |
|---|---|
| Adverse event (or adverse experience): | Any untoward medical occurrence in a patient or clinical investigation subject administered a pharmaceutical product and which does not necessarily have to have a causal relationship with this treatment |
| Adverse drug reaction (ADR): | In the preapproval clinical experience with a new medicinal product or its new usages, particularly as the therapeutic dose(s) may not be established, “all noxious and unintended responses to a medicinal product related to any dose should be considered adverse drug reactions” |
| IND: | Investigational New Drug application, filed with the FDA after preclinical testing, is complete asking for permission to proceed with human tests |
| Efficacy pharmacology: | Evaluation of a drug’s characteristics, effects, and uses with regard to the target illness and its interactions with living organisms |
| Healthy volunteer: | A healthy person who agrees to participate in a clinical trial for reasons other than medical and receives no direct health benefit from participating |
| Human subject: | An individual who is or becomes a participant in research, either as a recipient of the test article or as a control. A subject may be either a healthy human or a patient (21 CFR 50.3) |
| Nonclinical studies: | Studies in living systems (animals or cells) other than humans |
| Phase I: | Initial safety trials on a new medicine in which investigators attempt to establish the dose range tolerance for single and multiple doses in about 20–80 healthy volunteers |
| Phase II: | Pilot clinical trials to evaluate efficacy, safety, and therapeutic dose ranges in selected populations of about 100–300 subjects who have the disease or condition to be treated, diagnosed, or prevented |
| Phase III: | Multicenter studies in populations of perhaps 100–3000 subjects (or more) for whom the medicine is eventually intended |
| Phase IV: | Postmarketing trials to provide additional details about the product’s safety, efficacy, and additional uses |
| Preclinical studies: | Animal studies that support phase I safety and tolerance studies and must comply with good laboratory practice (GLP). Other preclinical studies are done in discovery research laboratories to support drug efficiency claims |
| Serious adverse event: | A serious adverse event (experience) or reaction is any untoward medical occurrence that at any dose: <ul style="list-style-type: none"> • Results in death • Is life threatening • Requires inpatient hospitalization or prolongation of existing hospitalization • Results in persistent or significant disability/incapacity • Is a congenital anomaly/birth defect |
| Subject/trial subject: | An individual who participates in a clinical trial, either as recipient of the investigational product(s) or as a control. (ICH) See also healthy volunteers, human subject (ICH 1.57) |
| Unexpected adverse drug reaction: | An adverse reaction, the nature or severity of which is not consistent with the applicable product information (e.g., investigator’s brochure for an unapproved investigational medicinal product) |

Source: Data from Machin et al. (2007) and Edwards et al. (2010).

pronounced to be unethical (Mackintosh, 2001). Together, several parts of the US Code of Federal Regulations (CFR) (21 CFR 50, 21 CFR 54, 21 CFR 56, 21 CFR 312) constitute the good clinical practice (GCP) regulations for studies conducted in the United States. The regulations detail the responsibilities of sponsors, investigators, and institutional review boards (IRBs) and also outline monitoring practices to ensure regulatory and study design compliance and subject safety. Similarly, the ICH guidelines on GCP provide detailed instructions for investigators, institutions, sponsors, and IRBs.

As with preclinical matters, during the 1990s, the ICH has brought together regulatory agencies and industry representatives from the United States, Europe, and Japan—and observers from all over the world—to agree to a single set of technical requirements for the registration of pharmaceuticals for human use. This process is now almost complete. The ICH Guideline for Good Clinical Practice has been adopted by the three lead regions and by many other countries (ICH, 1996). As developing nations begin establishing

practices for the testing and registration of new molecular entities, many are using ICH guidelines as standards.

Thus, during the past 50 years, the conduct of clinical drug research has improved because of regulations, guidelines, and policies put in place to protect subjects. Individual pharmaceutical companies have used these guidelines and regulations as the basis for their standard operating procedures (SOPs), technical operations policies, and training programs to direct work processes and staff in their research. Most companies have created quality assurance (QA) units to oversee their researchers’ adherence to agency guidelines and regulations and to their own company policies and practices.

In the United States, no clinical drug research can begin without prior FDA review of the IND application (with minor exceptions for research imaging agents and radiopharmaceuticals), which includes the human testing protocol and associated preclinical testing results. While formal FDA approval of such an application is not legally a requirement, a 30-day waiting period for FDA review before initiating

clinical trials and assent before proceeding is a prudent goal. Regulators in some countries require only notification of intent to initiate first-in-human studies. Before moving on to phase II or phase III studies, pharmaceutical companies and other sponsors must submit the information gathered to date for agency reviews.

For new chemical entities (NCEs), new indications or new formulations companies must file an IND application, which must be approved by the FDA before a drug can be used in humans. The requirements for ADR reporting for INDs are thus known as the IND regulations. Before a new product can be marketed, companies must file a new drug application (NDA) and have it approved by the FDA. The requirements for ADR reporting after marketing are thus known as the NDA regulations. Both sets of regulations can apply to a drug at the same time; for instance, the NDA regulations apply to any marketed forms, but the IND regulations apply to a new indication or formulation. At the time of writing, the current regulations are in Title 21 of the *CFR* (21 *CFR*) as follows:

- 21 *CFR* 312.32 Safety reports for investigational products subject to an IND application (published 1987)
- 21 *CFR* 314.80 Postmarketing reporting of ADEs (NDA) (published 1985)
- 21 *CFR* 600.80 Postmarketing reporting of adverse experiences for licensed biological products (includes vaccines) (published 1994)

See also *CFR* website: <http://www.access.gpo.gov/nara/cfr/index.html>. An August 1997 guideline *Postmarketing Adverse Experience Reporting for Human Drug and Licensed Biological Products: Clarification of What to Report* defined the minimum data relevant for a safety report as:

- An identifiable patient
- An identifiable reporter
- A suspect drug or biological product
- An AE or fatal outcome

If any of these items remain unknown after being actively sought, a report should not be submitted to the FDA. The guideline also clarifies that adverse experiences derived during planned contacts and active solicitation of information from patients (e.g., company-sponsored patients support programs, disease management programs) should be handled as safety information from a postmarketing study (i.e., for expedited reporting, events must be serious and unexpected, with a reasonable possibility that the drug caused the event).

See also Center for Drug Evaluation and Research (CDER) guidance page <http://www.fda.gov/Drugs/20GuidanceComplianceRegulatoryInformation/Guidances/default.htm>.

These regulations and the guidelines have most recently been extensively indexed in a publication in the *Drug Information Journal*: Curran and Sills (1997).

In the *Federal Register* of October 27, 1994, the FDA published a rule to amend the regulations to provide consistency with certain standardized definitions, procedures, and formats developed by the ICH and CIOMS and finally published detailed amended expedited safety reporting regulations, implementing ICH in October 1997 (FDA, 1997a). A revision of the associated guideline has also been proposed by the FDA, but was not available at the time of writing.

These new regulations for expedited reporting became effective 180 days later on April 6, 1998. Key points from the new regulations are described later. The amendments to the periodic postmarketing safety reporting regulations are delayed awaiting further consideration of the ICH E2C guideline. More recently, as of the start of 2008, all new drug clinical trials involving patients must be registered with ClinicalTrials.gov as part of the IND process. The form required to register, FDA-3674, is provided as Figure 28.3.

28.1.2.1 IND Regulations In 1994 the FDA proposed to amend requirements for clinical study design and conduct an annual sponsor reporting in the IND regulations as a result of events with fialuridine. In the light of comments received, the FDA withdrew the proposed amendments and will develop a guidance document with recommendations on study design and monitoring of investigational drugs used to treat serious and potentially fatal illnesses, with particular attention to detection of AEs similar to those caused by underlying disease.

28.1.2.2 Increased Frequency Reports The requirement for increased frequency reports for serious expected ADRs with marketed products is revoked. This was also published in FDA (1997b). The rationale for this was that despite receiving many such reports, only a small number of drug safety problems were identified.


28.1.2.3 Reporting Forms FDA form 3500/3500A (refer back to Figure 28.1) is the standard form for notifying expedited reports and can also be used by companies to submit IND safety reports. Foreign cases may be reported on the CIOMS I form.

28.1.2.4 Definitions The definition of “serious” has been revised to make it consistent with ICH E2A and is the same for INDs and NDAs (see preceding text).

The definition of “unexpected” for IND reporting is “Any adverse drug experience, the specificity or severity of which is not consistent with the current investigator brochure; or if an investigator brochure is not required or available, the specificity or severity of which is not consistent with the risk information described in the general investigational plan or

(a)

See OMB Statement on Reverse. Form Approved: OMB No. 0910-0616, Expiration Date: 06-30-2008

| | | |
|--|---|--|
|  DEPARTMENT OF HEALTH AND HUMAN SERVICES Food and Drug Administration Certification of Compliance, under 42 U.S.C. § 282(j)(5)(B), with Requirements of ClinicalTrials.gov Data Bank (42 U.S.C. § 282(j)) | | |
| (For submission with an application/submission, including amendments, supplements, and resubmissions, under §§ 505, 515, 520(m), or 510(k) of the Federal Food, Drug, and Cosmetic Act or § 351 of the Public Health Service Act.) | | |
| SPONSOR / APPLICANT / SUBMITTER INFORMATION | | |
| 1. NAME OF SPONSOR/APPLICANT/SUBMITTER <div style="border: 1px solid black; height: 20px; width: 100%;"></div> | 2. DATE OF THE APPLICATION/SUBMISSION WHICH THIS CERTIFICATION ACCOMPANIES <div style="border: 1px solid black; height: 20px; width: 100%;"></div> | |
| 3. ADDRESS (Number, Street, State, and ZIP Code) <div style="border: 1px solid black; height: 40px; width: 100%;"></div> | 4. TELEPHONE AND FAX NUMBER (Include Area Code) (Tel.) <div style="border: 1px solid black; width: 100%;"></div> (Fax) <div style="border: 1px solid black; width: 100%;"></div> | |
| PRODUCT INFORMATION | | |
| 5. FOR DRUGS/BIOLOGICS: Include Any/All Available Established, Proprietary and/or Chemical/Biochemical/Blood/Cellular/Gene Therapy Product Name(s) FOR DEVICES: Include Any/All Common or Usual Name(s), Classification, Trade or Proprietary or Model Name(s) and/or Model Number(s) (Attach extra pages as necessary) | | |
| | | |
| APPLICATION / SUBMISSION INFORMATION | | |
| 6. TYPE OF APPLICATION/SUBMISSION WHICH THIS CERTIFICATION ACCOMPANIES <input type="checkbox"/> IND <input type="checkbox"/> NDA <input type="checkbox"/> ANDA <input type="checkbox"/> BLA <input type="checkbox"/> PMA <input type="checkbox"/> HDE <input type="checkbox"/> 510(k) <input type="checkbox"/> PDP <input type="checkbox"/> Other | | |
| 7. INCLUDE IND/NDA/ANDA/BLA/PMA/HDE/510(k)/PDP/OTHER NUMBER (If number previously assigned) <div style="border: 1px solid black; width: 100%; height: 20px;"></div> | | |
| 8. SERIAL NUMBER ASSIGNED TO APPLICATION/SUBMISSION WHICH THIS CERTIFICATION ACCOMPANIES <div style="border: 1px solid black; width: 100%; height: 20px;"></div> | | |
| CERTIFICATION STATEMENT / INFORMATION | | |
| 9. CHECK ONLY ONE OF THE FOLLOWING BOXES (See instructions for additional information and explanation) <input type="checkbox"/> A. I certify that the requirements of 42 U.S.C. § 282(j), Section 402(j) of the Public Health Service Act, enacted by 121 Stat. 823, Public Law 110-85, do not apply because the application/submission which this certification accompanies does not reference any clinical trial. <input type="checkbox"/> B. I certify that the requirements of 42 U.S.C. § 282(j), Section 402(j)v of the Public Health Service Act, enacted by 121 Stat. 823, Public Law 110-85, do not apply to any clinical trial referenced in the application/submission which this certification accompanies. <input type="checkbox"/> C. I certify that the requirements of 42 U.S.C. § 282(j), Section 402(j) of the Public Health Service Act, enacted by 121 Stat. 823, Public Law 110-85, apply to one or more of the clinical trials referenced in the application/submission which this certification accompanies and that those requirements have been met. | | |
| 10. IF YOU CHECKED BOX C, IN NUMBER 9, PROVIDE THE NATIONAL CLINICAL TRIAL (NCT) NUMBER(S) FOR ANY "APPLICABLE CLINICAL TRIAL(S)," UNDER 42 U.S.C. § 282(j)(1)(A)(i), SECTION 402(j)(1)(A)(i) OF THE PUBLIC HEALTH SERVICE ACT, REFERENCED IN THE APPLICATION/SUBMISSION WHICH THIS CERTIFICATION ACCOMPANIES (Attach extra pages as necessary) NCT Number(s): <div style="border: 1px solid black; width: 100%; height: 20px;"></div> | | |
| The undersigned declares, to the best of her/his knowledge, that this is an accurate, true, and complete submission of information. I understand that the failure to submit the certification required by 42 U.S.C. § 282(j)(5)(B), section 402(j)(5)(B) of the Public Health Service Act, and the knowing submission of a false certification under such section are prohibited acts under 21 U.S.C. § 331, section 301 of the Federal Food, Drug, and Cosmetic Act. Warning: A willfully and knowingly false statement is a criminal offense, U.S. Code, title 18, section 1001. | | |
| 11. SIGNATURE OF SPONSOR/APPLICANT/SUBMITTER OR AN AUTHORIZED REPRESENTATIVE (Sign) | 12. NAME AND TITLE OF THE PERON WHO SIGNED IN NO. 11 (Name) <div style="border: 1px solid black; width: 100%; height: 20px;"></div> (Title) <div style="border: 1px solid black; width: 100%; height: 20px;"></div> | |
| 13. ADDRESS (Number, Street, State, and ZIP Code) (of person identified in No. 11 and 12) <div style="border: 1px solid black; height: 40px; width: 100%;"></div> | 14. TELEPHONE AND FAX NUMBER (Include Area Code) (Tel.) <div style="border: 1px solid black; width: 100%;"></div> (Fax) <div style="border: 1px solid black; width: 100%;"></div> | 15. DATE OF CERTIFICATION <div style="border: 1px solid black; width: 100%; height: 20px;"></div> |

FDA-3674 (1/08) (FRONT)

PSC Graphics: (301) 443-1090 EF

FIGURE 28.3 FDA form 3674.

(b)

| Instructions for Completion of Form FDA 3674 | | |
|---|---|---|
| <p>Certification of Compliance, under 42 U.S.C. § 282(j)(5)(B), with Requirements of ClinicalTrials.gov Data Bank (42 U.S.C. § 282(j)) Form 3674 must accompany an application/submission, including amendments, supplements, and resubmissions, submitted under §§ 505, 515, 520(m), or 510(k) of the Federal Food, Drug, and Cosmetic Act or § 351 of the Public Health Service Act.</p> | | |
| <p>1. Name of Sponsor/Applicant/Submitter - This is the name of the sponsor/applicant/submitter of the drug/biologic/device application/submission which the certification accompanies. The name must be identical to that listed on the application/submission.</p> | | |
| <p>2. Date - This is the date of the application/submission which the certification accompanies.</p> | | |
| <p>3. & 4. - Provide complete address, telephone number and fax number of the sponsor/applicant/submitter.</p> | | |
| <p>5. Product Information - For Drugs/Biologics: Provide the established, proprietary name, and/or chemical/biochemical/blood product/cellular/gene therapy name(s) for the product covered by the application/submission. Include all available names by which the product is known. For Devices: Provide the common or usual name, classification, trade or proprietary or model name(s), and/or model number(s). Include all available names/model numbers by which the product is known.</p> | | |
| <p>6. Type of Application/Submission - Identify the type of application/submission which the certification accompanies by checking the appropriate box. If the name of the type of application/submission is not identified, check the box labeled "Other."</p> | | |
| <p>7. IND/NDA/ANDA/BLA/PMA/HDE/510(k)/PDP/Other Number - If FDA has previously assigned a number associated with the application/submission which this certification accompanies, list that number in this field. For example, if the application/submission accompanied by this certification is an IND protocol amendment and the IND number has already been issued by FDA, that number should be provided in this field.</p> | | |
| <p>8. Serial Number - In some instances a sequential serial number is assigned to the application. If there is such a serial number, provide it in this field.</p> | | |
| <p>9. Certification - This section contains three different check-off boxes.</p> <p>Box A should be checked if the sponsor/applicant/submitter has concluded that the requirements of 42 U.S.C. § 282(j), section 402(j) of the Public Health Service Act, do not apply because no clinical trials are included, relied upon, or otherwise referred to, in the application/submission which the certification accompanies.</p> <p>Box B should be checked if the sponsor/applicant/submitter has concluded that the requirements of 42 U.S.C. § 282(j), section 402(j) of the Public Health Service Act, do not apply at the time of submission to any clinical trial that are included, relied upon, or otherwise referred to, in the application/submission which the certification accompanies. This means that, at the time the application/submission is being made, the requirements of 42 U.S.C. § 282(j), section 402(j) of the Public Health Service Act, do not apply to any of the clinical trials included, relied upon, or otherwise referred to, in the application/submission which this certification accompanies.</p> <p>Box C should be checked if the sponsor/applicant/submitter has concluded that the requirements of 42 U.S.C. § 282(j), section 402(j) of the Public Health Service Act, do apply at the time of submission to some or all of the clinical trials that are included, relied upon, or otherwise referred to, in the application/submission which the certification accompanies. This means that, at the time the application/submission is being made, the requirements of 42 U.S.C. § 282(j), section 402(j) of the Public Health Service Act, apply to one or more of the clinical trials included, relied upon, or otherwise referred to, in the application/submission which this certification accompanies.</p> | | |
| <p>10. National Clinical Trial (NCT) Numbers - If you have checked Box C in number 9 (Certification), provide the NCT Number obtained from www.ClinicalTrials.gov for each clinical trial that is an "applicable clinical trial" under 42 U.S.C. § 282(j)(1)(A)(i), section 402(j)(1)(A)(i) of the Public Health Service Act, and that is included, relied upon, or otherwise referred to, in the application/submission which the certification accompanies. Type only the number, as NCT will be added automatically before number. Include any and all NCT numbers assigned to the clinical trials included, relied upon, or otherwise referred to, in the application/submission which this certification accompanies. Multiple NCT numbers may be required for a particular certification, depending on the number of "applicable clinical trials" included, relied upon, or otherwise referred to, in the application/submission which the certification accompanies.</p> | | |
| <p>11. Signature of Sponsor/Applicant/Submitter or an Authorized Representative - The person signing the certification must sign in this field.</p> | | |
| <p>12. Name and Title of Person Who Signed in number 11. - Include the name and title of the person who is signing the certification. If the person signing the certification is not the sponsor/applicant/submitter of the application/submission, he or she must be an authorized representative of the sponsor/applicant/submitter.</p> | | |
| <p>13. & 14. - Provide the full address, telephone and fax number of the person who is identified in number 11 and signs the certification in number 11.</p> | | |
| <p>15. Provide the date the certification is signed. This date may be different from the date provided in number 2.</p> | | |
| Paperwork Reduction Act Statement | | |
| <p>Public reporting burden for this collection of information is estimated to average 15 minutes and 45 minutes (depending on the type of application/submission) per response, including time for reviewing instructions. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden to the applicable address below.</p> | | |
| Food and Drug Administration Center for Drug Evaluation and Research Central Document Room Form No. FDA 3674 5901-B Ammendale Road Beltsville, MD 20705-1266 | Food and Drug Administration Center for Biologics Evaluation and Research 1401 Rockville Pike Rockville, MD 20852-1448 | Food and Drug Administration Center for Devices and Radiological Health Program Operations Staff (HFZ-403) 9200 Corporate Blvd. Rockville, MD 20850 |
| <p><i>An agency may not conduct or sponsor, and a person is not required to respond to, a collection of information, unless it displays a currently valid OMB control number.</i></p> | | |

FDA-3674 (1/08) (BACK)

FIGURE 28.3 (Continued)

elsewhere in the current application, as amended. For example, under this definition, hepatic necrosis would be unexpected (by virtue of greater severity) if the investigator brochure only referred to elevated hepatic enzymes or hepatitis. Similarly, cerebral thromboembolism and cerebral vasculitis would be unexpected (by virtue of greater specificity) if the investigator brochure only listed cerebral vascular accidents. ‘Unexpected,’ as used in this definition, refers to an adverse drug experience that has not been previously observed (e.g., included in the investigator brochure) rather than from the perspective of such experience not being anticipated from the pharmacological properties of the pharmaceutical product.”

28.1.2.5 Time Frames The time period for submitting written IND safety reports has been revised from 10 working days to 15 calendar days. For telephone reports (fatal and life-threatening unexpected reactions), it has been revised from three working days to seven calendar days. Such reports can also be made by fax. Telephone reporting was previously restricted to clinical studies conducted under the IND, but under the new rule, telephone reporting within seven calendar days applies to any unexpected fatal or life-threatening reaction from any source.

The time period for submitting NDA alert reports (serious and unexpected) has been revised from 15 working days to 15 calendar days.

Wherever human drug research is conducted, national regulations call for an independent ethical review of the study plan. In countries where a guideline on GCP is used, ethics review bodies are made up of medical professionals from the institution, nonmedical personnel, and community members. Sponsor companies and involved investigators have no voting representation on these review boards, nor may they be present during the voting on the research approval. The investigator conducting the study may, however, present the protocol and answer questions at the IRB review meeting. The company sponsoring the trial is not allowed to participate in IRB meetings as a matter of routine, although a representative might be invited to explain or clarify the protocol to the ethics review body.

Informed consent must be obtained from study participants—in writing—before any study-related activities are performed. Regulations clearly describe the required elements of the consent document and the consent process to be followed. A good informed consent process can help ensure that potential subjects understand the nature of the studies they will enter, the type of treatments they will undergo, alternative therapies currently available, and any particular hazards they might experience. They must be informed that they can withdraw from the study at any time without penalty. Subjects are to be asked for their consent to release information from their medical records and told that the

medical information may be inspected by the sponsor, company, and regulatory agency representatives. They are to be informed that the results of the trials may be used publicly, but anonymously.

Drug supplies must be accounted for throughout the trial and reconciled at the end of the trial. These practices are designed to prevent the misuse or inappropriate redistribution of the investigational drug and to help ensure compliance with the protocol.

AEs, unexpected drug reactions, and drug side effects experienced and reported by subjects or observed by clinical investigators are all to be recorded and promptly reported to sponsor companies. It is clear, however, that there remain differences in reporting standards between companies and between countries (Hayashi and Walker, 1996). There is also, however, a marked difference in reporting standards and rates in US clinical trials between different medical areas (Ioannides and Lau, 2001). The investigators involved with subject care and the pharmaceutical company sponsor must then analyze each event for “causality” and “relatedness” to administration of the drug—did this reaction occur because of the drug or because of something else such as the progression of the disease symptoms, other medications being taken, or unrelated causes? Bradford Hill’s criteria for causation (Hill, 1966), originally developed in conjunction with evaluating the relationship between smoking and cancer, is the accepted fundamental standard for such assessments (see Table 28.10) (Hill, 1966). Such safety information is then to be forwarded to the appropriate IRB or ethics committee and regulatory agencies. The sponsor company is to send periodic updates to investigators, alerting them to new serious or unexpected drug reactions.

Company physicians are expected to continuously analyze the adverse drug event data coming in from worldwide trials for trends and patterns that could foretell a drug safety problem. Drug companies frequently set up data and safety monitoring boards (DSMBs), composed of noncompany medical experts and statisticians who impartially evaluate safety as the study progresses and are responsible for alerting the sponsor to unanticipated problems. Regulatory agencies also watch for trends, because they are often in the best position to see safety trends across classes of drugs from any different companies.

Sponsor companies use QA units independent of the clinical research group to audit medical operations. Their role is to ensure that regulatory standards and company policies and procedures for clinical research are being followed in all countries where research is being conducted.

Regulatory agencies oversee sponsor organizations, clinical trial processes, and clinical trial sites to verify that sponsors are conducting trials appropriately. Existing FDA regulations are conforming to ICH guidelines. When deficiencies are noted, agency inspectorates can restrict and penalize the offending

TABLE 28.10 Hill Criteria for Evaluation Causation

| Consistency and Unbiasedness of Findings | Confirmation of the Association by Different Investigators, in Different Populations, Using Different Methods |
|---|---|
| Strength of association | Two aspects: the frequency with which the factor is found in the disease and the frequency with which it occurs in the absence of the disease. The larger the relative risk, the more the hypothesis is strengthened |
| Temporal sequence | Obviously, exposure to the factor must occur before onset of the disease. In addition, if it is possible to show a temporal relationship, as between exposure to the factor in the population and frequency of the disease, the case is strengthened |
| Biological gradient (dose–response relationship) | Finding a quantitative relationship between the factor and the frequency of the disease. The intensity or duration of exposure may be measured |
| Specificity | If the determinant being studied can be isolated from others and shown to produce changes in the incidence of the disease, for example, if thyroid cancer can be shown to have a higher incidence specifically associated with fluoride, this is convincing evidence of causation |
| Coherence with biological background and previous knowledge | The evidence must fit the facts that are thought to be related, for example, the rising incidence of dental fluorosis and the rising consumption of fluoride are coherent |
| Biological plausibility | The statistically significant association fits well with previously existing knowledge |
| Reasoning by analogy | Use of knowledge of other cases for which a high similarity occurs |
| Experimental evidence | This aspect focuses on what happens when the suspected offending agent is removed. Is there improvement? The evidence of remission—or even resolution of significant medical symptoms—following explantation obviously would strengthen the case |

Note: It must be noted that not all criteria must be met to establish causality—indeed, such is a rare case.

academic IRB, investigator, and/or the sponsor company. Investigator sites (including, in extreme cases, entire universities) can be prohibited from conducting clinical research. Company studies can be rejected by regulatory agencies.

28.1.2.6 Continuous Safety Monitoring Medical staff members at the clinical trial site and at the sponsor company are expected to be continuously alert to ADRs or unexpected and serious medical problems that might be attributed to the new medication being tested. The investigator and the site staff examine subjects and take vital sign measurements on the schedule designated by the protocol, the guide for study conduct. Each drug, as shown in its preclinical studies, has unique characteristics and the potential for AEs or side effects that investigators need to watch for during the clinical trial. Because of this, safeguard activities are built into the protocol, such as the time intervals between subject visits, how often subjects are to be questioned and examined, the specific medical tests to be run at various time points, and the special diagnostic tests or interviews to be conducted. Staff members at the clinical site must record the medical information from these tests and from interviews and medical histories. Site staff must also transfer information from the medical source documents to the case report forms (CRFs) specific to the study. The CRFs contain key information required for the protocol. Clinical research associates and physicians are required to review the information regularly and to immediately report anything alarming to the IRB and regulatory agencies for further evaluation.

28.1.2.7 Sponsor Pharmacovigilance Dedicated departments in pharmaceutical companies—often called pharmacovigilance groups—receive, review, analyze, follow up on, and appropriately distribute safety-related information from new drug trials. These groups sometimes staff hot lines and question and answer services to provide up-to-date answers to drug-related questions. Safety information is to be reported to regulatory agencies at specified intervals and at milestone time points throughout all phases of drug development. The postapproval aspects of this effort are the core of the subject of Chapter 29.

Sponsor monitoring is another important oversight process to ensure quality, compliance, and subject safety. Monitors may be employees of the sponsor's medical staff or a contract research organization, or may be independent contractors. In each case, they represent the sponsor and visit investigator sites regularly, perhaps every 4–8 weeks. They examine subject records in detail and verify that the correct information was transferred to the clinical trial CRFs—a process called source data verification. During their site visits, monitors also examine administrative and regulatory documents, including drug supply and dosing records, AE documentation and reporting, the informed consent process and forms, and CRFs.

Ensuring subject safety must be of the foremost concern during the entire drug development process, especially in clinical trials. The gradual dosing of a drug candidate in healthy human volunteer subjects is performed under tightly controlled conditions and under the direction and scrutiny of

physicians trained in clinical pharmacology. A drug candidate's progress toward broader testing in a population of individuals with the disease or condition to be treated also moves prudently and in well-defined steps.

28.2 LIMITATIONS ON/OF CLINICAL TRIALS

Before looking more closely at the definition, structure, and designs of trials, one should understand their limitations. These are regulatory, economic, legal, and due to custom.

First, the most recent (October 2008) revision of the Declaration of Helsinki (58th World Medical Association General Assembly) called for discontinuing the use of placebo-controlled trials in patients. While this is not currently binding on US trials (the FDA has specifically said that they will not mirror this as a requirement) and is intended to protect the health of participating patient subjects by precluding having some denied existing efficacious treatments (which would be the effect in most—but not all—cases; see Al-Khatib et al., 2001), it will also likely cause the numbers of subjects required in a trial to increase. This will further stretch the economic aspects of limitations on the power of trials to assess potential drug safety in what will be the intended patient population. Trials are already very expensive—each additional subject enrolled costs \$15 000 or more (potentially much more) in a phase II or III trial.

The legal (or rather, litigation) limitations are that any AE in a trial (or resulting from it) exposes a sponsor to potential litigation. Accordingly, trials are designed to exclude not only those individuals who are not in the precisely designed subject disease population but also those who represent potential additional risk subpopulations (the elderly, the young, those currently taking other drugs, minorities, and women who are or may become pregnant) are likely to eventually use the drug when it enters the marketplace.

Custom (continuing to do things as they have previously been done) also limits the power of trials to identify safety issues. While there are now regulatory inducements to include more women, the young, and ethnic minorities in trials, the first two groups still are not proportionately incorporated because of both the perceived risks of AEs that they represent and because historically they have not been. Ethnic minorities, particularly African Americans, present a different problem in that there is a historically based resistance to participation in such trials.

28.3 THE CLINICAL TRIAL PROCESS

While the numbers of animals involved in research is tracked closely and is well known, such is not the case for human subjects involved in clinical trials. We simply do not know

how many are involved in such trials in the United States, much less worldwide. Though the National Institutes of Health (NIH) does track closely how many dollars and individuals are involved in research it funds (\$18.0 billion and 8 million subjects in 1997), the same is not true for privately funded research (where the numbers are greater). And while there is now a website where one can examine the numbers and types of efficacy trials open, the same is not true for phase I tolerance and PK trials—where most potential drugs cease development.

Clinical drug development is often described as consisting of four distinct phases (phases I–IV). It is important to recognize that the phase of development provides an inadequate basis for classification of clinical trials because one type of trial may occur in several phases (see Figure 28.4). Table 28.11 presents a preferable (objective-based) classification of trial types. It is important to appreciate that the phase label is a description, not a set of requirements. It is also important to realize that these temporal phases do not imply a fixed order of studies since for some drugs in a development plan the typical sequence will not be appropriate or necessary. For example, although human pharmacology studies are typically conducted during phase I, many such studies are conducted at each of the other three stages, but nonetheless sometimes labeled as phase I studies. Figure 28.4 demonstrates this close but variable correlation between the two classification systems. The distribution of the points of the graph shows that the types of study are not synonymous with the phases of development.

Drug development is ideally a logical stepwise procedure in which information from small early studies is used to support and plan later larger, more definitive studies. To develop new drugs efficiently, it is essential to identify characteristics of the investigational drugs in the early stages of development and to plan an appropriate development based on this profile.

Initial trials provide an early evaluation of short-term safety and tolerability and can provide PD and PK information needed to choose a suitable dosage range and administration schedule for initial exploratory therapeutic trials. Later confirmatory studies are generally larger and longer and include a more diverse patient population. Dose-response information should be obtained at all stages of development, from early tolerance studies to studies of short-term PD effect to large efficacy studies. Throughout development, new data may suggest the need for additional studies that may commonly be part of an earlier phase. For example, blood-level data in a late trial may suggest a need for a drug–drug interaction study, or AEs may suggest the need for further dose-finding and/or additional nonclinical studies. In addition, to support a new marketing application approval for the same drug, for example, for a new indication, PK or therapeutic exploratory studies are considered to

Correlation between development phases and types of study

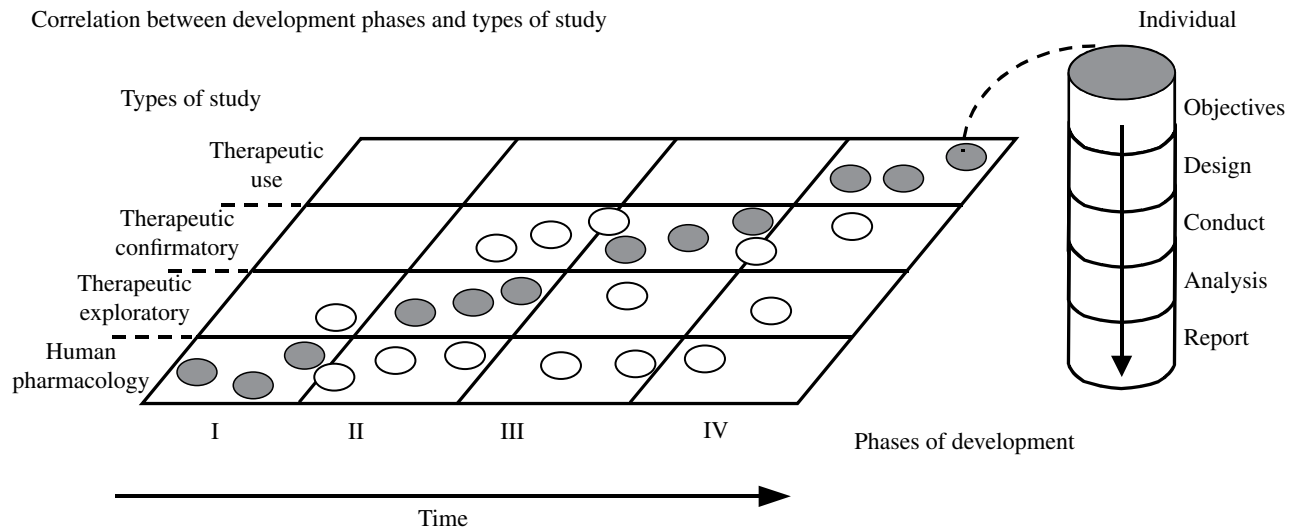


FIGURE 28.4 Matrix illustrating relationship between phases of development and types of study by objective that may be conducted during each clinical development of new medicinal product. The shaded circles show the types of the study most usually conducted in a certain phase of development; the open circles show certain types of study that may be conducted in that phase of development but are less usual. Each circle represents an individual study. To illustrate the development of a single study, one circle is joined by a dotted line to an inset column that depicts the elements and sequence of an individual study.

TABLE 28.11 An Approach to Classifying Clinical Studies according to Objective

| Type of Study | Objective of Study | Study Examples |
|---------------------------------------|--|--|
| Human pharmacology (phase I) | <ul style="list-style-type: none"> Assess tolerance Define/describe pharmacokinetics and pharmacodynamics Explore drug metabolism and drug interactions Estimate activity | <ul style="list-style-type: none"> Dose tolerance studies Single- and multiple-dose PK and/or PD studies Drug interaction studies |
| Therapeutic exploratory (phase Ib/II) | <ul style="list-style-type: none"> Explore use for the targeted indication Estimate dosage for subsequent studies Provide basis for confirmatory study design, end points, and methodologies | <ul style="list-style-type: none"> Earliest trials of relatively short duration in well-defined narrow patient populations, using surrogate or pharmacological end points or clinical measures Dose–response exploration studies |
| Therapeutic confirmatory (phase III) | <ul style="list-style-type: none"> Demonstrate/confirm efficacy Establish safety profile Provide an adequate basis for assessing the benefit/risk relationship to support licensing Establish dose–response relationship | <ul style="list-style-type: none"> Adequate and well-controlled studies to establish efficacy Randomized parallel dose–response studies Clinical safety studies Studies of mortality/morbidity outcomes Large simple trials |
| Therapeutic use (phase III/IV) | <ul style="list-style-type: none"> Refine understanding of benefit/risk relationship in general or special populations and/or environment Identify less common adverse reactions Refine dosing recommendation | <ul style="list-style-type: none"> Comparative studies Comparative effectiveness studies Studies of mortality/morbidity outcomes Studies of additional endpoints Large simple trials Pharmacoeconomic studies |

be in phase I or phase II of development (Stone, 2006; Green and Beneditti, 2012; Gallin and Ognibene, 2012).

Phase I starts with the first-in-man administration of an IND. Although human pharmacology studies are typically identified with phase I, they may also be indicated at other points in the development sequence. Studies in this phase of development usually have nontherapeutic objectives and

may be conducted in healthy volunteer subjects or certain types of patients, for example, patients with mild hypertension. Drugs with significant potential toxicity (cytotoxic drugs) or risks due to route of administration (such as intrathecal) are usually studied in patients. Studies in this phase can be open or baseline controlled or may use randomization and blinding to improve the validity of observations.

Studies conducted in phase I typically may involve one or a combination of the following:

- a. *Estimation of initial safety and tolerability* The initial and subsequent administration of an IND into humans is usually intended to determine the tolerability of the dose range expected to be needed for later clinical studies and to determine the nature of adverse reactions that can be expected. These studies typically include both single- and multiple-dose administration.
- b. *Pharmacokinetics* Characterization of a drug's absorption, distribution, metabolism, and excretion continues throughout the development plan. Such preliminary characterization is an important goal of phase I. Pharmacokinetics may be assessed via separate studies or as a part of efficacy, safety, and tolerance studies. PK studies are particularly important to assess the clearance of the drug and to anticipate possible accumulation of parent drug or metabolites and potential drug-drug interactions. Some PK studies are commonly conducted in later phases to answer more specialized questions. For many orally administered drugs, especially modified release products, the study of food effects on bioavailability is important. Obtaining PK information in subpopulations such as patients with impaired elimination (renal or hepatic failure), the elderly, children, women, and ethnic subgroups should be considered. Drug-drug interaction studies are important for many drugs and are generally performed in phases beyond phase I. But studies in animals and *in vitro* studies of metabolism and potential interactions may lead to doing such studies earlier.
- c. *Assessment of pharmacodynamics* Depending on the drug and the end points studied, PD studies and studies relating drug blood levels to response (PK/PD studies) may be conducted in healthy volunteer subjects or in patients with the target disease. In patients, if there is an appropriate measure, PD data can provide early estimates of activity and potential efficacy and may guide the dosage and dose regimen in later studies.
- d. *Early measurement of drug activity* Preliminary studies of activity or potential therapeutic benefit may be conducted in phase I as a secondary objective. Such studies are generally performed in later phases but may be appropriate when drug activity is readily measurable with a short duration of drug exposure in patients at this early stage. Frequently such evaluations are done in what are called "phase Ib" studies.

Phase II is usually considered to start with the initiation of studies in which the primary objective is to explore therapeutic efficacy in patients.

Initial therapeutic exploratory studies may use a variety of study designs, including concurrent controls and comparison with baseline status. Subsequent trials are usually randomized and concurrently controlled to evaluate the efficacy of the drug and its safety for a particular therapeutic indication. Studies in phase II are typically conducted in a group of patients who are selected by relatively narrow criteria, leading to a relatively homogeneous population, and are closely monitored.

An important goal for this phase is to determine dose levels and regimen for phase III trials. Early studies in this phase often utilize dose-escalation designs to give an early estimate of dose-response, and later studies may confirm the dose-response relationship for the indication in question by using recognized parallel dose-response designs (could also be deferred to phase III). Confirmatory dose-response studies may be conducted in phase II or deferred until phase III. Doses used in phase II are usually but not always less than the highest doses used in phase I.

Additional objectives of clinical trials conducted in phase II may include evaluation of potential study end points, therapeutic regimens (including concomitant medications), and target populations (e.g., mild vs. severe disease) for further study in phase II or III. These objectives may be served by exploratory analyses, by examining subsets of data, and by including multiple end points in trials.

Phase III usually is considered to begin with the initiation of studies in which the primary objective is to demonstrate or confirm therapeutic benefit.

Studies in phase III are designed to confirm the preliminary evidence accumulated in phase II that a drug is safe and effective for use in the intended indication and recipient population. These studies are intended to provide an adequate basis for marketing approval. Studies in phase III may also further explore the dose-response relationship, or explore the drug's use in wider populations, in different stages of disease, or in combination with another drug. For drugs intended to be administered for long periods, trials involving extended exposure to the drug are ordinarily conducted in phase III, although they may be started in phase III. ICH E1 and ICH E7 describe the overall clinical safety database considerations for chronically administered drugs and drugs used in the elderly. These studies carried out in phase III complete the information needed to support adequate instructions for use of the drug (official product information).

Phase IV begins after drug approval. Once rare, there are now commonly required therapeutic use studies that go beyond the prior demonstration of the drug's safety, efficacy, and dose definition.

Studies in phase IV are all studies (other than routine surveillance) performed after drug approval and related to the approved indication. They are studies that were not considered necessary for approval but are often important for

optimizing the drug's use. They may be of any type but should have valid scientific objectives. Commonly conducted studies include additional drug–drug interaction, dose–response or safety studies, and studies designed to support use under the approved indication, for example, mortality/morbidity studies and epidemiological studies.

28.3.1 Development of an Application Unrelated to Original Approved Use

After initial approval, drug development may continue with studies of new or modified indications, new dosage regimens, and new routes of administration or additional patient populations. If a new dose, formulation, or combination is studied, additional human pharmacology studies may be indicated, necessitating a new development plan.

The need for some studies may be obviated by the availability of data from the original development plan or from therapeutic use.

28.3.1.1 Special Considerations A number of special circumstances and populations require consideration on their own when they are part of the development plan.

Studies of Drug Metabolites Major active metabolite(s) should be identified and deserve detailed PK study. Timing of the metabolic assessment studies within the development plan depends on the characteristics of the individual drug.

Drug–Drug Interactions If a potential for drug–drug interaction is suggested by metabolic profile, by the results of nonclinical studies, or by information on similar drugs, studies on drug interaction during clinical development are highly recommended. For drugs that are frequently coadministered, it is usually important that drug–drug interaction studies be performed in nonclinical and, if appropriate, in human studies. This is particularly true for drugs that are known to alter the absorption or metabolism of other drugs or whose metabolisms or excretion can be altered by effects by other drugs.

Special Populations Some groups in the general population may require special study because they have unique risk/benefit considerations that need to be taken into account during drug development or because they can be anticipated to need modification of use of the dose or schedule of a drug compared to general adult use. PK studies in patients with renal and hepatic dysfunction are important to assess the impact of potentially altered drug metabolism or excretion. Specific ICH and FDA documents address such issues for geriatric patients and patients from different ethnic groups. The need for nonclinical safety studies to support human clinical trials in special populations is addressed in the ICH M3 document.

A key issue is thus when to perform kinetic studies in special patient groups (pediatrics, elderly, patients with renal or hepatic disease) and how. As the elderly are the majority users of many medicines, the subject of evaluating new drugs in the elderly is a major issue which is discussed in detail elsewhere. Unless a medication is unlikely to be used in the elderly, some data will be required by regulators for registration. However, an important issue in early phase development is whether to perform a separate elderly volunteer kinetic/tolerability study before elderly patients are included in later phase clinical trials. The major argument for doing so includes the possible reluctance of clinical investigators to enroll elderly patients without such data being available because of safety concerns; however, the utility of such studies has been questioned. The subjects in elderly volunteer studies are usually in much better health than the general population they are intended to represent. Also, the elderly may differ from the young not so much in terms of mean kinetic parameters, but the variability in the elderly may be much greater. The relatively small sample size (typically 12–18) may not allow a good estimation of the variability within the elderly population. For these reasons the FDA have recommended that information about the kinetics of a drug in the elderly should come from a larger group representative of the target population, and this can be done in the efficacy clinical trials. Although these data are useful, it is not always an acceptable substitute for a specific elderly volunteer PK study because the information is only available after many patients have been exposed rather than before and clinical investigators may be reluctant to enroll patients without such data in advance. In practice, for a drug likely to be given to the elderly, an elderly volunteer study should be performed soon after a young healthy volunteer study to expand the potential population for efficacy studies as much as possible. If elderly patient are then included in the main efficacy/safety studies, the population approach can then be used to explore the PK variability in this subset of the population and whether this is associated with an altered clinical outcome.

A similar rationale can be used to decide whether special kinetic (and possibly dynamic) studies should be performed in patients with renal or hepatic disease. For example, if the compound is largely metabolized to inactive metabolites, renal function can reasonably be expected not to have a major effect on kinetics. However, regulators usually will want some information as some expectations do exist to the previous assumption. An example is the “futile cycle” involving some NSAIDs—where prolonged residence of inactive acyl glucuronide metabolites in the plasma in patients with renal disease allows breakdown back to the parent molecule resulting in accumulation (Sallustio et al., 1989). As outlined earlier, a population approach could be used to screen for an effect of disease on drug kinetics, but some investigators may need reassurance before enrolling

patients in trials. A small study in patients with advanced renal disease may be able to provide this reassurance. Liver disease can be handled similarly for drugs which are primarily eliminated renally.

The safety of marketed drugs could be significantly improved if the subject groups involved in phase II and III trials better reflected the patient populations that will use drugs. By excluding “representatives” for what will clearly be subpopulations utilizing a drug (those using other drugs or with other diseases), many clear safety questions go unasked. The same is, to some degree, true about preclinical animal studies where only healthy young animals are employed rather than (perhaps) some disease model groups which might serve as better predictors of patient safety concerns.

28.4 INSTITUTIONAL REVIEW BOARDS (IRBS)/ETHICS COMMITTEES IN THE CLINICAL TRIAL PROCESS

Clinical drug trials represent research with human subjects (Cato et al., 2002). All research involving human subjects that is supported by the federal government or the results of which are to be used in applications for drug or device approval must be conducted in accordance with regulations promulgated by the HHS (45 CFR 46) and the FDA (21 CFR 56). The regulations of both the HHS and the FDA require that an IRB “...shall review and have authority to approve, required modifications in (to secure approval), or disapprove all research activities covered by [the] regulations” (45 CFR 46, 21 CFR 56).

The review of clinical drug trials by IRBs raises a number of interesting and difficult issues. These relate to the origin and sponsor of the proposed trial, the nature of the institution the IRB serves, and the manner in which the norms for determining ethical conduct in clinical trials can be applied to specific trials.

Here the ethical principles underlying research involving human subjects, the legal authority for IRBs, and the regulatory requirements affecting the operations of IRBs are reviewed. It will then discuss the role of IRBs in reviewing clinical trials by examining how IRBs can assess the scientific design of trials, the competency of the investigator, the manner of selecting subjects for the trial, the balance of risks and benefits, informed consent, and provisions for compensating for research-related injuries.

28.4.1 Legal Authority and Responsibilities for IRBs

The legal authority for IRBs derives from two parallel sets of federal regulations. One set of regulations was promulgated by the HHS and implements the 1974 amendments to the Public Health Service Act (DHEW, 1974). These regulations are codified in Title 45 of the CFR, Part 46. The second set

of regulations was promulgated by the FDA under the Federal Food, Drug, and Cosmetic Act (FD&C Act). These regulations are codified in Title 21 of the CFR; regulations pertaining to IRBs are in Part 50 and those pertaining to informed consent are in Part 56.

The FDA has the legal authority to regulate clinical investigations in the United States when the investigational products move across state or national boundaries. Under the FDA regulations, review and approval by an IRB are required for any experiment that involves a test article and one or more human subjects, either patients or healthy persons, and that is subject to the requirements for prior submission to the FDA (21 CFR 50.3). Such review is also required for any experiment, the results of which are intended to be submitted later to, or held for inspection by, the FDA. While there are as of this writing (December 2015) efforts to ease these requirements in certain cases (the “Right to Try” movement) and some US states have passed law to this effect, a fundamental change remains unlikely.

The regulations of the FDA are identical or similar to those of the ICH and HHS in nearly all essential respects. Such differences as do exist reflect the different statutory authority under which the separate sets of regulations were promulgated and the difference in mission between the FDA and the NIH, the agency within HHS charged with overseeing the implementation and enforcement of the HHS regulations. The difference in mission between the FDA and the NIH is reflected in the FDA’s approach to compliance with its regulations utilizing its traditional tools of inspections and audits.

The FDA regulations specify requirements for IRB membership, function, and operation and the criteria according to which approval may be given for conducting research. Since these requirements are similar, a single committee should be established to undertake the activities required by both sets of regulations. Additionally, the FDA regulations allow a wide variety of ways in which private practitioners not affiliated with an institution can obtain necessary IRB review of their clinical research activities. The basic ethical tenets governing the actions of an IRB should include the following (Sharp, 2001):

1. Risks to subjects are minimized (i) by using procedures which are consistent with sound research design and which do not unnecessarily expose subjects to risk and (ii) whenever appropriate by using procedures already being performed on the subjects for diagnostic or treatment purposes.
2. Risks to subjects are reasonable in relation to anticipated benefits, if any, to subjects, and the importance of the knowledge that may reasonably be expected to result....
3. Selection of subjects is equitable....

4. Informed consent will be sought from each prospective subject or the subject's legally authorized representative....
5. Informed consent will be appropriately documented....

Additional requirements are that adequate provisions exist for monitoring the data collected, that adequate provisions exist to protect the privacy of subjects and maintain the confidentiality of the data, and that appropriate safeguards be included to protect the rights and welfare of subjects who are "...vulnerable to coercion or undue influence...or persons who are economically or educationally disadvantaged..." (45 CFR 46.1).

Before a trial initiates, formal review by an IRB that agrees to assume this additional function or by IRBs formed by a local or state health agency, a medical school, a medical society, a state licensing board, or a nonprofit or for-profit independent group is required. All IRBs, regardless of sponsorship, that are assuming responsibilities for reviewing and approving clinical research protocols subject to FDA authority must comply with the IRB regulations set out by the FDA.

28.4.2 Duties of IRBs

IRBs are required to review and have the authority to approve, require modifications in, or disapprove all research activities covered by the regulations (45 CFR 46.109(b)). They must require that information given to subjects as part of informed consent is in accordance with the general requirements for informed consent that are set out in the regulations. Additionally, they may require that other information be given to subjects when they judge that such information would further protect the rights and welfare of the subjects (45 CFR 46.109(c)).

IRBs must require documentation of informed consent in all studies except those specified in the regulations in which documentation may be waived. Clinical drug trials are not among the classes of studies in which documentation of informed consent may be waived.

IRBs must provide written notification to investigators and institutions of their decisions to approve, require modifications in, or disapprove proposed research activities. Decisions to disapprove a proposed research proposal must be accompanied by a statement of reasons for the decision and provide the investigator an opportunity to respond in person or in writing.

IRBs must conduct continuing reviews of research they approve at least once each year. More frequent reviews may be required if the risk of a particular research project so warrants. IRBs have the authority to suspend or terminate approval of research that is not being conducted in accordance with their requirements or that has been associated with unexpected serious harm to subjects. Such action must be accompanied by a statement of reasons for it and be

communicated to the investigator, appropriate institutional officials, and the Secretary of HHS.

The regulations require that IRBs must follow the written procedures that are set out in the assurances they have filed with HHS, review proposed research at convened meetings at which a majority of the IRB members are present, vote approval by a majority of members present at the meeting, and be responsible for reporting to the appropriate institutional official and the Secretary of HHS "...any serious or continuing noncompliance by investigators with the requirements and determination of the IRB."

Institutions that are cooperating in multi-institutional studies, such as clinical drug trials, must each review and approve the proposed studies. Such institutions may, however, use joint review, rely on the review of another qualified IRB, or utilize similar arrangements to avoid duplication of efforts (Cato et al., 2002; Edwards et al., 2010).

28.4.3 Informed Consent

Assuring that adequate provisions exist for securing informed consent is a central duty of IRBs, and that which is seemingly the most visualized when it fails (Office of the Inspector General, 2000a, b). The requirements for informed consent are specified in international guidelines and the federal regulations. These require that investigators "shall seek such consent only under circumstances that provide the prospective subject...sufficient opportunity to consider whether or not to participate and that minimize the possibility of coercion or undue influence. The information that is given to the subject...shall be in language that is understandable to [him]" (45 CFR 46.116). The regulations further stipulate that "No informed consent, whether oral or written, may include any exculpatory language through which the subject...is made to waive or appear to waive any of the subject's legal rights, or releases or appears to be release the investigator, the sponsor, the institution or its agents from liability from negligence."

The federal regulations specify the information that shall be provided to each subject:

1. A statement that the study involves research, an explanation of the purposes of the research and the expected duration of the subject's participation, a description of the procedures to be followed, and identification of any procedures which are experimental
2. A description of any reasonably foreseeable risks or discomforts to the subjects
3. A description of any benefits to the subject or to others which may reasonably be expected from research
4. A disclosure of appropriate alternative procedures or courses of treatment, if any, that might be advantageous to the subject

5. A statement describing the extent, if any, to which confidentiality of records identifying the subject will be maintained
6. For research involving more than minimal risk, an explanation as to whether any compensation and an explanation as to whether medical treatments are available if injury occurs and, if so, what they consist of, or where further information may be obtained
7. An explanation of whom to contact for answers to pertinent questions about the research and research subject's rights and whom to contact in the event of a research-related injury to the subject
8. A statement that participation is voluntary, refusal to participate will involve no penalty or loss of benefits to which the subject is otherwise entitled, and the subject may discontinue participation at any time without penalty or loss of benefits to which the subject is otherwise entitled

In addition to these basic elements of informed consent, IRBs shall also require that information shall be provided, where indicated, to the effect that (i) the particular treatment or procedure being tested may involve risks to the subject that are currently unforeseeable, (ii) foreseeable circumstances may exist under which continued participation by the subject may be terminated by the investigator without regard to the subject's consent, (iii) additional costs to the subject may result from participation in the research, (iv) the consequences of a decision to withdraw, and (v) significant findings that may influence a subject's continued participation will be related to the subject.

In addition to the elements enumerated in the federal regulations, IRBs must consider whether consent forms should include the fact of randomization in the case of prospective randomized clinical trials.

Those who feel that the fact of randomization need not be disclosed to prospective subjects argue that since the alternative treatments to be tested are not known to produce significantly different results and since the physician would have to make an arbitrary selection of one treatment or the other for a particular patient, notification that selection of treatment is by computer rather than by the patient's own physician does not provide additional protection for the subjects and is unnecessary. The response to this contention is that a subject's ability to exercise full autonomy over what will be done with his or her own body is best served by notifying the subject as to how the treatment will be selected and by whom, even if the selection process is equally arbitrary whatever process is used.

The weight of the arguments favors the notion that for consent to be fully informed, subjects must be notified that their treatments will be allocated in a random manner, that is, selected by a process other than the judgment of their own

physician. The meaning of the concept of randomization and the fact that it will be the manner by which treatment is selected is therefore considered to be an important and integral part of informed consent for participation in randomized clinical trials.

Implicit in the elements that comprise informed consent for subjects participating in clinical trials is that subjects will be notified of the nature of their disease. Current bioethical thinking views this to be essential in order for patients/subjects to give legally effective informed consent. The current practice in the United States is that informed consent to participate in clinical trials requires that patients be notified of their diagnosis. Accordingly, a statement regarding the diagnosis is required in consent forms for participation in clinical trials that are sponsored by national cooperative groups. The Tuskegee study on syphilis is an excellent example of the results of not informing patients (Jones, 1993). It is of interest that other Western countries do not feel that it is necessary or even appropriate to inform patients of their diagnosis as part of the consent process.

Increased incidents in clinical trials have led to recognition of the weaknesses of informed consent procedures (Bohazchuk, 1998). Actually, the FDA has reported that such deficiencies are the poorest area of GCP compliance for more than 12 years (FDC Reports, 2000).

28.5 DRUG FORMULATIONS AND EXCIPIENTS

It should never be lost sight that one of the major reasons for the 1938 FD&C Act was a public health disaster caused by a drug formulation mistake. In the 1930s, the Massengill Company's use of diethylene glycol in an elixir of sulfanilamide led to 105 deaths. This same disaster was, by the way, repeated in Haiti in 1995 and 1996 (O'Brien et al., 1998). Such considerations are also overlooked in clinical safety evaluations, though the history of them directly and indirectly causing problems, even to the current day, is extensive (Winek, 2000).

Preclinical animal studies are usually performed with simple formulations which are appropriate for the route investigated in the (nonhuman) species involved. While similar simple formulations or approaches (such as capsules) are also employed for first-in-man studies, as development proceeds, efforts are made to develop formulations which optimize bioavailability. This may lead to effects not seen in earlier animal (or, indeed, human) studies—a factor that should be kept in mind in both study design and interpretation.

It is essential that formulations used in clinical trials should be well characterized, including information on bioavailability wherever feasible. The formulation should be appropriate for the stage of drug development. Ideally, the supply of a formulation will be adequate to allow testing in a series of studies that examine a range of doses. During drug

development, different formulations of a drug may be tested. Links between formulations, established by bioequivalence studies or other means, are important in interpreting clinical study results across the development program.

Safety limitations on formulations usually arise from sterility, purity (particularly acceptable levels of endotoxins), and local tissue tolerance concerns at the site of administration for drugs other than oral.

28.5.1 Route of Administration

Just as new drugs must be tested in animals by the route to be used in man, so must they be tested in volunteers using the intended route for patients. But there are clear benefits in testing all drugs when going into man for the first time using IV infusions, even if systemic exposure in patients will be achieved by another route. These benefits relate primarily to the fact that IV infusion allows for precise control of drug administration:

- In the event of a serious or otherwise distressing AE during the infusion, drug delivery can be halted.
- As the drug is delivered directly into the bloodstream, this ensures 100% exposure and overcomes problems relating to bioavailability which may occur with other routes, but in particular dosing by the oral route when the drug may be destroyed in the GI tract or metabolized presystemically in the gut wall or in the liver.
- Delivery of the full dose into the bloodstream, coupled with a uniform delivery rate, results in less variability in plasma or tissue concentrations of drug than is possible using oral dosing, where not only the extent but also the rate of absorption from the GI tract can vary considerably between subjects. Less intersubject variability in plasma concentrations of drug in turn enables the study to be done using smaller numbers of subjects and also offers advantages for drugs anticipated to have narrow therapeutic ratios.
- IV dosing allows the true disposition kinetics of the drug to be evaluated and makes the assessment of PK/PD relationships easier to perform.
- PK scaling between species, that is, animals to man, is made simpler as fewer assumptions need to be made about extent and rate of exposure in man. This in turn helps in further dose selection for human studies.
- Blinding of studies is made easier when IV dosing is used, that is, there is no need to produce matching placebos, while IV dosing overcomes any problems relating to taste, which can make it difficult in blind studies involving oral dosing.

The primary disadvantage of using IV dosing for first-time-in-human studies is that additional resource will be

needed to be spent in toxicology, establishing dosage form stability, and for mutation development on a drug which might fail at the first hurdle in man, as indeed many do. For this reason investigators often prefer to administer drugs for the first time in healthy volunteers using the route to be used in patients and dose intravenously to establish the drugs' PK profile only when they feel reasonably certain that it is likely to be a candidate for further development.

28.6 PHASE I DESIGNS

Phase I clinical trials are the first studies in which a new drug is administered to human subjects. The primary purpose of phase I studies of new drugs is to establish a safe dose and schedule of administration (O'Grady and Linet, 1990; Gad, 2009). Other purposes are to determine the types of side effects and toxicity and organ systems involved, to assess evidence for efficacy, and to investigate basic clinical pharmacology of the drug. Not all of these goals can be met completely in any phase I trial, in part because the number of patients treated is small. However, well-conducted phase I studies can achieve substantial progress toward each of these goals. Phase I trials are not synonymous with dose-response studies, but they have many characteristics in common.

The initial phase of clinical testing has the following objectives:

- a. Establish a dose-response PD profile by using initial doses projected to be therapeutic in humans. The dose required is predicted on the basis of blood levels found in animal screens.
- b. Determine the PK profile for initial titration and maintenance of steady state for chronically administered drugs.
- c. Design a safe dosage regimen for efficacy testing in adults, pediatric, or elderly.
- d. Estimate efficacy information necessary to make sample size determinations for phase II studies and establish adequate duration of treatment.
- e. Determine the drug interaction potential when concurrent medications are administered, as well as food interaction, assess the enzyme induction potential, and assess the need for therapeutic drug monitoring during efficacy testing.
- f. Establish the requirements for the final formulation.

The initial strategy for phase I is to conduct a single-dose safety study in normal volunteers. The first trial demands close 24 h supervision in a clinical setting. Ethical considerations may, however, demand that only patients be used—for example, when evaluating an anticancer agent with predictable toxicity. A repeat-dose tolerance and PK study in normal or patient volunteers is then conducted for chronically

administered drugs. These studies will provide the necessary safety information to support efficacy testing.

Sometimes investigators say that phase I studies are not “clinical trials” because there is no treatment comparison being made (except that frequently a placebo is employed). Such treatment comparisons are not a prerequisite for experiments. Because phase I trials rely on investigator-controlled treatment administration and subsequent structured observations, they are clinical trials.

In the development of cytotoxic drugs in oncology, dose finding usually means establishing a “maximum tolerated dose” (MTD). This is the dose associated with serious but reversible side effects in a sizable proportion of patients and the one that offers the best chance for a favorable therapeutic ratio. Side effects from cytotoxic drugs tend to be serious and are referred to as toxicities. Investigators are interested not only in the organ systems involved but also in the duration, reversibility, and probability of specific toxicities. In this setting, evidence of efficacy is usually weak or nonexistent, because many patients receive what turn out to be subtherapeutic doses of the drug.

For all phase I studies, learning about basic pharmacokinetics (clinical pharmacology) is important and includes measuring drug uptake, metabolism, distribution, and elimination. This information is vital to the future development and use of the drug and is helpful in determining the relationship between blood levels and side effects, if any. These goals indicate that the major areas of concern in designing phase I trials will be selection of patients, choosing a starting dose, rules for escalating doses, and methods for determining the MTD or safe dose.

If basic pharmacology were the only goal of a phase I study, the patients might be selected from any underlying disease and without regard to functioning of specific organ systems. However, phase I studies are usually targeted for patients with the specific condition under investigation. For example, in phase I cancer trials, patients are selected from those with a disease type targeted by the new drug. Because the potential risks and benefits of the drug are unknown, patients often are those with relatively advanced disease. It is usually helpful to enroll patients with a normal cardiac, hepatic, and renal function. Because bone marrow suppression is a common side effect of cytotoxic drugs, it is usually helpful to have normal hematologic function as well when testing new drugs in cancer patients. In settings other than cancer, the first patients to receive a particular drug might have less extensive disease or even be healthy volunteers.

28.6.1 First Administration: Single Dose Escalating (SDE)

First-time administration of single doses of new drugs is undertaken using a wide range of study design, but essentially there are several basic designs available which are

modified to meet the needs of a given study. Fundamental to all designs is that in the interests of safety, successive subjects are exposed to increasing doses of the drug. The fact that doses are titrated upward either in the same subject or groups of subjects and not randomized to remove the potential for bias can be argued as a design weakness, but there is no alternative. Nevertheless, an ordered dose-response can be taken as reasonable evidence of a drug-related effect. In addition, the use of placebo, which enables studies to be conducted on either a single- or a double-blind basis, will help to minimize bias. For this reason, placebo control is an integral part of a phase I study. Unwanted feelings or sensations are common occurrences in everyday life; hence it is to be expected that AEs will be encountered during phase I studies. AEs may be drug-related, study-related, or result from something which has nothing to do with the drug or the study. They may act singly or in combination. For example, headache, which is one of the most common if not the most common symptom reported by volunteers taking part in phase I studies, can result from any one of the following: fasting, caffeine withdrawal, feeling anxious about the study, an impending attack of influenza, or a combination of all four factors. Thus without placebo control it becomes difficult to differentiate between headache which is drug related and headache which is nondrug related. A placebo is not only of value in helping to distinguish between drug- and nondrug-related subjective effects, but it also plays a role in the interpretation of results from laboratory and other safety tests and pharmacological tests which may be influenced by diverse factors such as diet, physical activity, mental state, circadian or other biological rhythms, and asymptomatic illnesses, for example, subclinical viral infections.

The different designs available for a first-time-in-human study all have their own advantages and disadvantages (Fleiss, 1986; Spilker, 1991; Nylen, 2001; Gallin and Ognibene, 2012). At the end of the day, it is down to the investigator to weigh up the pros and cons of each and then to choose the design which best meets the aims of the study. In an attempt to examine their strengths and weaknesses, let us consider some designs open to an investigator who wishes to undertake a single rising dose safety and tolerability study with a new drug. A typical protocol might require:

- Placebo control
- The dose be increased from x (first dose) to $64x$ (top dose)
- Twofold increases in successive doses (within or between subjects)
- A 7-day within subject washout period
- A minimum of four subjects to receive each dose level

Some design options are shown in Tables 28.12, 28.13, 28.14, 28.15, and 28.16, while the implications for going with one or other in terms of subject numbers, number of clinic visits, highest first dose given to a subject, biggest

TABLE 28.12 Phase I Study—Type A

| Group | Number of Subjects Who Received Each Treatment | | | | | | |
|-------|--|----------|------------|------------|------------|-------------|-------------|
| | Placebo | <i>x</i> | 2 <i>x</i> | 4 <i>x</i> | 8 <i>x</i> | 16 <i>x</i> | 32 <i>x</i> |
| 1 | 2 | 3 | | | | | |
| 2 | 2 | | 3 | | | | |
| 3 | 2 | | | 3 | | | |
| 4 | 2 | | | | 3 | | |
| 5 | 2 | | | | | 3 | |
| 6 | 2 | | | | | | 3 |
| 7 | 2 | | | | | | |

Note: *x*, first dose.

TABLE 28.13 Phase I Study—Type B

| Group | Number of Subjects Who Received Each Treatment | | | | | | |
|-------|--|----------|------------|------------|------------|-------------|-------------|
| | Placebo | <i>x</i> | 2 <i>x</i> | 4 <i>x</i> | 8 <i>x</i> | 16 <i>x</i> | 32 <i>x</i> |
| 1 | 2 | 3 | 1 | | | | |
| 2 | 2 | | 2 | 1 | | | |
| 3 | 2 | | | 2 | 1 | | |
| 4 | 2 | | | | 2 | 1 | |
| 5 | 2 | | | | | 2 | 1 |
| 6 | 2 | | | | | | 2 |
| 7 | 2 | | | | | | |

Note: *x*, first dose.

TABLE 28.14 Phase I Study—Type C

| No. of Visits ^a | Volunteer No. | Treatment | | | |
|----------------------------|---------------|-------------|-------------|-------------|-----|
| 2 | 1–4 | <i>x</i> | 2 <i>x</i> | 4 <i>x</i> | (P) |
| 2 | 5–8 | 4 <i>x</i> | 8 <i>x</i> | 16 <i>x</i> | (P) |
| 2 | 9–12 | 16 <i>x</i> | 32 <i>x</i> | 64 <i>x</i> | (P) |

Note: *x*, first dose; (P), randomized placebo.

^aVolunteers receive three doses of drug on one visit and placebo on the other visit.

TABLE 28.15 Phase I Study—Type D

| No. Visits ^a | Volunteer No. | Treatment | | | |
|-------------------------|---------------|-------------|-------------|-------------|-----|
| 4 | 1–4 | <i>x</i> | 2 <i>x</i> | 4 <i>x</i> | (P) |
| 4 | 5–8 | 4 <i>x</i> | 8 <i>x</i> | 16 <i>x</i> | (P) |
| 4 | 9–12 | 16 <i>x</i> | 32 <i>x</i> | 64 <i>x</i> | (P) |

Note: *x*, first dose; (P), randomized placebo.

^aVolunteers receive a single dose of drug on each of three visits and placebo on one visit.

TABLE 28.16 Phase I Study—Type E

| Visit | Volunteer No. | | Volunteer No. | | Volunteer No. | |
|-------|---------------|-------------|---------------|-------------|---------------|-------------|
| 1 | 1 | P | 5 | <i>x</i> | 9 | 2 <i>x</i> |
| | 2 | <i>x</i> | 6 | P | 10 | 2 <i>x</i> |
| | 3 | <i>x</i> | 7 | 2 <i>x</i> | 11 | P |
| | 4 | <i>x</i> | 8 | 2 <i>x</i> | 12 | 4 <i>x</i> |
| 2 | 1 | 4 <i>x</i> | 5 | P | 9 | 4 <i>x</i> |
| | 2 | 4 <i>x</i> | 6 | 4 <i>x</i> | 10 | P |
| | 3 | 4 <i>x</i> | 7 | 4 <i>x</i> | 11 | 8 <i>x</i> |
| | 4 | P | 8 | 4 <i>x</i> | 12 | 8 <i>x</i> |
| 3 | 1 | 8 <i>x</i> | 5 | 16 <i>x</i> | 9 | P |
| | 2 | 8 <i>x</i> | 6 | 16 <i>x</i> | 10 | 16 <i>x</i> |
| | 3 | P | 7 | 16 <i>x</i> | 11 | 16 <i>x</i> |
| | 4 | 16 <i>x</i> | 8 | P | 12 | 16 <i>x</i> |
| 4 | 1 | 16 <i>x</i> | 5 | 32 <i>x</i> | 9 | 64 <i>x</i> |
| | 2 | P | 6 | 32 <i>x</i> | 10 | 64 <i>x</i> |
| | 3 | 32 <i>x</i> | 7 | P | 11 | 64 <i>x</i> |
| | 4 | 32 <i>x</i> | 8 | 64 <i>x</i> | 12 | P |

Note: *x*, first dose; P, placebo.

increment in dose, and time to complete the study are given in Table 28.17.

Design types A and B not only require more than three-fold the number of volunteers needed for the other designs but also give the clearest picture of the pharmacokinetics and tolerance of a single dose. This requirement is compounded by the fact that one-third of the volunteers will receive placebo. The number of volunteers on placebo in each group is open to the investigator's choice, but a balanced (even number of each) design again is easiest to interpret. Whichever way, types A and B require large numbers of subjects, which could present problems when recruiting suitable subjects. The situation is made more difficult when numbers on the volunteer panel are limited (which often is the case) and when one is attempting to recruit the best available volunteers who also satisfy the inclusion/exclusion criteria for a first-time-in-human study. It can also be argued that if the drug under test proved to be toxic, then more subjects would be exposed to its harmful effects. On the other hand, if the drug turns out to be well tolerated, it can be argued equally well that exposing a larger number of subjects is a better basis on which to proceed to the next study.

Types A and B, however, have two clear advantages over the other designs. First, as only one visit to the clinic is required, this will encourage the volunteer to take part in and complete the study. Second, they are ideal designs for drugs with long (or unclear) pharmacological, clinical, or chemical half-lives when a 7-day washout period is an inadequate time for the drug effects to disappear or for it to be cleared from the body.

In the interests of safety, the lower the dose the volunteer is given on the first exposure to the drug, the better. However, as it is impractical to start everyone off with dose *x*, the next

TABLE 28.17 Comparisons of Study Types

| | A | B | C | D | E |
|--|---------|---------|---------|------------------------|------------------------|
| No. of volunteers | 12 | 12 | 12 | 42 | 42 |
| No. of clinic visits | 4 | 4 | 2 | 1 | 1 |
| Highest first dose to a subject | 16x | 8x | 16x | 64x | 64x |
| Largest within subject increment in dose | ×2 | ×16 | ×2 | — | — |
| Times to do study | 9 weeks | 4 weeks | 3 weeks | 3–6 weeks ^a | 3–6 weeks ^a |

^a Dependent upon whether dosing takes place once or twice weekly.

best thing one can do is to keep the first dose given to a volunteer in each group as low as possible within the confines of the design of the study. In this respect, type E works best and types A and B do badly.

With types D and C, twofold increments in dose are uniformly made throughout the whole dose range. This is in contrast to type E, in which the size of dosage increments over the dose range within subjects varies between 2- and 16-fold. Thus type E might be an unwise choice for a drug anticipated to have a narrow therapeutic index or a steep dose–response.

Assuming the study goes according to plan (which is often not the case in first-administration studies) and depending upon the study design used, it will take between 3 and 9 weeks to complete. However, although type C (in which the dose is increased stepwise on the same study day) offers the advantage of speed, the fact that it can only really be used for drugs given by the IV route and for drugs with rapid onsets and offsets of action limits its usefulness in practice.

28.6.2 First Administration in Humans: Multiple Dose Escalating (MDE)

In clinical practice, drugs are often prescribed for illnesses which require regular treatment for days, weeks, months, or years. For drugs used in this way (or for which off-label use is likely to be this way), testing on a repeat-dose basis in volunteers is required to evaluate safety and tolerability before treating patients. As with first-time single-dose studies, first-time repeat-dose studies can be undertaken using different designs but with the emphasis again on safety and tolerability. The cornerstone design is a randomized, rising dose, placebo-controlled group comparative evaluation. Whichever design is used, the investigator has to decide upon an appropriate dosing schedule. The choice of a unit dose and dosing interval depends primarily upon the results from the single-dose study. To illustrate this point, if one assumes that the top dose (i.e., 64x) given in the previously described single-dose study proved to be well tolerated, then one might opt for the dosing schedule given in Table 28.18. Of course, the frequency of dosing will depend upon the PD and/or PK profile of the drug. Ideally, dosing should be continued until steady-state plasma concentrations of drug have been achieved, but this may not be practical for drugs with long

TABLE 28.18 Design and Dosing Schedule for a First Repeat-Dose Phase I Study

| Group | Day 1 | Day 2–6 | Day 7 |
|-------|-------|------------|-------|
| 1 | 8x | 8x b.i.d. | 8x |
| 2 | 16x | 16x b.i.d. | 16x |
| 3 | 32x | 32x b.i.d. | 32x |

Note: b.i.d. = twice daily.

half-lives. More often than not volunteers are dosed for 7–10 days, but in certain circumstances if toxicological clearance is available and there is a definite need to do so, volunteers may be dosed for 4 weeks. Even if the intent is to dose more than once daily (as in Table 28.18 where twice-daily (BID) dosing is required), giving single doses on the mornings of day 1 and the last day of dosing (i.e., day 7) offers certain advantages. For example, it allows for a longer period to assess tolerability before the second dose of drug is given to a volunteer who more than likely will not have been exposed to the drug previously. It also enables comparisons to be made between drug plasma concentration time profiles over 24 h and the elimination kinetics of the drug at the start and end of dosing.

In the interests of safety, doses are increased between groups sequentially, and as a rule dosing is completed in the previous group before dosing is started in the next group. However, if groups are to be dosed for more than 7–10 days or a large number of increments in dose is planned, particularly if more than one dosing frequency is under test, the investigator might choose to overlap dosing between one group and the next, thus enabling the study to be completed in a reasonable time frame. Within each group, volunteers are randomly allocated to receive drug or placebo. The size of the groups usually varies between 6 and 12 with the numbers of subjects receiving drug and placebo in a group being subject to investigator preference.

28.6.2.1 Number of Subjects In a phase I trial, a sufficient number of subjects must be included in a study if valid conclusions are to be drawn from the results. Studies in healthy volunteers and patients are inherently flawed when it comes to assessing safety and tolerability because of the small numbers of subjects involved, and only the most guarded of conclusions are possible. It is easier to draw valid

conclusions in respect to drug action involving PD, surrogate, or clinical end points because one is able to specify beforehand the magnitude of the difference which constitutes a useful drug effect and thus calculate the numbers of studies, except that regulatory authorities rather than the investigator specify the criteria which have to be met to enable different formulations to be judged bioequivalent.

28.7 CLINICAL TRIAL SAFETY INDICATORS

One major purpose of preclinical (animal) toxicity studies of a potential new drug is to identify the toxic effects which most commonly occur at doses nearest to those to be used in man. These observations serve to help ensure that care is taken to detect any such effects in humans. Additionally, a broad range of other indicators of adverse drug action may be identified to ensure that their occurrence is looked for. These are also commonly called safety parameters.

Because of the relatively small numbers of volunteers and patients involved, only the most common of drug-related AEs are likely to be detected during early studies (O'Grady and Joubert, 1997). For example, to have a 95% chance of picking up three subjects who have experienced an adverse reaction (with no background incidence) which occurs in 1 in every 100 subjects treated with the drug, it would need to be given to 650 subjects. Matters are made worse when the AE in question also occurs in the general population, which is usually the case with the kind of symptoms reported by volunteers and patients taking part in drug studies. No matter how good the study design, nothing can compensate for this problem of inadequate numbers. In this respect, all of the study designs described earlier are more or less equally adequate or inadequate as the case may be.

Monitoring for drug-related AEs employs the same or similar methods in both volunteers and patients. In both cases assessments of tolerability and safety are based upon symptom reports, routine laboratory safety screens, EKG monitoring, and on occasions special tests designed to detect unwanted effects associated with a particular class of drug. The chances of obtaining reliable information on a drug's safety profile are enhanced by detailed and careful monitoring, such as special biomarkers for cardiotoxicity (Braunwald, 2008). Symptoms may be reported spontaneously or elicited in reply to standard questions. Open questions such "how are you feeling?" are to be preferred to leading questions on the basis that they result in fewer reports of AEs. If leading questions are used, they need to be carefully worded. A certain amount of basic information is required on all AEs, that is, type, severity, time of onset in relation to time of dosing, duration, and causality. Attributing the cause of an unwanted effect to the drug or some other factor can be difficult particularly when little is known about the drug, as is often the case at the state of initial studies in

volunteers or patients. Rechallenge with the drug ideally using the same dose or, if need be (because the event caused a degree of discomfort), a reduced dose is probably the single best way of proving or disproving a causal relationship. But if done the rechallenge procedure must be designed using placebo as comparator under double-blind conditions. Obviously rechallenges can be done only if the AE was reversible, did not cause excessive discomfort, and most importantly was not life threatening. The question of assessing attributions or causality is considered in detail later in this chapter.

28.7.1 Overall Approach to Assessing Safety

28.7.1.1 Choosing Safety Parameters Choosing the appropriate safety parameters for a clinical trial depends on a number of factors. A selected list of examinations and tests commonly used to assess the safety of medicines is given in Table 28.19. The majority of these tests will not be conducted in most drug trials. An assessment of the quantity and quality of prior experience and previous data obtained with the therapeutic is essential to enable one to decide which specific safety tests to incorporate in a medicine trials. The choice of safety parameters requires both data in areas where there are indications of potential (or actual) safety problems to monitor and also additional experience and data with a new drug. Until a sufficient body of safety data has accumulated, more laboratory parameters of safety are generally included than will be needed at a later date. The nature of the clinical trials and efficacy tests used may dictate that certain safety parameters should or should not be included (e.g., in testing a new anticancer medicine, it may be necessary to perform a bone marrow biopsy and smear to confirm the lack of toxicity, and in assessing an agent in anesthetized patients, the appropriate tests to ensure the patient's safety while under anesthesia must be performed). If, on the basis of preclinical pharmacological or toxicological data, any toxicity is either anticipated or considered possible, then an attempt should be made to evaluate patients for those possible problems. The anticipated use(s) of a therapeutic will also influence which safety parameters are chosen for evaluation (e.g., ophthalmological tests would be included for drugs intended for ocular use).

28.7.1.2 Measuring Safety Parameters After specific safety parameters are chosen, it is necessary to determine how thorough an evaluation of each parameter should be conducted. It is also possible that different types of examinations would be suitable at different points of a clinical trial. For example, a physical examination may be specified to include more or fewer measurements or facets, and a complete examination may not be necessary or even suitable during some periods of clinical trial.

TABLE 28.19 Selected List of Examinations and Tests Used to Evaluate Safety

| |
|---|
| A. Clinical examinations |
| 1. Physical |
| 2. Vital signs (usually considered as part of the physical examination) |
| 3. Height and weight (state of dress is usually specified, e.g., socks) |
| 4. Neurological or other specialized clinical examinations |
| B. Clinical laboratory examinations |
| 1. Hematology (see Table 28.23) |
| 2. Clinical chemistry (see Table 28.23) |
| 3. Urinalysis (see Table 28.23) |
| 4. Virology (viral cultures or viral serology) |
| 5. Immunology or immunochemistry (e.g., immunoglobulins, complement) |
| 6. Serology |
| 7. Microbiology (including bacteriology and mycology) |
| 8. Parasitology (e.g., stool for ova and protozoa) |
| 9. Pulmonary function tests (e.g., arterial blood gas) |
| 10. Other biological tests (e.g., endocrine, toxicology screen) |
| 11. Stool for occult blood (specify hemoccult or guaiac method) |
| 12. Skin tests for immunologic competence |
| 13. Medicine screen (usually in urine) for detection of illegal or nonprotocol-approved medicines |
| 14. Bone marrow examination |
| 15. Gonadal function (e.g., sperm count, sperm motility) |
| 16. Genetics studies (e.g., evaluate chromosomal integrity) |
| 17. Stool analysis using <i>in vivo</i> dialysis |
| C. Probe for adverse reactions |
| D. Psychological and psychiatric tests and examinations |
| 1. Psychometric and performance examinations |
| 2. Behavioral rating scales |
| 3. Dependence liability |
| E. Examinations requiring specialized equipment (selected examples) |
| 1. Audiometry |
| 2. Electrocardiogram (EKG) |
| 3. Electroencephalogram (EEG) |
| 4. Electromyography (EMG) |
| 5. Stress test |
| 6. Endoscopy |
| 7. Computed tomography (CT) scans |
| 8. Ophthalmological examination |
| 9. Ultrasound |
| 10. X-rays |
| 11. Others |

Vital signs may be measured with the patient in a supine, seated, and/or erect position. Both supine and erect positions are usually used if orthostatic changes are being evaluated. The need for such data will depend on the situation, but the position of the patients for this examination, as well as the period of time desired for stabilization, should be noted in the protocol.

28.7.1.3 Parameters That Measure either Safety or Efficacy

Certain parameters may, of course, be either safety or efficacy parameters, or both. The electroencephalogram (EEG) is an example. Blood pressure is another. It is thus important to establish clearly in the protocol whether each parameter is being incorporated in the protocol for safety or efficacy evaluations. Almost any safety parameter can be used for measuring efficacy.

28.7.1.4 Appropriateness of Each Parameter for the Clinical Trial and Patient

There are four categories of appropriateness of safety tests used in clinical trials:

1. Appropriate for patients, but not necessary for the clinical trial. All of these tests should be included in the clinical trial. They indirectly benefit the trial because they may be monitored for progress or trends or they may simply ensure that patients are receiving appropriate care.
2. Appropriate for the clinical trial, but not necessary for the patients. These tests should be included in the clinical trial if they do not place the patient at unacceptable risk or discomfort. If any tests are deemed unethical in the context of the trials and the patients enrolled, then they should be excluded.
3. Appropriate for both patients and the clinical trial. All of these tests should be included in the clinical trial.
4. Appropriate for neither patients or the clinical trial. All of these tests should be identified and excluded from the clinical trial.

28.7.2 Precautions

Clinical laboratory parameters must be specified individually in the protocol. Abbreviations such as “SMA-6” or “SMA-12” are not acceptable, as different laboratories include different tests in their “SMA-6” (or “SMA-12”) battery, and using these abbreviations without an explanation can adversely affect the clarity of the protocol and possibly lead to the collection of data on divergent parameters at different sites. Other precautions to consider prior to initiating a clinical trial are to decide if (i) severely abnormal results should be routinely confirmed, (ii) samples should be divided and sent to two separate laboratories when specified abnormalities are determined, (iii) additional tests should be routinely requested if specified abnormalities are observed, (iv) medical consultants should examine patients whenever severe abnormalities are observed, and (v) aliquots of known concentrations of standard drugs should be sent to laboratories for confirmatory measurements and interlaboratory evaluation.

28.7.2.1 Summary of Tests Common dermatological tests are shown in Table 28.20 and ophthalmological tests in Table 28.21. Note that any of these tests could be utilized as

TABLE 28.20 Selected Examples of Safety Measurements and Tests for a Specialized Dermatological Examination

1. Biopsy
2. Erythema at site of lesion
3. Absorption of medications systemically (e.g., blood levels)
4. Signs and symptoms of absorption
5. Interactions with standard treatment (e.g., ultraviolet light)

TABLE 28.21 Procedures and Tests Performed in Ophthalmological Examination

1. Ophthalmological history (attention is paid to patient family history plus patient's diseases and drug reactions)
2. Visual acuity corrected (i.e., with glasses present)
3. External ocular examination (i.e., check for inflammation, ptosis, nystagmus, tearing, proptosis, and other abnormalities)
4. Extraocular muscle testing
5. Pupil size and evaluation (in darkened room with controlled illumination)
6. Slit-lamp biomicroscopy (with dilated pupils)
7. Tonometry (ocular pressure)
8. Ophthalmoscopy with fundus photographs
9. Visual field testing and color vision testing
10. Gonioscopy^a
11. Lacrimation^a (Schirmer's test)

^aThese tests are of minimal value in determining ocular toxicity and are not recommended for routine use in ophthalmological examination to detect drug toxicities.

TABLE 28.22 Selected Considerations Pertaining to Laboratory Data

1. Ask the laboratory to maintain assayed samples that are of particular importance; if questions arise as to the accuracy of results, it might be possible to retest the original samples
2. If laboratory problems are anticipated, divide the initial (and subsequent) samples and send them to two different laboratories or to the same laboratory at two different times
3. If laboratory samples for a complete blood count are going to remain unexamined for a long period of time (e.g., sample obtained on Sunday), prepare a fresh smear so that a comparison may be made with one made 24 or more hours later, because abnormalities may occur when a sample lies around even when it is kept at an appropriate temperature

measures of efficacy if they addressed the clinical trial objectives. Selected pointers are given in Table 28.22. Specific tests that may be used in hematology, clinical chemistry, and urinalysis are shown in Table 28.23, adult and pediatric behavioral rating scales in Tables 28.24 and 28.25, and psychometric and performance tests in Table 28.26.

28.7.2.2 Choosing Laboratory Tests There is no standardized series of laboratory parameters that are evaluated in all clinical trials, nor is there a single standard for drugs in phases I, II, or III. There are, however, broad general guide-

lines for laboratory tests that are performed at each stage of clinical development.

28.7.2.3 Tests in Phase I In phase I clinical trials, there is the greatest need to obtain a wide variety of laboratory evaluations as part of developing the safety profile on a new medicine. This entails an evaluation of the basic hematology, clinical chemistry, and urinalysis parameters (Table 28.23). There will never be 100% agreement among investigators and/or clinical scientists as to which specific tests constitute a "basic" workup.

28.7.2.4 Tests in Later Phases The total number of normal laboratory values that is sufficient to collect on a new drug to demonstrate safety is impossible to specify. Numerous factors must be considered, such as the toxicological profile on other safety parameters and the expected use of the drug in patients. It is important to determine if a therapeutic agent is to be used topically or parenterally, whether it is to be used in generally healthy patients or in seriously ill patients, whether it is a "me-too" drug or a totally novel drug chemically, and whether it will be life-saving or provide a minimal therapeutic effect. The number of laboratory tests performed usually decreases as an investigational drug moves closer to the market, but one or more tests may be added to the list in Table 28.23 and studied in great detail.

28.7.2.5 Tests in Medical Practice The ordering of laboratory tests in medical practice (as opposed to phase I clinical trials) is extremely inefficient and often irrational. This suggests the need in some clinical situations to develop logical protocols and algorithms for physicians to follow in ordering tests, particularly when the technology is changing (e.g., hepatitis), in therapeutic areas in which an excessive number of tests are often ordered (e.g., thyroid tests), or when hospitals have developed their own approaches to diagnosis (e.g., use of cardiac isoenzymes in diagnosing a myocardial infarction).

28.7.2.6 Less Commonly Used Methods Evaluations of virtually any biological fluid, tissue, or sense (taste, smell, hearing, sight, and touch) can be conducted to ascertain the safety of a drug (several have been reported to affect taste in some patients, and there are many other examples involving medicine-induced effects on one of the other senses). The choice of tests will depend on experience with the medicine and suspicions about possible problems. Drugs should also be reviewed for teratogenic potential, drug dependence, liability, and carcinogenicity.

28.7.2.7 Identifying the Most Important Laboratory Analytes to Monitor in a Clinical Trial A choice often must be made among the numerous laboratory analytes that could be measured in a clinical trial. This choice is based on

TABLE 28.23 Hematology, Clinical Chemistry, and Urinalysis Parameters Usually Evaluated during the Development of a New Therapeutic Agent

| | |
|---|--|
| <p>A. Hematology</p> <ol style="list-style-type: none"> 1. Red blood cell (RBC) count 2. Hemoglobin 3. Hematocrit 4. White blood cell (WBC) count and differential 5. Platelet estimate or platelet count 6. Red blood cell indices (MCV, MCH, MCHC)^a 7. Prothrombin (PT), partial thromboplastin time (PTT) 8. Reticulocytes 9. Fibrinogen 10. Any additional tests suggested by previous data <p>B. Clinical chemistry</p> <ol style="list-style-type: none"> 1. Albumin 2. Albumin/globulin ratio 3. Alkaline phosphatase (and/or its isoenzymes) 4. Amylase 5. Bilirubin, total and direct 6. Bicarbonate (carbon dioxide) 7. BUN/creatinine ratio 8. Calcium 9. Chloride 10. Cholesterol (and/or a lipid panel) 11. Creatinine 12. Creatine phosphokinase (CPK) 13. γ-Glutamyl transferase (GGT) 14. Globulin 15. Glucose, nonfasting or fasting 16. Glucose-6-phosphate dehydrogenase (G6PD) 17. Glutamic oxaloacetic transaminase (SGOT), now frequently referred to as aspartate aminotransferase (AST) 18. Glutamate pyruvate transaminase (SGPT), now frequently referred to as alanine aminotransferase (ALT) 19. Iron (and/or other related parameters such as ferritin and total iron binding capacity) 20. Lactic acid dehydrogenase, total (LDH and/or its isoenzymes) 21. Inorganic phosphorus 22. Potassium | <ol style="list-style-type: none"> 23. Sodium 24. Total iron binding capacity 25. Total protein 26. Triglycerides 27. Blood urea nitrogen (BUN) 28. Uric acid <p>C. Hormones and/or other chemical substances in blood</p> <p>D. Urinalysis^b</p> <ol style="list-style-type: none"> 1. Appearance and color 2. Specific gravity 3. Acetone 4. Protein 5. Glucose 6. PH 7. Bile 8. Urobilinogen 9. Occult blood 10. Microscopic evaluation of sediment <ol style="list-style-type: none"> a. Red blood cells (number per high-power field) b. White blood cells (number per high-power field) c. Casts (describe and give number per high- or low-power field) d. Crystals (describe and given number per high-power field) e. Bacteria (generally rated as few, many, or loaded) f. Epithelial cells (number per low-power field) <p>E. Other urine tests sometimes evaluated</p> <ol style="list-style-type: none"> 1. Creatinine (actual values are preferable to estimated values) 2. Electrolytes (usually sodium, potassium, and chloride) 3. Protein 4. Specific hormones or chemicals 5. 24 h collections for specific evaluations |
|---|--|

^aMCH, mean corpuscular hemoglobin=hemoglobin divided by RBC count; MCHC, mean corpuscular hemoglobin concentration=hemoglobin divided by hematocrit; MCV, mean corpuscular volume=hematocrit divided by RCB count.

^bSample codes used to quantify several parameters in the urinalysis are the following: protein, glucose, ketones, bilirubin: 0, none or negative; 0.5, trace or positive (qualitative); 1, + or 1+; 2, ++ or 2+; 3, +++ or 3+; 4, ++++ or 4+. Epithelial cells, crystal, WBC, RBC, casts: 0, none or negative; 0.5, rare, occasional, few present, trace (1–5); 1, several, mild (6–10); 2, moderate (11–25); 3, many, much (26–50); 4, loaded, severe (>50). Bacteria: 0, none or negative; 0.5, rare, trace, occasional, few several (1–10); 1, mild (11–50); 2, moderate (51–75); 3, many, numerous (76–100); 4, loaded, severe (>100).

(i) past experience with the treatment(s) being evaluated, (ii) therapeutic claim, (iii) cost of the tests, (iv) convenience of obtaining samples, (v) resources available, (vi) state-of-the-art concept of the data's importance, and (vii) the ability of data obtained to convince both regulators and medical practitioners. To arrive at a decision given these and other previously discussed factors may be difficult.

28.7.2.8 Uses of Specific Laboratory Tests to Discover, Confirm, and/or Exclude a Disease Some tests can confirm the diagnosis of a disease (e.g., tissue histology from a bronchoscopic biopsy to confirm lung cancer), but cannot be used to exclude the disease or discover the disease in routine screening. Other tests can be used both to confirm and to exclude the diagnosis of a disease (e.g., glucose tolerance

TABLE 28.24 Adult Behavioral Rating Scales^a

| Scale | Scale Rated by | |
|--|----------------|---------|
| | Professional | Subject |
| 1. Anxiety Status Inventory (ASI) | X | |
| 2. Beck Depression Inventory (Beck) | | X |
| 3. Brief Psychiatric Rating Scale (BPRS) | X | |
| 4. Carroll Depression Scale | | X |
| 5. Clinical Global Impression (CGI) | X or | X |
| 6. Clyde Mood Scale | X | |
| 7. Covi Anxiety Scale | X | |
| 8. Crichton Geriatric Rating Scale | X | |
| 9. Depression Status Inventory | X | |
| 10. Hamilton Anxiety Scale (HAMA) | X | |
| 11. Hamilton Depression Scale (HAMD) | X | |
| 12. Hopkins Symptom Checklist (HSCL) | | X |
| 13. Inpatient Multidimensional Psychiatric Scale (IMPS) | X | |
| 14. Nurses' Observation Scale for Inpatient Evaluation (NOSIE) | X | |
| 15. Plutchik Geriatric Rating Scale (PLUT) | X | |
| 16. Profile of Mood States (POMS) | | X |
| 17. Sandoz Clinical Assessment-Geriatric | X | |
| 18. Self-Report Symptom Inventory (SCL-90) | | X |
| 19. Wittenborn Psychiatric Rating Scale (WITT) | X | |
| 20. Zung Self-Rating Anxiety Scale (SAS) | | X |
| 21. Zung Self-Rating Depression Scale (SDS) | | X |

^aStandard abbreviations are used (see *ECDEU Assessment Manual for Psychopharmacology*, Guy, 1976). Additional tests are described in *Mental Measurements Yearbook* (Buros, 1978).

TABLE 28.25 Pediatric Behavioral Rating and Diagnostic Scales

| |
|--|
| 1. Children's Behavior Inventory (CBI) |
| 2. Children's Diagnostic Classification (CDC) |
| 3. Children's Diagnostic Scale (CDS) |
| 4. Children's Psychiatric Rating Scale (CPRS) |
| 5. Clinical Global Impression (CGI) |
| 6. Conners Parent Questionnaire (PQ) |
| 7. Conners Parent/Teacher Questionnaire (PTQ) |
| 8. Conners Teacher Questionnaire (TQ) |
| 9. Devereux Child Behavior Rating Scale |
| 10. Devereux Elementary School Behavior Rating Scale |
| 11. Dosage Record and Treatment Emergent Symptom Scale (DOTES) |
| 12. Stereotyped Behavior in Retarded |

test for diabetes mellitus) but are too inconvenient to be used to discover the disease in routine screening. The uses of each laboratory test to discover, confirm, or exclude a disease should be considered before a test is simply added to a

TABLE 28.26 Psychometric and Performance Tests^a

| Test | For Use in | |
|---|------------|----------|
| | Adults | Children |
| 1. Bender-Gestalt Test | X | X |
| 2. Conceptual Clustering Memory Test | X | X |
| 3. Digit Symbol Substitution Test | X | X |
| 4. Embedded Figures Test | X | X |
| 5. Frostig Development Test of Visual Perception | | X |
| 6. Goodenough-Harris Figure-Drawing Test (GOOD) | | X |
| 7. Peabody Picture Vocabulary Test | | X |
| 8. Porteus Mazes | X | X |
| 9. Reaction time | X | X |
| 10. Vigilance tests | X | X |
| 11. Wechsler Adult Intelligence Scale (WAIS) | X | |
| 12. Wechsler Intelligence Scale for Children (WISC) | | X |
| 13. Wechsler Memory (WMEM) Scale | X | X |
| 14. Wide Range Achievement Test (WRAT) | | X |

^aAdditional tests are described in *Mental Measurements Yearbook* (Buros, 1978).

clinical trial protocol. This ensures that the test is appropriate in the context of the planned clinical trial.

28.7.2.9 Hematology A basic hematology evaluation usually includes determination of hemoglobin, hematocrit, red blood cell (RBC) count, white blood cell (WBC) count, RBC indices (mean corpuscular hemoglobin (MCH), mean corpuscular hemoglobin concentration (MCHC), and mean corpuscular volume (MCV)), and a platelet count. The WBC differential count is usually not required as part of a basic hematological workup unless a specific parameter of the differential count is being evaluated. Nonetheless, a WBC differential count is often obtained in phase I and generally provides useful (though often negative) information. Other hematological parameters (some of which are indicated in Table 28.23) are not usually obtained unless there is a specific reason to do so.

28.7.3 Clinical Chemistry

A measurement of renal function (creatinine and/or BUN) is an "essential" test for most clinical studies, as is the inclusion of a panel of liver function tests (SGOT, SGPT, LDH, CPK, GGT, and/or alkaline phosphatase). The specific tests chosen to be included in a study are somewhat dependent on both the investigator's and/or clinical scientist's experiences and the characteristics of the drug. Other important parameters to measure include serum electrolytes and at least some of the tests listed in Table 28.23.

28.7.3.1 Drug Levels in Plasma Drug levels may also be measured in a clinical trial. Such levels are usually part of a PK analysis but also provide important safety data. This information would be particularly relevant in cases of suspected or actual drug overdosage, drug interactions, to correlate medicine levels with toxic events, or in other situations. It must be clarified whether free levels of the drug and/or the protein bound will be measured by the laboratory.

As discussed before, plasma levels are a surrogate for drug levels in the body and may be misleading.

28.7.3.2 Total Blood That May Be Taken from Patients

The total amount of blood that may be taken from a subject in most therapeutic trials should be limited to one unit (about 460 mL) per 8-week period.

28.7.4 Urinalysis

Most clinical laboratories have established a standard battery of tests that includes most or all of the basic parameters listed in Table 28.23. If a dipstick is used to test the urine for several parameters, it is useful to use one that measures occult blood, even if a microscopic examination will count the number of RBC per high-power field. The means of obtaining the specimen should be indicated (i.e., normal voiding sample, clean catch, midstream, catheterization, suprapubic tap, or cystoscopy), especially in clinical trials in which an antidiuretic or antibiotic (or other relevant drug) is being tested.

It is usually unnecessary to obtain a microscopic examination on all urinalyses unless there are reasons to believe that important information and data may be lost. This is particularly true after it has been demonstrated that the test treatment does not affect the parameters measured in the microscopic evaluation of urine.

28.7.5 Urine Screens

A urine screen can be used to confirm generally that patients being screened or entering the baseline period of the clinical trial are not using agents (legal or otherwise) contraindicated in the protocol. It can also be used on a scheduled or random basis during the study to confirm that patients are not using such agents. The urine screen is limited in that it is unable to detect positive compliance with the protocol and only measures certain aspects of compliance failure. If a urine tests will be conducted at unannounced times in the clinical trial, then this point must be mentioned in the informed consent.

The number of agents tested in the urine screen is generally determined individually for each clinical trial, since there is a wide variety of possible drugs that may be measured. The choice of drugs to screen will be based on their relative importance for the trial plus the cost and reliability of the methodology. Results of urine screens are usually best viewed in qualitative (i.e., present or absent) rather than quantitative

terms. The identification of specific drugs in a patient's urine may help in explaining unusual adverse reactions, laboratory abnormalities, or other events. Urine screens may detect the presence of the therapeutic under study. If the urine screen is able to detect the presence of the study drug and this is reported as an unknown drug that is present or as a false positive for another drug, then it could essentially unblind a double-blind clinical trial. To prevent this situation from occurring, data from urine screens may be reported to a nonblinded monitor rather than to the investigator. If a sample of the study drug is put in urine at a physiological concentration and sent to the laboratory, the possibility of cross-reactivity with known agents may be assessed prior to initiation of the trial.

28.7.5.1 Type of Container to Be Used The specific type of container used to collect blood or urine samples is sometimes indicated in a protocol, especially if a special anticoagulant or additive is required or if other specific conditions of sample collection and handling are required. It is generally not necessary to provide this information for commonly requested laboratory tests.

28.7.5.2 Use of International System Units Although the international system of laboratory analyte units is almost universally agreed upon, many people in the United States resist using it. Typically, these are physicians (and others) who desire to retain the system with which they were trained, which makes more sense to them.

28.7.6 Identifying New Diagnostic Laboratory Tests

Numerous laboratory tests are periodically performed as aids in the diagnosis of disease states. The standards that must be met before a new test is accepted are extremely high, particularly in terms of calculated rates of false-positive and false-negative results. A five-step process leading up to acceptance of a new diagnostic test is proposed.

28.7.7 Ophthalmological Examination

Various parts of the ophthalmological examination are shown in Table 28.21. The most important common ophthalmological test to evaluate patients for the occurrence of chronic drug-induced toxicity is slit-lamp examination. Specific types of drugs with known potential for ocular toxicity may require that special attention be directed to other evaluations shown in Table 28.21. Most drugs that are to be taken systemically require at least some evaluation of ocular safety prior to approval for marketing.

28.7.8 Dermatological Examinations

A few selected safety measurements and tests for specialized dermatological examination are listed in Table 28.20.

In evaluating the safety of drugs using laboratory or other tests, it is important to develop data that helps establish the nature and magnitude of any issue or problem (real or potential) that arises with abnormal laboratory data. Data obtained must also measure the strength of the association between the drug and the event noted or of the serial trends that are observed. While this information is being collected, the definitive courses of action in dealing with the issue or problem can be developed and evaluated. These countermeasures may take the form of (i) periodic monitoring (i.e., prothrombin (PT) or partial thromboplastin time (PTT) times for patients receiving anticoagulants), (ii) cessation of medicine treatment, (iii) decreasing the dose or changing the dose schedule, (iv) initiating countertreatment, (v) use of specific antidotes to counter or reverse medicine effects, (vi) increasing surveillance of the patient, or (vii) various other alternatives.

28.7.9 Cardiovascular Safety

As cardiovascular effects can be fatal or extremely debilitating, irreversible, and rapidly occurring, a number of steps are taken to allow prompt detection and evaluation of such effects. Functionally, blood pressure and heart rate and cardiac electrophysiology (either episodically by EKG or continuously even when mobile by a Holter monitor) can now be supplemented by measuring a range of biomarkers (Table 28.27). Additionally, indicators of inflammation which are frequently associated with heart failure (Table 28.28) should be evaluated and screened for.

28.7.10 Deaths in Clinical Trials

Certain ADRs may be sufficiently alarming so as to require very rapid notification to regulators in countries where the medicinal product or indication, formulation, or population for the medicinal product are still not approved for marketing because such reports may lead to consideration of suspension of, or other limitations to, a clinical investigations program. Fatal or life-threatening unexpected ADRs occurring in clinical investigations qualify for very rapid reporting. Regulatory agencies should be notified (e.g., by telephone, by facsimile transmission, or in writing) as soon as possible, but no later than seven calendar days after first knowledge by the sponsor that a case qualifies, followed by a report that is as complete as possible within eight additional calendar days. This report must include an assessment of the importance and implication of the findings, including relevant previous experience with the same or similar medicinal products.

Determining the cause of deaths in clinical trials is extremely important, but this goal is often difficult or impossible to achieve. Investigators should be prepared to present reasons to family members to convince them of the importance

TABLE 28.27 Biomarkers in Heart Failure

| |
|---|
| Inflammation ^{a,b,c} |
| C-reactive protein |
| Tumor necrosis factor α |
| Fas (APO-1) |
| Interleukins 1, 6, and 18 |
| Oxidative stress ^{a,b,d} |
| Oxidized low-density lipoproteins |
| Myeloperoxidase |
| Urinary biopyrrins |
| Urinary and plasma isoprostanes |
| Plasma malondialdehyde |
| Extracellular matrix remodeling |
| Matrix metalloproteinases |
| Tissue inhibitors of metalloproteinases |
| Collagen propeptides |
| Propeptide procollagen type I |
| Plasma procollagen type III |
| Neurohormones ^{a,b,d} |
| Norepinephrine |
| Renin |
| Angiotensin II |
| Aldosterone |
| Arginine vasopressin |
| Endothelin |
| Myocyte injury ^{a,b,d} |
| Cardiac-specific troponins I and T |
| Myosin light-chain kinase I |
| Heart-type fatty acid protein |
| Creatine kinase MB fraction |
| Myocyte stress ^{b,c,d,e} |
| Brain natriuretic peptide |
| N-terminal probrain natriuretic peptide |
| Midregional fragment of proadrenomedullin |
| ST2 |
| New biomarkers ^b |
| Chromogranin |
| Galectin 3 |
| Osteoprotegerin |
| Adiponectin |
| Growth differentiation factor 15 |

Source: Adapted from Braunwald (2008).

^a Biomarkers in this category aid in elucidating the pathogenesis of heart failure.

^b Biomarkers in this category provide prognostic information and enhance risk stratification.

^c Biomarkers in this category can be used to identify subjects at risk for heart failure.

^d Biomarkers in this category are potential targets of therapy.

^e Biomarkers in this category are useful in the diagnosis of heart failure and in monitoring therapy.

of conducting an autopsy. Such an autopsy should include examination of the brain, whenever possible.

Any history of drug or alcohol abuse by a patient should trigger a request for appropriate blood and urine tests. Blood samples should always be taken to assess the levels of study

TABLE 28.28 Deleterious Effects of Biomarkers of Inflammation in Heart Failure**Known**

Left ventricular dysfunction
 Pulmonary edema
 Cardiomyopathy
 Decreased skeletal muscle blood flow
 Endothelial dysfunction
 Anorexia and cachexia

Potential^a

Receptor uncoupling from adenylate cyclase
 Activation of the fetal-gene program
 Apoptosis of cardiac myocytes

Source: Adapted from Hartupée and Mann (2013).

^aEffects shown in animals but not yet in humans.

drugs and any concomitant agents used. The drug containers should always be analyzed to confirm their contents. This usually entails sending these drugs to their manufacturer.

The circumstances surrounding the patients' death should be as well documented as possible, including a description of all possible influences of the clinical trial procedures on the death, even influences that are clearly independent of the medicine(s) being tested. Even procedures in a clinical trial apparently unrelated to a patient's death may have contributed to the death in some way. For example, these procedures could include (i) the requirement for excessive physical exertion, (ii) prolonged periods of psychologically difficult testing that lead to extreme fatigue, or (iii) giving patients many (e.g., 30) large capsules to ingest per day that lead to choking or aspiration.

Evaluation of the data surrounding the death by physicians who are unassociated with the clinical trial lends additional credibility to the report and conclusions. Physician biases probably will strongly influence their decision regarding the association of a patient's death with the clinical trial, and this factor must be considered in interpreting their report. This is particularly true for developing survival curves in cancer or other often fatal diseases, when deaths unrelated to the disease or to the treatment are excluded from the analysis.

28.7.11 Behavioral Rating Scales, Performance, Personality, and Disability Tests

A number of behavioral ratings scales and psychometric and performance tests, listed in Tables 28.24, 28.25, and 28.26, are briefly summarized in the following, since many of these scales and tests may be used to evaluate safety as well as efficacy. The following comments on the tests provide only a few highlights; readers who are interested in more details are advised to obtain additional information before choosing the tests that appear most relevant to be included in their particular protocol.

These scales may be used either as part of a clinical trial or as major end points in an efficacy trial. Here they are described as a means of obtaining ancillary data on psychological factors in a clinical trial. If these scales are used to demonstrate efficacy, it is mandatory to include only those scales known to be valid.

Unless otherwise noted, all of the adult and children's behavioral scales are given once pretreatment and at least once posttreatment (depending on the trial design, subject drug pharmacokinetics, and length of the trials). Investigators may schedule additional evaluations with these tests, but this is usually not done at less than weekly or biweekly intervals. Many tests provide data on both a total score and subtest (factor) scores. The times given to complete tests are subject to significant variation depending on the anxiety and characteristics of the patient and/or the experience of the professional. The times listed do not include either scoring or preliminary and/or necessary observations of the patient.

28.7.12 Adult Behavioral Rating Scales

28.7.12.1 Anxiety Status Inventory The Anxiety Status Inventory (ASI) scale is the professional-rated version of the Zung Self-Rating Anxiety Scale (SAS). Both tests (ASI and SAS) contain 20 items, each with a four-point scale, and are designed for use in adults diagnosed as having anxiety neurosis. Both assess anxiety as a clinical disorder rather than a "feeling state." The tests rate either the present time or the average status of the patient during the week preceding the evaluation. The ASI takes up to 15–20 min to complete and gives two scores: state anxiety and trait anxiety.

28.7.12.2 Beck Depression Inventory The Beck Depression Inventory (Beck) test may be used to measure the depth of depression as a rapid screen for depressed patients. It is a self-rating scale of 21 items (13 in a shortened form), with each item rated on a four-point scale. It measures the immediate present and has been used in antidepressant medicine trials. The original 21-item scale can be completed in about 10 min and the test is able to discriminate between anxiety and depression. No subtests are present in the Beck.

28.7.12.3 Brief Psychiatric Rating Scale The Brief Psychiatric Rating Scale (BPRS) is used primarily in adult inpatients to evaluate treatment response in medicine trials and in nonmedicine clinical treatment, but it is also used in some outpatient trials. Abbreviated instructions are printed on the form. Ratings are based on observations of patients. Originally developed for psychopharmacologic research, this test contains 18 symptoms, each rated on a seven-point severity scale. It requires approximately 20 min to complete and rates the period of time since the last test. If the test is being used for the first time, it rates the previous week. Five

separate subscales are obtained: anxiety/depression, anergia, thought disturbance, activation, and hostility/suspiciousness.

28.7.12.4 Carroll Rating Scale for Depression The Carroll Rating Scale for depression (52-item self-rating scale) is scored with yes or no answers by patients. It was designed to match closely the information content and specific items included in the Hamilton Rating Scale. It has been validated by comparisons with both the Hamilton Depression Scale (HAMD) and Beck and requires approximately 20 min to complete. Seventeen components of depression are measured.

28.7.12.5 Clinical Global Impressions Although the ECDEU *Assessment Manual for Psychopharmacology* (Guy, 1976) provides a formal test for the Clinical Global Impression (CGI) Scale, numerous investigators have modified the three major questions as well as the scales used in order to fit this test to their own clinical trials. The three questions, which may be applied in almost all phase II and phase III clinical trials, are:

1. *Severity of illness* “Considering your total clinical experience with this particular population, how mentally ill (the investigator may substitute a more appropriate term if this is not applicable) is the patient at this time?”
2. *Global improvement* “Rate total improvement, whether or not in your judgment it is due entirely to medicine treatment.”
3. *Efficacy index* “Rate this item on the basis of medicine effect only.” This utilizes a rating of both efficacy and adverse reactions and divides the efficacy score by the adverse reaction score to form a ratio (“efficacy index”).

Severity of illness is the only one of these three rated pretreatment. All three questions may be rated posttreatment, and additional ratings are possible during a clinical trial. The CGI measure, which is widely used in all types of medicine trials, is generally well accepted.

A scale of two to nine gradations is usually used for questions 1 and 2, although five or so gradations are probably most common. A typical five-point scale for question 2 would be that the patient is rated as 1 (much worse), 2 (minimally worse), 3 (unchanged), 4 (minimally improved), or 5 (markedly improved).

28.7.12.6 Clyde Mood Scale The Clyde Mood Scale test may be used as either a self-rated or observer-rated scale. It contains 48 items to measure mood and has been shown to be sensitive to medicine effects. The test takes 5–15 min to complete and measures the immediate present in a patient or

normal individual. The test gives six scores: friendly, aggressive, clear thinking, sleepy, unhappy, and dizzy.

28.7.12.7 Covi Anxiety Scale The Covi Anxiety Scale is a global observer’s rating scale of patient anxiety. There were three items that are each rated on a 0–5 scale. The test is simple to use and requires only a few minutes to complete.

28.7.12.8 Crichton Geriatric Rating Scale The Crichton Geriatric Rating Scale test measures the level of behavioral function in elderly psychiatric patients using a five-point scale on 11 items. It rates either the present or the period within the last week and takes 5–10 min to complete.

28.7.12.9 Depression Status Inventory The Depression Status Inventory (DSI) Scale is the professional’s version of the Zung Self-Rating Depression Scale (SDS). Each of the two scales (DSI and SDS) consists of the same 20 items rated on a four-point scale and is applied to adults with depressive symptomatology. The DSI is completed by the professional, and the SDS is completed by the patient. Both tests take about 5–10 min to complete. The DSI rates either the present situation or the last week prior to the test, and a total score is obtained.

28.7.12.10 Hamilton Anxiety Scale The Hamilton Anxiety Scale (HAMA) was designed to be used in adult patients who already have a diagnosis of anxiety neurosis rather than for making a diagnosis of anxiety in patients who have other problems. The test contains 14 items, each with a five-point scale, and is completed by a physician or psychologist. The test emphasizes the patient’s subjective state. The two subscales determined are somatic anxiety and psychic anxiety.

28.7.12.11 Hamilton Depression Scale The HAMD is one of the most widely used tests to evaluate the severity of depressive illness quantitatively in adults. The most widely used form of this test contains 21 items covering a broad range of symptomatology, with a three- to five-point scale for most items. The minimum time required to complete this test is usually 10–20 min, and it requires a skilled interviewer. Either the present time or the period within the last week is rated. Six subscales are obtained in the HAMD: anxiety/somatization, weight, cognitive disturbance, diurnal variation, retardation, and sleep disturbance.

28.7.12.12 Hopkins Symptom Checklist The Hopkins Symptom Checklist (HSCL) is a scale that has been used to measure the presence and intensity of various symptoms in outpatient neurotic patients. It is a 58-item self-rating scale and has generally been replaced by the Self-Report Symptom Inventory (SCL-90). It measures the symptoms during the past week and requires approximately 20 min to complete.

There are five subtests: somatization, obsessive/compulsive, interpersonal sensitivity, depression, and anxiety.

28.7.12.13 Inpatient Multidimensional Psychiatric Scale

The Inpatient Multidimensional Psychiatric Scale (IMPS) is used to measure psychotic syndromes in hospitalized adults capable of being interviewed. The 89 items are rated on the basis of a psychiatric interview. This test has been well validated and requires 10–15 min following a 35–45 min interview. There are 10 scores: excitement, hostile belligerence, paranoid projection, grandiose expansiveness, perceptual distortions, anxious intropunitiveness, retardation and apathy, disorientation, motor disturbances, and conceptual disorganization.

28.7.12.14 Nurses' Observation Scale for Inpatient Evaluation

The Nurses' Observation Scale for Inpatient Evaluation (NOSIE) (30-item test) is used by nursing personnel to rate a patient's behavior on the ward, with a five-point scale for each item. This test is widely used and is well accepted for adult inpatients. The test, which rates the most recent 3 days, is relatively easy to use and requires 3–5 min to complete.

28.7.12.15 Plutchik Geriatric Rating Scale

The Plutchik Geriatric Rating Scale (PLUT) (31-item test) is designed to measure the degree of geriatric functioning in terms of both physical and social aspects. The three-point scale for each item is completed on the basis of direct observation of the patient's behavior and takes 5–10 min to complete. The subscales measure overall dysfunction, aggressive behavior, sleep disturbance, social isolation, sensory impairment, work and activities, and motor impairment.

28.7.12.16 Profile of Mood States

Profile of Mood States (POMS) self-rating scale is used in both normal and psychiatric outpatients to evaluate feelings, affect, and mood. It has been widely used in medicine trials. The 65 adjectives included in this test may be used to rate the present and/or previous week. This test requires approximately 5–10 min to complete and provides scores for six subtests: tension/anxiety, depression/dejection, anxiety/hostility, vigor, fatigue, and confusion.

28.7.12.17 Sandoz Clinical Assessment-Geriatric

The Sandoz Clinical Assessment-Geriatric (SCAG) test measures 18 individual symptoms plus a global rating using a seven-point scale similar to those used in the BPRS. It measures the present period or that within the last week, requires about 10–15 min to complete, and does not contain subtests.

28.7.12.18 Self-Report Symptom Inventory

Each of the 90 items in the SCL-90 uses a five-point scale of distress. It was designed as a general measure of symptomatology for use by adult psychiatric outpatients in either a research or

clinical setting. It rates either the present or previous week. It requires about 15 min for the patient to complete this form and about 5 min for a technician to verify identifying information. This test is sensitive to drug effects and may be used with inpatients. Nine subscales are measured: somatization, obsessive/compulsive, interpersonal sensitivity, depression, anxiety, anger/hostility, phobic anxiety, paranoid ideation, and psychoticism.

28.7.12.19 Wittenborn Psychiatric Rating Scale

The ECDEU version (Wittenborn Psychiatric Rating Scale (WITT)) is a 17-item test shortened from the original 72-item test. All but one item use a four-point scale, and the test takes 5–10 min to complete. It is used in both in- and outpatients and rates either the present or previous week. This test is not intended to make diagnoses but to reflect changes within one patient and to provide a basis for comparing different patients. This test provides descriptive, as opposed to etiological or prognostic, information on patients and includes the following subscales: anxiety, somatic/hysterical, obsessive/compulsive/phobic, depressive retardation, excitement, and paranoia.

28.7.12.20 Zung Self-Rating Anxiety Scale

The SAS test requires approximately 5–10 min to complete.

28.7.12.21 Zung Self-Rating Depression Scale

The SDS test requires approximately 5–10 min to complete.

28.7.13 Pediatric Behavioral Rating and Diagnostic Scales

Many of the behavioral rating scales described for adults are not suitable for use in the pediatric population. Special tests have been designed, and a number of pediatric behavioral rating scales are presented in Table 28.25. General comments on these tests are presented later. A further description of rating scales used in pediatric medicine trials is given in the *ECDEU Assessment Manual for Psychopharmacology* by Guy (1976). His article is a practical guide to identifying appropriate scales for a particular situation. Connors discusses the two broad approaches of many pediatric rating scales as either "rating current behaviors, symptoms or states; or ... describing basic traits, dispositions, and personality characteristics." The choice of one of these two approaches depends in part on the purpose of using a scale in a medicine trial. Three general purposes have been suggested for using a behavioral test (prediction, measurement of change, and classification). The choice of one of these three purposes usually implies that one of the two specific approaches implicit in the pediatric behavioral scales will be more appropriate:

1. To be able to predict something about a patient, choose a scale that rates basic traits.

2. To measure change in a patient, choose a scale that rates current symptoms.
3. To assess a patient's classification, choose a scale that rates either basic traits or current symptoms, depending on the purpose of the classification.

The type of patient population and the desired format of the test to be used in a clinical trial also influence the particular scale(s) chosen.

An evaluation system that can be used in a wide variety of pediatric inpatients is the Children's Behavior Inventory.

28.7.13.1 Children's Behavior Inventory The Children's Behavior Inventory (CBI) is a 139-item, two-point (yes–no) scale to record maladaptive behavior in children aged 1–15 years. Relatively little training is needed to administer this test. It is easily used by nurses, teachers, graduate students, psychologists, and others. This test usually requires at least 2 h of observation of the child, but better reliability is achieved if behavior is observed over an 8 h period. Nine subtest scores are provided: anger/hostility, conceptual dysfunctioning, fear and worry, incongruous behavior, incongruous ideation, lethargy/dejection, perceptual dysfunctioning, physical complaints, and self-deprecation.

28.7.13.2 Children's Diagnostic Classification Children's Diagnostic Classification (CDC) test may be used instead of the Children's Psychiatric Rating Scale (CPRS) to arrive at a diagnosis. This differs from the CPRS in that it is highly directed and leads the observer to a diagnosis. It rates the current status of the child and may be used at pretreatment and/or the termination of the clinical trial.

28.7.13.3 Children's Diagnostic Scale The Children's Diagnostic Scale (CDS) is used in children up to 15 years of age to assist in the diagnosis and classification of the child's condition. It contains 13 items, eight of which have a seven-point scale. The others are specific diagnostic questions. It measures current status only and is mainly used at the start of a study, although it may be used at the termination of the study as well.

28.7.13.4 Children's Psychiatric Rating Scale The CPRS is a comprehensive scale to assess a wide range of psychopathologies in children up to age 15. It contains 63 items, with a seven-point scale derived from the BPRS. This test rates 28 items by direct observation of the child based on behavior expressed during the interview and rates other items based on the child's reports of events that occurred either over the preceding week or that occurred during the interview. Scores of 15 separate clusters of the rated items are provided as well as the overall score.

28.7.13.5 Clinical Global Impression See adult behavioral rating scale description of the CGI.

28.7.13.6 Conners Parent Questionnaire The Conners Parent Questionnaire (PQ) is a 94-item checklist of symptoms that evaluates common behavior disorders using a four-point scale in children up to 15 years of age and takes 15–20 min to complete. It is used once pretreatment and may be repeated but is often replaced after the first use by the 11-item Conners Parent/Teacher Questionnaire (PTQ). There are eight subscales: conduct problem, anxiety, impulsive/hyperactive, learning problem, psychosomatic, perfectionism, antisocial, and muscular tension.

28.7.13.7 Conners Parent/Teacher Questionnaire See previous descriptions for Conners PQ and below for the Conners Teacher Questionnaire (TQ). The PTQ is used in conjunction with either the PQ to TQ and yields a total score only (i.e., no subscales are given). The PTQ takes about 5 min to complete and is not used pretreatment.

28.7.13.8 Conners Teacher Questionnaire The TQ form was designed to obtain teacher evaluations of children up to age 15 in terms of their interactions with peers and their ability to cope with the school environment and requirements. There are 41 items, and the first 39 have a four-point scale. Question 40 deals with the teacher's evaluation of the child's severity of illness, and question 41 deals with global improvement in four different areas. This test is used once at pretreatment and as needed afterward. It takes about 15 min to complete and covers either the present or any interval period up to 1 month. A shorter 11-item PTQ is often used after the initial use of the 41-item TQ. The five subscales included are conduct, inattentive/passive, tension/anxiety, hyperactivity, and social ability.

28.7.13.9 Devereux Child Behavior Rating Scale The Devereux Child Behavior Rating Scale contains 97 items and is similar to the Devereux Teacher Scale. It is used for emotionally disturbed and mentally retarded children aged 8–12 years. Besides being easy to use, this scale is well researched and discussed in the literature. It requires 10–20 min to complete by clinicians, child care workers, parents, or others and gives 17 scores. There is a Devereux Adolescent Behavior Rating Scale for children from ages 13 to 18.

28.7.13.10 Devereux Elementary School Behavior Rating Scale The Devereux Elementary School Behavior Rating Scale is a widely used test incorporating 47 items that have high test–retest reliability. It uses a checklist format and is easy to use (requires 10 min). There are 11 factor scores and 3 item scores.

28.7.14 Psychometric and Performance Tests

The psychometric and performance tests presented in Table 28.26 may be grouped as being applicable for use in either children or adults. In children, the tests measure

intellect (Goodenough–Harris Figure-Drawing Test (GOOD), Porteus Mazes, Wechsler Intelligence Scale for Children (WISC), Peabody), achievement (Wide Range Achievement Test (WRAT)), and motor performance (vigilance tests, reaction time). There are other tests that may be used to measure learning, although many of these tests utilize equipment and are not described. All of these tests (unless otherwise noted) are given once pretreatment, at least once posttreatment, and at additional times if desired by the investigator. The contribution of learning in the scores obtained at second and third testings is usually unknown. The methods used to motivate patients to perform to the best of their ability in all tests must be standardized and reported.

28.7.14.1 Bender-Gestalt Test The Bender-Gestalt Test is a nonverbal performance test in which the individual copies a design shown on a card. It is often used to identify a problem of visual perception and/or motor performance or minimal brain dysfunction in children.

The scoring used for children (age 4 or 5–11 years) differs from that used for adults (age 15–adult). This test measures perceptual maturity, possible neurological impairment, and emotional adjustment in children. It measures maturation, intelligence, psychological disturbance, and cortical impairment in adults. The test requires 10 min to complete. Scores may fluctuate from test to test and thus must be interpreted carefully.

28.7.14.2 Conceptual Clustering Memory Test For the Conceptual Clustering Memory Test, patients are given a list of 24 specific words from a number of different categories such as birds, cars, or types of drinks. The words are presented one at a time over 2 min, after which patients are asked to recall as many of the specific words as possible. The test measures the total recall as well as the degree to which words of a specific category (e.g., animals) are recalled from the cluster of words given in that category (e.g., dog, cat, cow).

28.7.14.3 Digit Symbol Substitution Test A subtest of the Wechsler Adult Intelligence Scale (WAIS), the Digit Symbol Substitution Test, measures sensorimotor integration and learning relationships of symbols. It has been used in many psychopharmacological studies. Subjects are given different forms of this test at each session. The test requires the patient to match as many of 100 symbols to their respective numerals, found in a code key, as possible within 60 s.

28.7.14.4 Embedded Figures Test For the Embedded Figures Test, patients are shown a complex design and must identify as quickly as possible a simple figure that is “embedded” within the design. Twenty-four embedded figures are included, and a maximum of 3 min is allowed for each one.

28.7.14.5 Frostig Developmental Test of Visual Perception The Frostig Developmental Test of Visual Perception (FROST) measures the development of perceptual skills in children from 4 to 8 years of age or in older children with learning difficulties. It may be administered individually (requires 30–45 min) or to groups (requires 40–60 min).

28.7.14.6 Goodenough–Harris Figure-Drawing Test The GOOD is a brief (10–15 min) easy-to-use test for children 4–15 years of age to measure intellectual maturity.

28.7.14.7 Peabody Picture Vocabulary Test The Peabody Picture Vocabulary Test is a rapid 10–15 min intelligence test for children aged 2.5–18 years that is useful when there is inadequate time to give the WISC.

28.7.14.8 Porteus Mazes The Porteus Mazes is a nonverbal test that has been shown to be sensitive to medicine effects in both children (over 3 years) and adults. The test has three series of mazes to prevent score improvement on retesting with the same test. It requires about 25 min and provides both a qualitative and quantitative score.

28.7.14.9 Reaction Time There are many different tests used to measure reaction time. These tests measure the period of time between the presentation of a stimulus to a patient and the onset of the resulting response. The signal is usually a visual or auditory stimulus, and the onset of a motor reaction, such as the lifting of a finger, arm, or leg or the pressing of a buzzer, is used to measure the speed of response.

In simple reaction times, a stimulus is presented that always requires the same response, even if the nature of the stimulus changes. A complex reaction time requires the patient to respond to some stimuli but not to others.

28.7.14.10 Vigilance Tests Numerous tests have been designed to measure vigilance. In these tests, patients are requested to respond in some manner to certain stimuli or occurrences but not to others. The stimuli may be controlled to present minimally perceived signals that require vigilance on the part of the patient.

28.7.14.11 Wechsler Adult Intelligence Scale The WAIS consists of 11 subtests—six verbal tests and five performance tests. This provides an age-related IQ in adults from 16 to 75 years of age, that is, the test measures intelligence of the person in relation to their age group and not to the entire population. It may be used either as an initial assessment or as a tool to measure change. The test, which takes 40–60 min to complete, provides 13 scores in verbal and performance categories plus a total score.

28.7.14.12 Wechsler Intelligence Scale for Children The WISC was extensively revised in 1974, and it became the

WISC-R, which requires 40–60 min to complete. This widely used scale in children from 6 to 16 years of age may be used for either screening or baseline data or as a measure of change. There is a “preschool and primary scale of intelligence” version that may be used in children from 4 to 6½ years of age (requires 50–75 min). The WISC-R has six verbal and six performance subtests.

28.7.14.13 Wechsler Memory Scale The Wechsler Memory (WMEM) Scale is a brief test that is used to measure memory deficits. There are two forms of the test, and they are generally alternated to avoid a training effect in children taking the test on two or more occasions.

28.7.14.14 Wide Range Achievement Test The WRAT is used in children from age 5 years to adults in college. It assesses basic skills in reading, spelling, and mathematics. It is simple, is easy to administer, and requires 20–30 min to complete.

28.7.15 Personality Tests

In addition to the aforementioned behavioral and performance tests, there are a number of well-known tests of personality that may provide useful information in select clinical studies. The most well known of these tests is the Minnesota Multiphasic Personality Inventory (MMPI). This test consists of 550 affirmative statements to which a true or false response is given and requires about 1 h to complete. It is given to adults over the age of 16 and is scored for 10 scales: depression, hysteria, hypochondriasis, psychopathic deviate, masculinity/femininity, paranoia, hypomania, schizophrenia, psychasthenia, and social introversion.

28.8 ASSESSMENT OF UNWANTED DRUG EFFECTS

Because of the relatively small numbers of volunteers and patients involved, only the most common of drug-related AEs are likely to be detected during early studies. For example, to have a 95% chance of picking up three subjects who have experienced an adverse reaction (with no background incidence) which occurs in 1 in every 100 subjects treated with the drug, it would need to be given to 650 subjects (Gad, 1998). Matters are made worse when the AE in question also occurs in the general population, which is usually the case with the kind of symptoms reported by volunteers and patients taking part in drug studies. No matter how good the study design, nothing can compensate for this problem of inadequate numbers. In this respect, all of the study designs described earlier are more or less equally adequate or inadequate as the case may be.

Monitoring for drug-related AEs employs the same or similar methods in both volunteers and patients. In both cases

assessments of tolerability and safety are based upon symptom reports, routine laboratory safety screens, ECG monitoring, and on occasions special tests designed to detect unwanted effects associated with a particular class of drug. The chances of obtaining reliable information on a drug's safety profile are enhanced by detailed and careful monitoring. Symptoms may be reported spontaneously or elicited in reply to standard questions. Open questions such as “how are you feeling?” are to be preferred to leading questions on the basis that they result in fewer reports of AEs. If leading questions are used they need to be carefully worded. A certain amount of basic information is required on all AEs, that is, type, severity, time of onset in relation to time of dosing, duration, and causality. Attributing the cause of an unwanted effect to the drug or some other factor can be difficult, particularly when little is known about the drug, as is often the case at the stage of initial studies in volunteers or patients. Rechallenge with the drug ideally using the same dose or if need be (because the event caused a degree of discomfort) a reduced dose is probably the single best way of proving or disproving a causal relationship; but if done the rechallenge procedure must be designed using placebo as a comparator under double-blind conditions. Obviously rechallenges can be done only if the AE was reversible, did not cause excessive discomfort, and most importantly was not life threatening.

28.8.1 Separation of Adverse Reactions from Placebo Reactions

Since adverse nondrug symptoms are common (Reidenberg and Lowenthal, 1968) and are not easily separated from drug-induced symptoms, both must be collected for analysis if a complete profile of adverse reactions is to be made. However, this technique can only be used in controlled studies, ideally with placebo, as well as with other standard drugs. The temptation to subtract the number of the particular AR in the placebo group from the number in the active drug group as follows:

Drug group – placebo group = number of adverse reactions to drugs should be resisted because:

- The difference may not be statistically significant and may have arisen by chance.
- Although the total number of events may be statistically different in the two treatment groups, it is also necessary to establish whether the numbers of patients afflicted with the AE are different and vice versa.
- Having established that there is a significant difference between the two treatment groups for the number of events and the number of patients afflicted, the severity of the ADRs in the two groups should be compared.

A further problem is that due to classification, some terms may include more than one type of abnormality (e.g.,

the incidence of “blurred vision” may be equal in both groups, but there may be several cases of tunnel vision with the trial drug, but because there is no code for tunnel vision it is coded under a more general terms). Another problem is that the symptoms forming a syndrome are often coded separately and individually there may be no difference between two drugs, but when the cases are examined there may be a combination of symptoms with one drug that warrant being called a syndrome. It is therefore essential to read the individual original description of the AEs before making a judgment. This area has been explored more fully by Bernstein, and he has added to the equation with the addition of bias:

Attributable AEs = drug group AEs – placebo group AEs \pm bias

Bias is equal to the baseline (*B*) frequency and severity of the AE multiplied by pharmacological clinical activity of the drug (AD) minus the pharmacological clinical activity of the placebo (AP):

$$\text{Bias} = B(\text{AD} - \text{AP})$$

The argument is that the disease or a symptom or sign of the disease and the drug ADR may interact as follows:

- Compliance—Early improvement may cause the patient to stop the drug, and the improvement of the ADR may be inappropriately assigned to tachyphylaxis of the ADR; failure of the disease to improve may persuade the patient to add a rescue drug or increase the dose of the study drug or even stop the drug; impaired mental or cognitive function due to the disease may affect compliance.
- The disease may alter the absorption, distribution, metabolism, or elimination of the drug (e.g., alteration of the blood–brain barrier by the disease may allow the drug to affect the brain).
- Observational bias of convalescence (e.g., severe pain causing insomnia may require morphine causing compensatory hypersomnia in excess of that caused by morphine alone).
- Observational bias by halo effects—Perception of an ADR may be swamped by the symptoms of the disease; thus as the disease symptoms resolve, the ADR becomes apparent.
- Unblinding—If the patient or physician are unblinded due to rapid improvement of the disease or an ADR, they may be led to expect ADRs with the active treatment.
- Pharmacological clinical activity bias—An AE that is already present due to the disease may be increased if is also an ADR of the drug or vice versa. For example, the

diarrhea of gastroenteritis may be alleviated by code in-containing preparations given to relieve pain, while the inertia of a severely depressed patient may be sufficiently resolved by an antidepressant to enable the patient to commit suicide.

ADRs that are similar to common nondrug AEs are rarely described or investigated sufficiently for a causal relationship for each individual event to be established. If they cannot be distinguished qualitatively, the correct quantitative procedure is to compare them using nonparametric statistics, giving the confidence limits for the incidences of ADRs. Small studies ($n < 30$) have little chance of separating ADRs from placebo or nondrug events unless they are very common and specific to the drug. The situation is worsened by the fact that members of a placebo group have a tendency to “catch” AEs from the active drug group, therefore changing a relatively specific ADR to a nonspecific event.

28.8.1.1 Base-Case Causality of Single-Event Adverse Drug Reactions

The analysis and evaluation of ADRs is a major problem in both the development of new drugs and the postmarketing surveillance period. Just as there are standards and requirements established as guidelines in the chemical, pharmacological, and toxicological phases preceding the marketing of a new drug, there are also guidelines for causality assessment of individual human cases in both pre- and postapproval for new drugs.

The most common problem of assessment is the single-event ADR case. Presented here is an approach to such single-event ADR cases. This methodology, relating to the use of therapeutic-, diagnostic-, and prophylactic-type drugs in a clinical setting, should permit the diagnostician to make one of three responses after an assessment of an ADR case: an assured *yes*, a firm *no*, or a reasoned admission of *uncertainty*. The clinicopathological picture presented by the ADR case is often not readily distinguishable from nondrug-induced diseases. The clinical and morphologic findings of ADRs have the same limited number of final common paths that characterize these other (nondrug) human illnesses.

There are three major requirements for establishing the occurrence of an ADR:

1. The possibility and likelihood of a causal relationship between the drug and the ADR must be confirmed by establishing its eligibility.
2. Linkage of the drug with the clinicopathological findings made.
3. The degree of certainty of this drug linkage should be determined.

As an initial background for developing this algorithm or methodology, Figure 28.5 is offered for consideration and

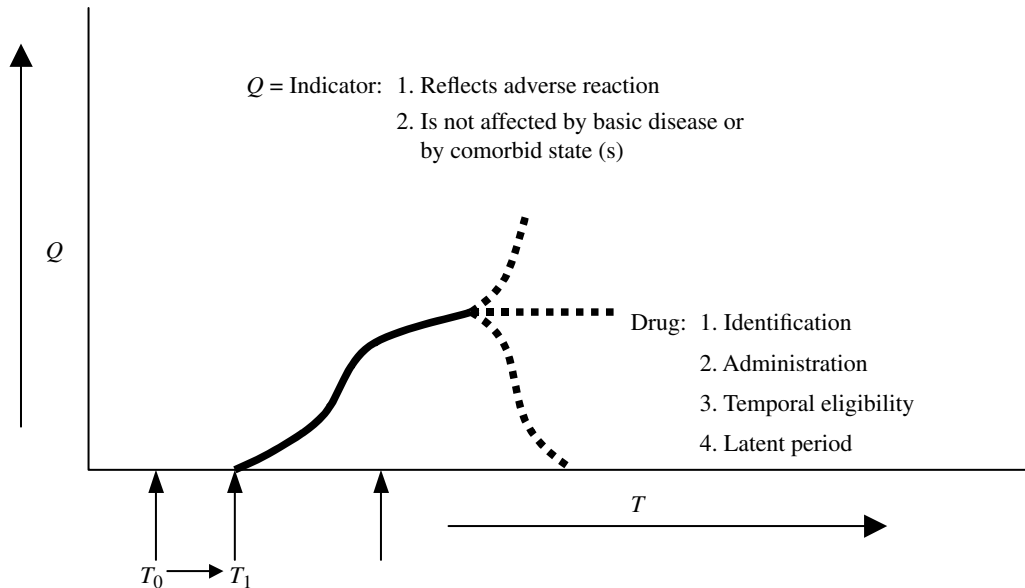


FIGURE 28.5 Adverse drug reaction (curve Q) plotted against time (abscissa T). Dashed lines show three courses an ADR can take: increasing severity to death, leveling off to chronicity, or return to abscissa, indicating recovery. Four criteria that must be met before drug is eligible to be empiric correlate of Q (adverse drug reaction) are listed.

orientation. This figure has the basic elements of a “time flowchart,” which has considerable utility in evaluating ADR cases.

In this graphical representation of an ADR, the ordinate (Q) represents any of the findings of an ADR. *Specifically*, Q may be a symptom (pain, nausea, etc.), a sign, a clinical laboratory result, a radiological finding, a morphological finding, or any combination of these. Synonyms for Q include marker, disease marker, signal, indicator, parameter, detector, response, and effect.

The abscissa is the time element (T), related to both the time of drug administration and the dating of disease marker data. Both are usually plotted on the same time flowchart in a particular case.

This graphical representation of an ADR case will be used frequently in the assessment of eligibility and linkage determinations of ADRs. The four eligibility criteria are also listed in Figure 28.5.

Administration of the Drug As it is with accurate identification of a drug, so it is that its “administration” must at times be held in question. Subject compliance with the study protocol is not a rare problem in clinical trials. Complete noncompliance sometimes occurs.

Temporal Eligibility The time factor in assessment of ADRs is a very important one and in some cases is of critical diagnostic importance. This is true not only in establishing “eligibility” of the drug but also in linking the drug to the reaction. On the other hand, the time element may be equally important in denying eligibility and also make linkage of the

drug with the clinicopathological picture a most unlikely possibility.

It is quite apparent that a drug cannot be responsible for an ADR if the latter is already in progress before the drug is first administered. This dyssynchronicity is sometimes seen in both trials and later in the marketplace.

Latent Period Latent period refers to the time interval between the initial administration of the drug and the onset of the ADR (in Figure 28.6, it is the interval T_0 to T_1). The latent period is not rigidly fixed or exactly predictable, but it tends to fall within certain limits.

Characteristically, strychnine deaths occur in seconds to minutes. Most anaphylactic deaths occur within 20–30 min after contact with the lethal antigen, while jaundice associated with most drugs has its onset within 3 days to 3 weeks after the beginning of therapy. The fatal pancytopenia following chloramphenicol appears in 1–3 months, while hepatic angiosarcoma related to aflatoxin has a latent period of one to several decades. The ultimate in length of latency is one to several generations from a drug-induced mutational germ cell change to its manifestation in a conceptus.

Consideration of the latent period in an ADR is of use in an ADR assessment in one of two ways: the latent period may be too long, or it may be too short.

In summary, identification, administration, temporal eligibility, and latent period are the four criteria for establishing the eligibility of a drug to have caused an ADR. Emphasis should be placed on obtaining sufficiently detailed time-related data on drug administration and on the appearance of

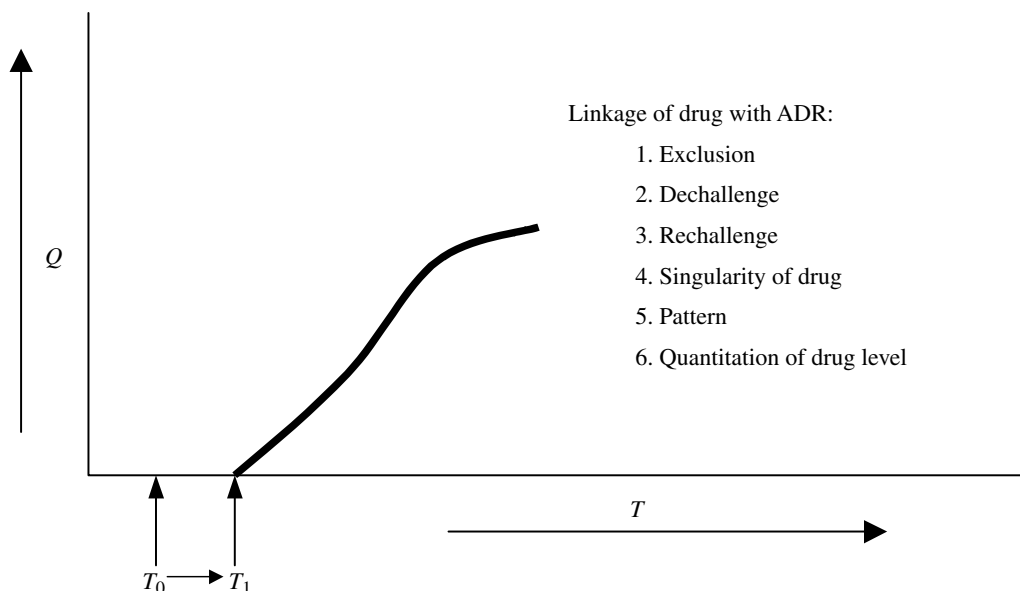


FIGURE 28.6 Six methods of linking drug with adverse drug reaction.

ADR markers. These data are a *sine qua non* in the assessment of drug eligibility.

Linking the Drug with the Clinicopathological Findings

The second major task in analyzing an ADR case is to establish a connection or linkage between the drug and the clinicopathological findings (making empiric correlates of the drug and these findings).

Figure 28.6 is a time flowchart representing an ADR that itemizes six ways of making this linkage.

Exclusion Exclusion consists of selecting one drug from a group of drug candidates by the use of the time flowchart.

The exclusion method also includes instances in which drug candidates are themselves excluded from causation status because a nondrug etiology is clearly demonstrable (environmental or occupational factors, radiation injury, the underlying disease of the patient, or a comorbid state) that can reasonably account for the clinicopathologic findings.

Dechallenge The principle involved in the dechallenge method of linkage is that if there is a reversible effect present, then removing the cause will eliminate the effect.

Rechallenge The principle involved in the rechallenge method of linkage of a therapeutic to an ADR is implied in the phrase *post hoc ergo propter hoc* (after this therefore because of this).

As applied, if a drug has been incriminated with a reaction and the ADR disappeared when the drug was discontinued, a rechallenge with this drug followed by a return of the ADR would increase the probability that the drug and the

ADR were empiric correlates. While intentional challenge is not often done, such a rechallenge may occur inadvertently.

Singularity of the Drug The principle involved in the singularity method of linking a drug with an ADR is based on two assumptions: only one drug was administered, and there was no basic disease or comorbid state that could be related to the ADR marker being used in the assessment.

Pattern The pattern method of linking a drug with an ADR shifts the focus of attention to the clinicopathological findings in an ADR and away from the identification of the causative drug. This shift of emphasis is necessary when detailed time-related drug and disease marker data are unavailable to the evaluator of the case. The site-process profile may then be used as a guideline for searching past experience and the literature for cases that have matching features. Matching features found in the literature may include associations with certain drugs or chemicals, which serves as a guideline for a focused examination of the patient's history for the causative agent.

This "pattern" method may also be used in excluding drugs. If the drugs or chemicals suggested by the morphologic findings are not identified or disclosed by historical or toxicological efforts, then the morphologic changes appear to remain nondrug or nonchemical related.

Quantitation of Drug Level Assessing an ADR case by quantitation of drug level brings our focus back to the search for and the identification of the causative agent by quantitative and objective data based on laboratory analysis of body fluids and/or viscera (Ozdemir et al., 2001). This method is applicable and strongest in the case of higher dose level.

The feasibility of this approach is based on the availability of dependable information on lethal levels from past experience or preclinical work. Without this comparison information, there is no judgmental significance to toxicological levels in the case at hand.

Quantitated levels of drugs have limitations in diagnostic value. In adverse reactions in the hypersensitivity, idiosyncratic, and pharmacogenetic categories, drugs have been administered in therapeutic (not toxic) amounts, and blood and other body fluids and tissue levels have been found to lie within therapeutic ranges. Such analyses will confirm any prior administration of the drugs, but the problem of the etiological differential diagnosis will still remain.

In addition to quantitative approach, qualitative identification can be of value in appropriate instances. In some cases of ADRs, more than one of the six methods of drug linkage that are listed in Figure 28.6 may be used in causation analysis. In fact, multiple methods in the same case strengthen the confirmation of the rejection of an ADR and its etiology.

Difficulties in Assessing ADRs Requirements for establishing eligibility and methods of linking a drug with an illness have been presented in the preceding discussion. This algorithm should constitute of blueprint for solving many if not most of the ADR problems in this area of medical diagnostics.

However, in the hands-on practice of the assessment of ADR cases, there are at least four major difficulties that stand in the way of such high diagnostic expectations. These four obstacles include the following:

1. *Incomplete information* Incomplete information is not unique to ADR evaluation but is common to all areas of medical practice. The lack of sufficiently detailed time-related data on drug administration and disease markers may make it impossible to render a reasoned judgment on many ADR cases, leaving them in their original and unsatisfactory anecdotal status. Denial to the evaluation of access to these required facts makes it impossible to make judgments on latent period and temporal eligibility; time flowcharts cannot be utilized in exclusion, dechallenge, and rechallenge techniques. The diagnostic database should also include information on any other drugs being administered or taken, concurrent comorbid states, and the existence of any preexisting occupational and environmental hazards.
2. *Polypharmacy* In recent times, polypharmacy is the rule rather than the exception. Patients with complicated and prolonged illnesses may have 20–30 medications in their medical background. Cases of this sort may be of such complexity that even with ideally complete drug and disease marker information, diagnostic success may be elusive.

3. *Lack of objective means of linking the drug to the ADR* Tests and procedures that specifically and causally connect a drug to an illness are lacking. Our high-tech laboratory instrumentation is capable of identifying and quantifying extremely low levels of drugs and chemicals, but this type of information falls short of establishing causation.
4. *The limited number of toxicologic responses in human disease* There are a limited number of generic morphological reaction patterns that diseases fit into (inflammatory, congenital, neoplastic, degenerative, infiltrative, vascular, and functional). In parallel, there are also a rather limited number of clinical symptoms and signs (pain, nausea, fever, lumps, etc.) that come to the attention of the practicing physician. There are a multitude of causes and a multitude of clinical conditions that funnel into these clinicopathological “final common paths.” The algorithm previously described is an attempt to move from the generic to the specific in analyzing ADR causation.

Of the previous four difficulties, only the first (incomplete information) is subject to at least some degree of improvement.

Degree of Certainty The third major task in analyzing and assessing ADRs is determining the degree of certainty one has as to the causal relationship between the drug and the clinicopathological findings. Interposed between the definitive causative and negative categories are three shades of certainty (probable, possible, and coincidental) that titrate between these two extremes. These degrees of certainty are defined as follows:

1. *Causative* Cases in this class are those in which there is no doubt that a drug has caused the reaction. This category is essentially limited to drug overdose cases or those cases in which the causative agent can be objectively identified (asbestos bodies, granuloma-encapsulated silica, etc.). Parenthetically, the overdose cases with drug levels in lethal ranges should have important negative findings: no anatomical cause of death at autopsy.
2. *Probable* This term is equivalent to the phrase “consistent with,” and cases in this category of certainty fall short of the “causative” designation because they lack an objective and quantitative laboratory finding that is the sine qua non of the causative category. Cases placed in this category have the following characteristics:
 - a. The criteria of temporal eligibility and appropriateness of latent period have been met.

TABLE 28.29 Major Causes of Acute Functional Adverse Drug Reactions

| Acute Adverse Drug Reaction | Example |
|--|---|
| Augmented (“supratherapeutic”) effect of interaction with the primary molecular target | Pronounced bradycardia with a β -blocker; pronounced hypotension with an angiotensin II receptor antagonist |
| Interaction with the primary molecular target present in nontarget tissues | Sedation caused by antihistamines |
| Interactions with secondary molecular targets | Interactions with the hERG cardiac channel leading to QT interval prolongation (e.g., some antipsychotics and antihistamines drugs) |
| Nonspecific effects | |
| Pharmacologically active metabolites | Inhibition of the hERG channel transcription by desmethylamiodarone, metabolite of amiodarone |

Source: Adapted from Gad (2009).

- b. The clinicopathological features are consonant with previous experience and literature precedent for the drug in question.
 - c. Other causes (the basic disease, comorbid states, and other modalities of therapy) have been eliminated from consideration.
 - d. One or several means of linkage of the drug to the ADR have been utilized: exclusion, dechallenge, rechallenge, singularity of the drug, and pattern.
3. *Possible* Cases are put in this category when the relationship between the drug and the clinicopathological findings can be neither confirmed nor denied. There are three subdivisions in this category:
- a. Cases with potential causes other than the drug in question. The clinicopathological picture could have been produced by the basic disease, a comorbid state, or some other modality of therapy.
 - b. Cases in which some of the criteria for eligibility and linkage have been met, but some have not because of lack of adequate information. Such a case could be put in this category temporarily while awaiting more information or placed here permanently if it were evident that further data would not be forthcoming.
 - c. Cases that have met all the criteria of eligibility and linkage but for which there is no known precedent literature. Such a case might be a new and emergent ADR. It could be placed in the “possible” group, awaiting the appearance of similar cases for cluster studies at a later time.
4. *Coincidental* Cases in this category include those that were indeed exposed to the drug in question but in which assessment of the case clearly reveals only an anecdotal association.
5. *Negative* This category applies to those cases in which the alleged drug was not or could not have been in the patient’s system at the time of the ADR. This circumstance could be related either to noncompliance, mislabeling of the drug, or historical misinformation.

Clinical trials are the final step in the drug development process and are the slowest to evolve and integrate new technology. The paradigms of plasma levels being directly related to (and predictive of) efficacy and safety are only slowly being supplanted by understanding molecular and receptor occupancy-based correlations. Table 28.29 provides a classification of such relationships to AEs.

REFERENCES

- Acuña G, Foernzler D, Leong D, Rabbia M, Smit R, Dorflinger E, Gasser R, Hoh J, Ott J, Borroni E, To Z, Thompson A, Li J, Hashimoto L, Lindpaintner K. (2002) Pharmacogenetic analysis of adverse drug effect reveals genetic variant for susceptibility to liver toxicity. *Pharmacogenomics J.* 2(5):327–334.
- Al-Khatib SM, Califf RM, Hasselblad V, Alexander JH, McCrory DC, Sugarman J. (2001) Placebo controls in short-term clinical trials of hypertension. *Science.* 292(5524):2013–2015.
- Association of American Universities. (2000) *Task Force on Research Accountability: Report on University Protection of Human Beings Who Are the Subjects of Research.* Washington, DC, June 28, 2000. Available at: www.aau.edu/HumSUBRpt06.28.00pdf (accessed April 7, 2016).
- Biomarkers Definitions Working Group. (2001) Biomarkers and surrogate endpoints: preferred definitions and conceptual frameworks. *Clin. Pharmacol. Ther.* 69:89–95.
- Bohazchuk W. (1998) Ethics committee and IRB audit results. *Good Clinical Practice*, pp. 46–55.
- Braunwald E. (2008) Biomarkers in heart failure. *N. Engl. J. Med.* 358:2148–2159.
- Buros OK. (1978) *The Eight Mental Measurements Yearbook.* Gryphon Press, Highland Park, NJ.
- Cato AE, Sutton L, Cato AE III. (2002) *Clinical Drug Trials and Tribulations*, 2nd Ed. Marcel Dekker, New York.
- CFR, Title 21, Parts 50, 54, 56, 312, 314. U.S. Government Printing Office, Washington, DC. Available at: www.access.gpo.gov/nara/cfr/ (accessed April 7, 2016).
- Curran CF, Sills JM. (1997) An index of United States federal regulations and guidelines which cover clinical safety surveillance of biological products. *Drug Inf. J.* 31(3):843–847.

- DHEW (1974) *The Institutional Guide to DHEW Policy on Protection of Human Subjects, Public Health Services Act*. Department of Health, Education, and Welfare (DHEW), Washington, DC.
- Daly AK, Aithal GP, Leathart JB, Swainsbury RA, Dang TS, Day CP. (2007) Genetic susceptibility to diclofenac-induced hepatotoxicity: contribution of UGT2B7, CYP2C8, and ABCC2 genotypes. *Gastroenterology*. 132(1):272–281.
- Danoff TM, Campbell DA, McCarthy LC, Lewis KF, Respasch MH, Saunders AM, Spurr NK, Purvs IJ, Roses AD, Xu C-F. (2004) A Gilbert's syndrome UGT1A1 variant confers susceptibility to tranilast-induced hyperbilirubinemia. *Pharmacogenomics J*. 4:49–53.
- Davis R. (1998) Holes are growing in medical testing safety nets. *USA Today* 19A–20A.
- Downing NS, Aminawung JA, Shah ND, Krumholz HM, Ross JS. (2014) Clinical trial evidence supporting FDA approval of novel therapeutic agents. *JAMA*. 311(4):368–377.
- Ehlert FJ. (2015) *Affinity and Efficacy*. World Scientific, London.
- Eikelboom JW, Mehta SR, Pogue J, Yusuf S. (2001) Safety outcomes in meta-analyses of phase II versus phase III randomized trials. *JAMA*. 285:444–450.
- Edwards LD, Fox AW, Stonier PD. (2010) *Principles and Practice of Pharmaceutical Medicine*, 3rd Ed. John Wiley & Sons, Inc, Hoboken, NJ.
- FDC Reports, Inc.. (2000) FDA may request 'intermediate' enforcement authority for clinical trials. *Pink Sheet*. 62(19): 16–17.
- Fleiss JL. (1986) *The Design and Analysis of Clinical Experiments*. John Wiley & Sons, Inc., New York.
- Food and Drug Administration (FDA). (1997a) *Expedited Safety Reporting Requirements for Human Drug and Biological Products*. Federal Register, October 7, 1997, Vol. 62, Number 194, pp. 52237–52253. Food and Drug Administration (FDA), Silver Spring, MD.
- Food and Drug Administration (FDA). (1997b) *Postmarketing Expedited Adverse Experience Reporting for Human Drug and Licensed Biological Products; Increased Frequency Reports*. Federal Register of June 25, 1997, Vol. 62, Number 122, pp. 34166–34168. Food and Drug Administration (FDA), Silver Spring, MD.
- Gad SC. (1998) *Statistics and Experimental Design for Toxicologists*. CRC Press, Boca Raton, FL.
- Gad SC. (Ed.) (2009) *Clinical Trials Handbook*. John Wiley & Sons, Inc., Hoboken, NJ.
- Gallin JJ, Ognibene FP. (2012) *Principles and Practice of Clinical Research*, 3rd Ed. Elsevier, New York.
- Ganter J. (1999) Responding to industry critics: if the industry doesn't address concerns raised by the consumer press, who will? *Appl. Clin. Trials*. 8:10–11.
- Green S, Beneditti J. (2012) *Clinical Trials in Oncology*, 3rd Ed. Chapman SC Hall, Boca Raton, FL.
- Guy W. (1976) *ECDEU Assessment Manual for Psychopharmacology*. U.S. DHEW Pub. No. (ADM) 76-338. U.S. Government Printing Office, Washington, DC.
- Harning R. (2007) *Planning and conducting clinical trials in oncology*. Tutorial # 75. Teijin America at DIA Meeting, June 17, 2007, Atlanta, GA.
- Hartup J, Mann DL. (2013) Positioning of inflammatory biomarkers in the heart failure landscape. *J. Cardiovasc. Transl. Res.* 6(4):485–492.
- Hayashi K, Walker AM. (1996) Japanese and American reports of randomized trials: difference in the reporting of adverse effects. *Control. Clin. Trials*. 17:99–110.
- Henney JE. (2000) *Remarks Presented at the Association of American Medical Colleges Council on Teaching Hospitals Spring Meeting*, May 11, 2000, Washington, DC.
- Hill B-H. (1966) The environment and disease; association or causation? *Proc. R. Soc. Med.* 58:295–300.
- Huang YS, Chern HD, Su WJ, Wu JC, Chang SC, Chiang CH, Chang FY, Lee SD. (2003) Cytochrome P450 2E1 genotype and the susceptibility to antituberculosis drug-induced hepatitis. *Hepatology*. 37(4):924–930.
- Hudson KL, Collins FS. (2015) Bringing the common rule into the 21st century. *N. Engl. J. Med.* 373(24):2293–2296.
- ICH (1996) *E6(R1) Guideline for Good Clinical Practice*. International Conference on Harmonisation (ICH), Geneva, Switzerland.
- ICH. (2006) *ICH E8: General Considerations for Clinical Trials*. International Conference on Harmonisation, Geneva, Switzerland.
- Ioannides JPA, Lau J. (2001) Completeness of safety reporting in randomized trials: an evaluation of 7 medical areas. *JAMA*. 284:437–443.
- Jones JH. (1993) *Bad Blood: Tuskegee Syphilis Experiment*. Free Press, Berkeley, CA.
- Kaplowitz N. (2001) Drug induced liver disorders: implications for drug development and regulation. *Drug Saf.* 24:483–490.
- Kesselheim AS, Myers JA, Avorn J. (2011) Characteristics of clinical trials to support approval of orphan vs nonorphan drugs for cancer. *JAMA*. 305(22):2320–2326.
- Kindmark A, Jawaid A, Harbron CG, Barratt BJ, Bengtsson OF, Andersson TB, Carlsson S, Cederbrant KE, Gibson NJ, Armstrong M, Lagerström-Fermér ME, Dellsén A, Brown EM, Thornton M, Dukes C, Jenkins SC, Firth MA, Harrod GO, Pinel TH, Billing-Clason SM, Cardon LR, March RE. (2008) Genome-wide pharmacogenetics investigation of a hepatic adverse event without clinical signs of immunopathology suggests an underlying immune pathogenesis. *Pharmacogenomics J*. 8(3):186–195.
- Lee C-J, Lee LH, Wu CL, Lee BR, Chen M-L. (2006) *Clinical Trials of Drugs and Biopharmaceuticals*. Taylor & Francis, Boca Raton, FL.
- Machin D, Day S, Green S. (2007) *Textbook of Clinical Trials*, 2nd Ed. John Wiley & Sons, Inc., Hoboken, NJ.
- Mackintosh DR. (2001) Bye-bye placebos. *Appl. Clin. Trials*. 10:47–49.
- Matoren GM. (1984) *The Clinical Research Process in the Pharmaceutical Industry*. Marcel Dekker, New York.
- Meille C, Gentet JC, Barbolosi D, Andre N, Doz F, Iliadis A. (2008) New adaptive method for phase I trials in oncology. *Clin. Pharmacol. Ther.* 83(6):873–881.

- Meinert CL. (1996) An open letter to the FDA regarding changes in reporting procedures for drugs and biologics proposed in the wake of the FIAU tragedy. *Control. Clin. Trials*. 17:273–284.
- Mello MM, Francer JK, Weinick M, Teden P, Beirer BE, Barnes M. (2013) Preparing for responsible sharing of clinical trial data. *N. Engl. J. Med.* 369:1851–1858.
- Muller PY, Brennan FR. (2009) Safety assessment and dose selection for first-in-human clinical trials with immunomodulatory monoclonal antibodies. *Clin. Pharmacol. Ther.* 85(3):247–258.
- Nylen RA. (2001) *The Ultimate Step-By-Step Guide to Conducting Pharmaceutical Clinical Trials in the USA*. RAN Institute, Tampa, FL.
- O'Brien KL, Selaniko JD, Heedivert C. (1998) Epidemic of pediatric deaths from acute renal failure caused by diethylene glycol poisoning. *JAMA*. 279:1175–1180.
- O'Donnell P. (2005) A new GCP bible is born. *Appl. Clin. Trials*. 14(6):38–42.
- O'Donohue J, Oien KA, Donaldson P, Underhill J, Clare M, MacSween RN, Mills PR. (2000) Co-amoxiclav jaundice: clinical and histological features and HLA class II association. *Gut*. 47(5):717–720.
- O'Grady J, Joubert PH. (1997) *Phase I/II Clinical Drug Trials*. CRC Press, New York.
- O'Grady J, Linet OI. (1990) *Early Phase Drug Evaluations in Man*. CRC Press, New York.
- Office of Inspector General. (2000a) *Protecting Human Research Subjects: Status of Recommendations*. OEI-01-97-00197. Department of Health and Human Services, Washington, DC, June 2000. Available at: <http://oig.hhs.gov/oei/reports/oei-01-97-00197.pdf>
- Office of Inspector General. (2000b) *Recruiting Human Subjects: Pressures in Industry-Sponsored Clinical Research*. OEI-01-97-00195. Department of Health and Human Services, Washington, DC, June 2000. Available at: <http://oig.hhs.gov/oei/reports/oei-01-97-00195.pdf>
- Ozdemir V, Shear NH, Kalow W. (2001) What will be the role of pharmacokinetics in evaluating drug safety and minimizing adverse effects? *Drug Saf.* 24:75–85.
- Pharmaceutical Research and Manufacturers Association (PhRMA) (2016) *Prescription Medicines, Costs in Context*. Available at: <http://www.phrma.org/sites/default/files/pdf/prescription-medicines-costs-in-context-extended.pdf>.
- Piantadosi S. (1997) *Clinical Trials: A Methodologic Perspective*. John Wiley & Sons, Inc., New York.
- Reidenberg MM, Lowenthal DT. (1968) Adverse non-drug reactions. *N. Engl. J. Med.* 279:678–679.
- Robinson WT. (2008) Innovative early development regulatory approaches: expIND, expCTA, microdosing. *Clin. Pharmacol. Ther.* 83(2):358–360.
- Rolan PE. (1997) The assessment of pharmacokinetics in early phase drug evaluations. In: *Handbook of Phase I/II Clinical Drug Trials* (Grady JC and Joubert PH, Eds.). CRC Press, Boca Raton, FL, pp. 169–175.
- Rolan PE, Mercer AJ, Weatherly BC, Holdich T, Meire H, Peck RW, Ridout G, Posner J. (1994) Examination of some factors responsible for a food-induced increase in absorption of atovaquone. *Br. J. Clin. Pharmacol.* 37:13–20.
- Roy B, Chowdhury A, Kundu S, Santra A, Dey B, Chakaborty M, Majumder PP. (2001) Increased risk of antituberculosis drug-induced hepatotoxicity in individuals with glutathione S-transferase M1 'null' mutation. *J. Gastroenterol. Hepatol.* 16(9):1033–1037.
- Sallustio BC, Purdie YJ, Birkett DJ, Meffin PJ. (1989) Effect of renal dysfunction on the individual components of the acyl-glucuronide futile cycle. *J. Pharmacol. Exp. Ther.* 25:288–294.
- Schmidt MJ. (2001) Human safety in clinical research. *Appl. Clin. Trials*, July 2001, 40–50.
- Shamoo AE. (2000) Future challenges to human subject protection. *The Scientist*, June 26, 2000: 14(13):35. Available at: www.the-scientist.com/yr2000/jun/opin_000626.html (accessed April 7, 2016).
- Sharma SK, Balamurugan A, Saha PK, Pandey RM, Mehra NK. (2002) Evaluation of clinical and immunogenetic risk factors for the development of hepatotoxicity during antituberculosis treatment. *Am. J. Respir. Crit. Care Med.* 166(7):916–919.
- Sharp SM. (2001) The ethical foundations of clinical trials. *Appl. Clin. Trials*, April 2001, 48–54.
- Simon T, Becquemont L, Mary-Krause M, de Waziers I, Beaune P, Funck-Brentano C, Jaillon P. (2000) Combined glutathione-S-transferase M1 and T1 genetic polymorphism and tacrine hepatotoxicity. *Clin. Pharmacol. Ther.* 67(4):432–437.
- Spilker B. (1991) *Guide to Clinical Trials*. Raven Press, New York.
- Stone J. (2006) *Conducting Clinical Research*. Mountainside MD Press, Cumberland, MD.
- Washington N, Washington C, Wilson CG. (2001) *Physiological Pharmaceutics*. Taylor & Francis, London.
- Watanabe I, Tomita A, Shimizu M, Sugawara M, Yasumo H, Koishi R, Takahashi T, Miyoshi K, Nakamura K, Izumi T, Matsushita Y, Furukawa H, Haruyama H, Koga T. (2003) A study to survey susceptible genetic factors responsible for troglitazone-associated hepatotoxicity in Japanese patients with type 2 diabetes mellitus. *Clin. Pharmacol. Ther.* 73(5):435–455.
- Wechsler J. (2001) Clinical trial safety and oversight top policy agenda. *Appl. Clin. Trials*. 10:18–21.
- Wilke RA, Lin DW, Roden DM, Watkins PB, Flockhard D, Zineh I, Giacomini KM, Krauss RM. (2007) Identifying genetic risk factors for serious adverse drug reactions: current progress and challenges. *Nat. Rev. Drug Discov.* 6:904–916.
- Willman D. (2000) Quickened Pace of Drug Approvals by FDA Taking Toll. *San Jose Mercury News*, December 28, 2000.
- Wilson D. (1998) A Cautionary Tale for Human "Guinea Pigs". *News and Observer*, May 8, 1998:4E.
- Winek CL. (2000) History of excipient safety and toxicity. In: *Excipient Toxicity and Safety* (Werner ML and Kotkoskie LA, Eds.). Marcel Dekker, New York, pp. 59–72.
- World Medical Association. (2008) *Declaration of Helsinki: Ethical Principles for Medical Research Involving Human Subjects*. Adopted by the 18th WMA General Assembly, Helsinki, Finland, June 1964; Amended by the 29th WMA, 1975; 35th WMA, 1983; 41st WMA, 1989; 48th WMA, Somerset West, Republic of South Africa, October 1996; 52nd WMA, Edinburgh, Scotland, October 2000; 58th WMA, Copenhagen, Denmark, October 2007. Available at: <http://www.wma.net/en/30publications/10policies/b3/>

POSTMARKETING SAFETY EVALUATION: MONITORING, ASSESSING, AND REPORTING OF ADVERSE DRUG RESPONSES (ADRs)

Once a new drug is approved, it proceeds to market either with or without postmarketing evaluation requirements. While this represents the end of a long road, it is also the start of yet another. While careful work during development (both in animals and humans) serves to provide the tools to greatly reduce the potential safety issues around a new drug, it cannot totally eliminate them. One needs only to look at Table 29.1 to appreciate the history of market withdrawals due to safety issues in the modern era (1961–2001) or turn to Table 29.2 to verify that the problem is still present in the first decade of the twenty-first century, though with lower incidence of individuals severely affected in each case.

Tracking and continuing to evaluate the safety of a therapeutic agent once it is on the market are a complex task, and now the field of the discipline (and regulations) called pharmacovigilance (PV). Philosophically, PV represents a special case of the general risk assessment model (the steps of which are summarized in Table 29.3). The World Health Organization (WHO) established the precursor for these efforts as the Programme for International Drug Monitoring in response to the thalidomide disaster detected in 1961.

More explicitly, PV (or postmarket clinical drug safety) is the discipline dedicated to the collection, detection, assessment, monitoring, and prevention of adverse effects with pharmaceutical products. The etymological roots for the word “PV” are *pharmakon* (Greek for drug) and *vigilare* (Latin for to keep watch). As such, PV heavily focuses on adverse drug reactions (ADRs) which are defined as any response to a drug which is noxious and unintended, including lack of efficacy, which occurs at doses normally used for prophylaxis, diagnosis, or therapy of disease, or for the modification of physiological function. Medication errors such as overdose and misuse and abuse of a drug are also of interest because they may result in an ADR.

Information received from patients and healthcare providers, as well as other sources such as the medical literature, plays a critical role in providing the data necessary for PV to take place. In fact, in order to market or to test a pharmaceutical product in most countries, adverse event (AE) data received by the license holder (usually a pharmaceutical company) must be submitted to the local drug regulatory authority (FDA, 2005; Mann and Andrews, 2007; Strom et al., 2012).

For the WHO drug safety reporting system (now participated with by more than 134 countries), Figure 29.1 summarizes the growth in number of event reports through 2012. The minimum requirements that the WHO and partners agree should be present in any national PV:

1. A national PV center with designated staff (at least one fulltime), stable basic funding, clear mandates, well-defined structures and roles, and collaborating with the WHO Programme for International Drug Monitoring
2. The existence of a national spontaneous reporting system, with a national individual case safety report (ICSR) form, that is, an ADR reporting form
3. A national database or system for collating and managing ADR reports
4. A national ADR or PV advisory committee able to provide technical assistance on causality assessment, risk assessment, risk management, case investigation, and, where necessary, crisis management including crisis communication
5. A clear communication strategy for routine communication and crises communication

Manufacturers are legally required to collect, analyze, and report such data both nationally (by the Food and Drug

TABLE 29.1 Therapeutic Products Withdrawn from the Marketplace Due to Safety Reasons in the United Kingdom and/or the United States 1961–2001

| Drug (INN) | Trade Name | Therapeutic Class | Reason(s) for Withdrawal/ Suspension of Product License | Launch Date | Countries | License | |
|-------------------------------------|---------------|--|--|----------------|------------------------------------|---------------------------------------|--------------------|
| | | | | | | Suspension or Market Withdrawal | Years on Market |
| Acetylsalicylic acid (pediatric) | Aspirin | Analgesic | Reye's syndrome | 1899 | United Kingdom | 1986 | 87 |
| Alclofenac | Prinalgin | Nonsteroidal anti- inflammatory drug (NSAID) | Skin and renal reactions Mutagenic metabolite | 1972 | United Kingdom | 1979 | 8 |
| Alfaxalone | Althesin | Anesthetic | Allergic-type reaction | 1972 | United Kingdom | 1984 | 12 |
| Aminogluthethimide | Elipten | Anticonvulsant | Endocrinological Reintroduced—Cushing's syndrome | 1960 1978 | United States United States | 1966 | 6 |
| Aminopyrine | | Analgesic | Hematological | 1900 1900 | United States United Kingdom | 1970 1975 | 70 75 |
| Astemizole | Hismanal | Antihistamine | Drug–drug interaction | 1988 | United States | 1999 | 11 |
| Azari bine | Triazure | Antipsoriatic | Neuropsychiatric coagulation disorders | 1975 | United States | 1976 | 1 |
| Benoxaprofen | Opren | NSAID | Cholestatic jaundice | 1980 | United Kingdom | 1982 | 2 |
| Benziodarone | Amplivix | Uricosuric | Photosensitivity Hepatic | 1982 1962 | United States United Kingdom | 1982 1964 | <1 2 |
| Bithionol | Actamer | Coronary dilator | Dermatological | ? | United States | 1967 | ? |
| Bromfenac sodium | Duract | Anthelminthic | Liver damage | 1996 | United States | 1998 | 3 |
| Cervastatin sodium | Baycol | NSAID | Muscle weakening | 1999 | United States | 2001 | 2 |
| Chlormadinone | Normenon | Cholesterol lowerers Hormone | Animal carcinogenicity | 1965 1966 | United States United Kingdom | 1970 1970 | 5 4 |
| Cisapride | Propulsid | Heartburn | Cardiovascular irregularities | 1993 | United States | 2000 | |
| Clioquinol | Enterovioform | Antidiarrheal | Neuropsychiatric | 1930 | United Kingdom | 1981 | 51 |
| Centoxin | HA-1A | Hu-anti-lipid | Gram-negative septicemia | 1930 1990 | United States United Kingdom | 1973 1992 | 43 2 |
| Danthron | Dorbanex | A IgM monoclonal Laxative | Animal carcinogenicity Reintroduced with restrictions | 1959 1964 | United States United Kingdom | 1987 1987 | 28 23 |
| Desensitizing vaccines | Various | Vaccine | Allergic-type reactions | ? | United Kingdom | 1989 | |
| Diamthazole | Asterol | Antifungal | Neuropsychiatric | ? | United States | 1970 | 17 |
| Dihydrostreptomycin | | Antibiotic | Neuropsychiatric | ? | United States | 1970 | ? |

| | | | | | | | |
|-------------------------------|-------------------|--------------------------|---|------|----------------|------|-----------|
| Dinoprostone | Propress | Hormone | Uterine hypertonus and fetal distress | 1989 | United Kingdom | 1990 | <1 |
| Dipyrene | | Analgesic | Hematological | 1930 | United States | 1977 | 47 |
| | | | | 1930 | United Kingdom | 1977 | 47 |
| Dithiazanine | | Anthelminthic | Metabolic | ? | United States | 1964 | |
| Domperidone (injection) | Motilium | Antiemetic | Cardiovascular | 1984 | United Kingdom | 1986 | 2 |
| Doxylamine | Bendectin | Antihistamine | Risk of overdose | 1956 | United States | 1983 | 27 |
| Dicyclomine | Debendox | | Fear of teratogenicity—dysmorphogenicity? | 1957 | United Kingdom | 1983 | 26 |
| Encainide | Enkaid | Antiarrhythmic | Cardiovascular | 1987 | United States | 1991 | 4 |
| Factor VIII | Factorate | Coagulation factor | Excess mortality risk | 1972 | United Kingdom | 1986 | 14 |
| Fenclofenac | Flenac | NSAID | Manufacture problem | 1978 | United Kingdom | 1984 | 6 |
| Feprazone | Methrazone | NSAID | Risk of AIDS transmission | | | | |
| Flosequinan | Manoplax | Heart failure | Multiple—especially skin—reactions | 1976 | United Kingdom | 1984 | 8 |
| | | | Multiple | | United Kingdom | 1984 | |
| Flosequinan | Manoplax | Heart failure | Increased mortality | 1992 | United Kingdom | 1993 | 9 months |
| Grepafloxacin | Raxar | Quinolone antibiotic | Lack of long-term efficacy | 1997 | United States | 1999 | 23 months |
| Growth Hormone (natural) | Crescormon | Hormone | QT interval prolongation | 1970 | United States | 1985 | 15 |
| | | | Manufacture problem | | United Kingdom | | |
| | | | Creutzfeldt–Jakob disease transmission | 1970 | United States | 1985 | |
| Guanethidine | Ganda (high dose) | Antiglaucoma eye drops | Ophthalmological | 1977 | United Kingdom | 1986 | 9 |
| Ibufenac | Dytransin | NSAID | Hepatic | 1966 | United Kingdom | 1968 | 2 |
| Indomethacin-R | Osmosin form | NSAID | Multiple gastrointestinal—36 fatal small intestine perforations | 1982 | United Kingdom | 1983 | 9 months |
| | | | | | United Kingdom | 1983 | |
| Indoprofen | Flosint | NSAID | Gastrointestinal | 1982 | United Kingdom | 1983 | 1 |
| | | | carcinogenicity | | United Kingdom | 1984 | ?2 |
| Iodinated casein strophanthin | Coratose | Anorexiant | Metabolic | ? | ? | 1964 | ? |
| — | | | | | | | |
| Mebanazine | Lotronex | Irritable bowel syndrome | Ischemic colitis | 2000 | United States | 2000 | 9 months |
| | Actomol | Antidepressant | Hepatic drug interactions | 1963 | United Kingdom | 1975 | 12 |
| Me-gestrol acetate | Volidan 21 | Hormone | Carcinogenicity | 1963 | United Kingdom | 1970 | 7 |
| | | | | | United Kingdom | 1969 | |
| | | | | | | 1975 | |

(Continued)

TABLE 29.1 (Continued)

| Drug (INN) | Trade Name | Therapeutic Class | Reason(s) for Withdrawal/ Suspension of Product License | Launch Date | Countries | License | |
|--|----------------------------------|-----------------------------|---|-----------------------|---|---------------------------------------|--------------------|
| | | | | | | Suspension or Market Withdrawal | Years on Market |
| Methandrostenolone Methapyrilene | Dianabol | Hormone H1 antihistamine | Endocrinological Carcinogenicity | 1960 1947 1950 | United States United States United Kingdom | 1982 1979 1979 | 12 31 29 |
| Metipranolol | Glauline | Antiglaucoma eye drops | Ophthalmological—uveitis (high dose) | ? | United Kingdom | 1990 | ? |
| Metofoline | Versidyne | Analgesic | Low-dose preparation | ? | | 1991 | |
| Mibefradil | Posicor | Calcium channel blocker | Experimental toxicity Lethal drug interaction (inhibited liver enzymes) | 1996 | United States | 1965 1998 | ? 3 |
| Mumps vaccine | Priorix | Vaccine | Neuropsychiatric | 1988 | United Kingdom | 1992 | 4 |
| Urabe AM9 strain Neomycin (injection) | | Antibiotic | Meningitis Misuse | 1988 ? | United States United States | 1992 1989 | 4 ? |
| Nialamide | Niamid | Antidepressant (MAOI) | Irrigation of open wounds Hepatic | 1959 | United Kingdom | 1978 | 19 |
| Nomifensine | Merital | Antidepressant | Drug interactions Hemolytic anemia | 1959 1977 | United States United Kingdom | 1974 1986 | 15 9 |
| Oxyphenbutazone | Tanderil | NSAID | Hepatotoxicity—fatal hepatitis Hematological Multiple | 1960 1962 | United States United Kingdom | 1985 1984 | 25 22 |
| Oxyphenisatin | Veripaque | Laxative | Hepatic | 1955 | United Kingdom | 1978 | 23 |
| Perhexiline maleate | Pexid | | Hepatic damage Peripheral neuropathy Renal carcinogenicity | 1957 1975 <1900 | United States United Kingdom United Kingdom | 1972 1985 1980 | 15 10 80 |
| Phenacetin | | Analgesic | | | | | |
| Fen-phen | Fenfluramine/ dexfenfluramine | Diet aid | Heart valve abnormality | 1900 1973 | United States United States | 1983 1997 | 83 24 |
| Phenformin | Insoral Dibotin | Antidiabetic | Metabolic | 1959 | United Kingdom | 1982 | 23 |
| Phenoxypiprazine | Drazine | Antidepressant (MAOI) | Hepatic Drug interactions | 1959 1961 | United States United Kingdom | 1977 1966 | 18 5 |
| Phenylpropanolamine | PPA | CTC ingredient | Hemorrhagic stroke | 1936 | United States | 2001 | 65 |

TABLE 29.1 (Continued)

| Drug (INN) | Trade Name | Therapeutic Class | Reason(s) for Withdrawal/ Suspension of Product License | Launch Date | Countries | License Suspension or Market Withdrawal | Years on Market |
|------------|--------------------|----------------------------------|---|----------------|------------------------------------|--|--------------------|
| Tryptophan | Pacitron | Low-protein diet Reintroduced | Eosinophilic myalgia syndrome | 1974 | United States | 1989 | 15 |
| Vitamin E | Optimax E-Ferol | Optics monitor Vitamin | Hematological Hepatic Renal | 1983 | United States | 1990 1984 | 1 |
| Zomepirac | Zomax | NSAID | Allergic-type reactions—fatal anaphylaxis | 1980 1981 | United States United Kingdom | 1983 1983 | 3 2 |
| Zimeldine | Zelmid | Antidepressant | Hepatotoxicity Neurological—peripheral neuropathy, Guillain-Barré syndrome | 1982 | United Kingdom | 1983 | 1 |

TABLE 29.2 List of Drugs Withdrawn Since 1990

| Year | Drug | Indication/Class | Causative Side Effect |
|------------|---|--------------------------------------|--|
| 1991 | Enkaid (4 years on market) | Antiarrhythmic | Cardiovascular (sudden cardiac death) |
| 1992 | Temafloxacin | Antibiotic | Blood and kidney |
| 1997 | Fenfluramine/dexfenfluramine (combo used since 1984) (24 years on market) | Diet pill | Heart valve abnormalities |
| 1998 | Seldane (terfenadine) | Antihistamine | Ventricular arrhythmias |
| | Posicor (mibefradil) (1 year on market) (inhibited liver enzymes) | Ca ²⁺ channel blocker | Lethal drug interactions |
| | Duract (bronfemic sodium) (early approval warnings of ↑ liver enzymes) | Pain relief | Liver damage |
| 1999 | Trovan (use severely restricted) | Antibiotic | Liver/kidney damage |
| 2000 | Hismanal | Antihistamine | Drug–drug interactions |
| | Rotashield | Rotavirus vaccine | Bowel obstruction |
| | Rezulin (approved December 1996) | Type 2 diabetes | Liver damage |
| 2001 | Propulsid | Heartburn | Cardiovascular irregularities/death |
| | Lotronex ^a | Irritable bowel syndrome | Ischemic colitis/death |
| | Phenylpropanolamine (ppa) | Otc ingredient | Hemorrhagic stroke |
| | Baychlor (baycol) | Cholesterol reducing (statin) | Muscular weakness/death |
| | Alatrofloxacin | Antibiotic | Liver toxicity |
| 2002, 2003 | None | None | None |
| 2004 | Serzone | Antidepressant | Liver failure and injury |
| | Vioxx ^b | Arthritis (cox-2 inhibitor) | Heart attack/cardiovascular (thrombosis) |
| 2005 | Tysabri ^a | Ms (multiple sclerosis) | Pml (progressive multifocal leukoencephalopathy) |
| | Bextra | Arthritis (cox-2 inhibitor) | Cardiovascular safety, skin reaction (sometimes fatal) |
| 2006 | Dolophine (methadone hydrochloride) | Treatment of moderate to severe pain | Respiratory depression and cardiac arrhythmias |
| 2007 | Zelnorm | Constipation | Cardiovascular safety |
| | Permax | Parkinson's disease | Heart valve damage |
| 2008 | Trasylol | Bleeding | Increased risk of death |
| | Acomplia | Weight loss | Severe depression/suicide |
| 2009 | Raptiva | Psoriasis | Pml |
| 2010 | Meridia | Weight loss | Cardiovascular safety |
| | Avandia | Diabetes | Cardiovascular safety |
| | Darvon and darvocet | Pain | Addiction, cardiovascular safety |
| | Mylotarg | Aml | Veno-occlusive disease, risk for death |
| 2011 | Vivaglobin | Primary immune deficiencies | Thrombolytic adverse extents (cv) |

^aLimited return to market in 2003. The new labels will also include a detailed warning on the “well-described, serious, and potentially life-threatening GI bleeding,” as well as GI ulceration and perforation. These products should be used with “extreme caution” in patients with a prior history of ulcer disease or GI bleeding, the warnings section will state, and although the risk of these problems increases with duration of use, even short-term therapy is not without risk.

^bThe possibility of skin reactions is also part of the warnings section: “NSAIDs including [brand name] can cause serious skin adverse events including exfoliative dermatitis, Stevens–Johnson syndrome, and toxic epidermal necrolysis.” These reactions played a big part in the FDA’s decision to ask Pfizer to suspend Bextra—the product had “no added advantage” and a “special risk” of serious skin reactions, the agency said.

TABLE 29.3 A Special Case of the General Risk Assessment Model in a Regulated Environment

| |
|--|
| Identify potential hazards (signals) |
| Collect signal data (ADRs) |
| Classify/categorize/assess signals (assign causality) |
| Code and analyze signals (ADR reports) |
| Report population ADR data |
| Perform risk/benefit assessment and then risk management |
| A similar approach used for medical devices |

Administration (FDA) in the United States and by equivalent organizations in other countries—but here the emphasis will be on the US situation) and internationally (by WHO). There are regulatory reporting systems (that is, where the reports go directly to government agencies) and organizational reporting systems (organized around method of distribution—such as hospital pharmacies (Hunziker et al., 1977; ASHP, 1995) or by product type, such as radiopharmaceuticals). Poison control centers also monitor

CASE STUDY: THALIDOMIDE

Thalidomide, α -N-[phthalimido] glutarimide, was introduced into the market in 1957 by the Chemie Grünenthal Company in Germany. A derivative of glutamic acid, it exists in humans in two concurrent, interchangeable, and optical isomeric forms (enantiomers), *S*(–) and *R*(+). Despite research showing its safety, the drug caused horrific birth defects when taken by pregnant women after it was approved for market use. The detrimental effects of thalidomide led Congress to pass the Kefauver–Harris amendment to the FD&C Act in 1962 which required far greater PV. The amendment also required proof of safety before any clinical trials (the current IND process) can be initiated.

The nonclinical testing for thalidomide included overdose experiments in rodents, which produced neither death nor congenital defects in offspring. Marketing therefore asserted that thalidomide had no toxic effects in humans. They advertised it as a sedative, hypnotic, and antiemetic, with particular emphasis on the reduction of symptoms related to morning sickness. By 1961 thalidomide was marketed in 46 countries under a variety of names. The increase in distribution was followed by an increase in severe, and somewhat unique, deformities in newborn children, though the causal relationship was far from obvious.

In 1961 a German doctor, Widukind Lenz, cast suspicion on thalidomide as the cause of the deformities, which was quickly and independently confirmed that same year by an Australian doctor, William McBride. Withdrawal of the drug from the market began in the late 1961. This was the first instance of a drug confirmed to have teratogenic effects. Previously, researchers and scientists assumed that the placenta protected the fetus from this type of harm entirely. Unfortunately, the removal of the drug from the market was slow. Children were born with deformities due to its use even 2 years after the recall began. By current estimates, thalidomide caused over 10,000 children to be born with birth defects.

The human fetus proved very vulnerable to the effects of thalidomide. Multiple studies indicate that exposure to a single 50 mg tablet raises the likelihood of teratogenic effects as high as 50%. There is, however, a limited time window in which thalidomide must be administered in order to cause teratogenic effects: between 20 and 36 days postfertilization or 34 and 50 days since the last menstrual period. Before this window, exposure to thalidomide generally causes miscarriage and afterwards causes no morphological effects. In contrast, rats (the original animal model tested) are relatively insensitive to the developmental effects of thalidomide. However, rabbits and especially macaques are sensitive models for these effects.

Morphological effects induced by fetal exposure to thalidomide include defects to the limbs, eyes, ears, internal organs, nerves, and CNS. Limb defects are the most common, occurring in over 90% of the cases. Effects are generally reductive in nature and symmetrical. Deformities ranged from absence of digits to complete absence of limbs (amelia). Phocomelia is the classic image of the thalidomide tragedy, a shortening of the long bones which can leave “flippers.” Ocular effects include cataracts, microphthalmos, anophthalmos, and colobomas. They usually occur in only one eye, though the other eye may be affected independently. Ocular effects generally occur simultaneously with aural effects and changes to facial musculature. Aural effects are generally symmetrical and include anotia and microtia. Internal organs which are affected include the heart, kidneys, genitals, and bowels. Many of the defects in internal organs do not develop until later in life, and therefore incidence is unknown as many patients die. CNS effects can include facial palsies, cranial nerve conduction problems, and later in life autism and epilepsy.

ADR cases and rates (Chyka, 1999; Chyka and McEommon, 2000). The regulatory systems for such PV in the United States are MedWatch (for human drugs), vaccine adverse event reporting system (VAERS) (for human vaccines) (Niu et al., 1998, 1999; Varrinchio, 1998), and the FDA–CVM system (for veterinary drugs) (Bukowski and Wartenberg, 1996; Keller et al., 1998). One key difference of the US systems from those in other countries is that they are (and remain) voluntary (largely a reflection of the primarily private—that is, not national government—healthcare system in the United States). The US regulatory requirement for PV was promulgated in March of 2005, *FDA Guidance for Industry: Good Pharmacovigilance Practices and Pharmacoepidemiologic Assessment* (part of PDUFA IIIa) (FDA, 2005), and then continued as new

technologies are added to therapeutic medicines and new challenges (and surprises) arise:

- 1993—Tambocor (flecainide) as signal case for class IC antiarrhythmic drugs (~50,000 deaths due to cardiac arrest caused by hypokalemia and *torsades de pointes*) induced by these drugs
- Principal cause of International Conference on Harmonization (ICH) preclinical safety pharmacology testing requirements for drugs
 - Revealed cardiac arrhythmia suppression trial (CAST) (Moor, 1995)

There is no compulsion for physicians, hospitals, or individuals to report AEs to either the manufacturer or

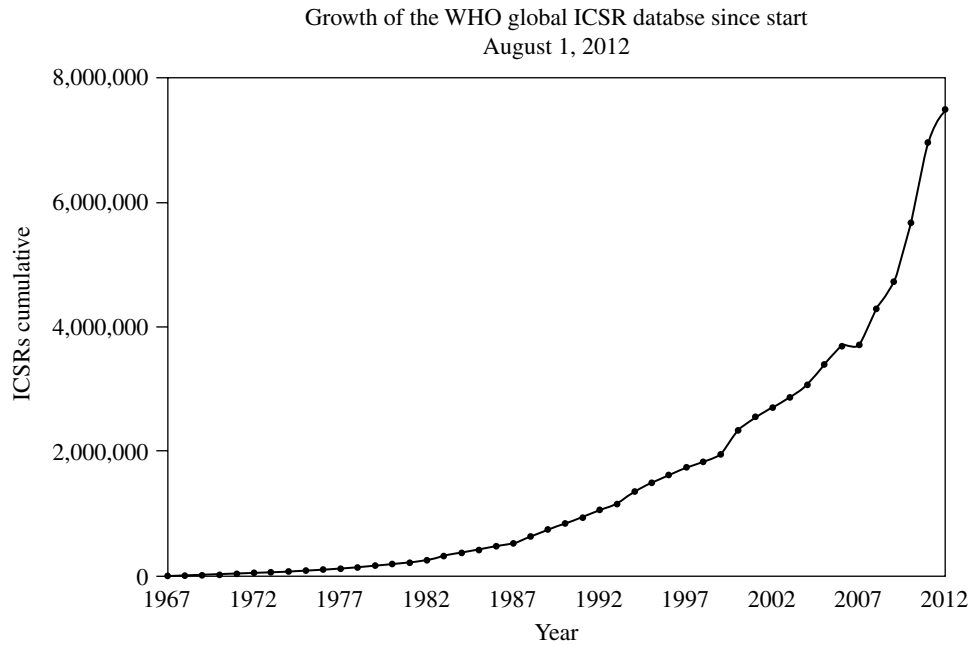


FIGURE 29.1 Global drug safety reports.

the government (though the marketing companies for a therapeutic are required to periodically summarize all reports of AEs that they know of on to the federal government). It is widely held that it is this voluntary (or “spontaneous”) aspect which limits the effectiveness of the US systems (Piazza-Hepp and Kennedy, 1995; Kennedy and Goldman, 1997; Goldman, 1998; Sharpe, 1998; White and Love, 1998; Brewer and Colditz, 1999). Studies have identified factors which influence (and limit) physician use of such systems (LaCalamita, 1995; Figueras et al., 1999), and newly marketed drugs are subject to a higher rate of underreporting of ADRs than are established drugs (Martin et al., 1998). It should be kept in mind that what we are considering here are adverse effects caused by the use of a drug as intended and not by a medication error. Medication errors are at least as serious a problem (and complex an issue) as ADRs (Antonow et al., 2000) but are beyond the scope of this volume. Both companies and the regulatory agencies must collect reports of AEs, evaluate them, and then decide on a correct course of action (ranging from doing nothing through improving labeling then on to restricting access and/or requiring ongoing or increased medical surveillance of patients and then to withdrawing the drug from the market).

For while new medications can save or mostly improve lives, once on the market side effects—harmful ones—are virtually certain to occur and must be promptly recognized.

ICH has sought to harmonize global PV efforts by promulgating a series of guidances:

- ICH E1—*The Extent of Population Exposure to Assess Clinical Safety for Drugs Intended for Long-Term Treatment of Non-Life Threatening Conditions*

- ICH E2C (R1)—*Clinical Safety Data Management: Periodic Safety Update for Marketed Drugs*
- ICH E2C (R2)—*Period Benefit–Risk Evaluation Report*
- ICH E2D—*Post-Approval Safety Data Management: Definitions and Standards for Expedited Reporting*
- ICH E2E—*Pharmacovigilance Planning*
- ICH E2F—*Development Safety Update Report*

Actual regulatory authority comes with implementation by regional authorities:

- European Union—adopted by CHMP, December 1, 2004, issued as CPMP/ICH/571603
- MHLW—Adopted September 16, 2005, PFSB/ELD Notification No. 0916001 and PFSB/SD Notification No. 0916001
- FDA—Published in the Federal Register, April 1, 2005, Vol. 70, No. 62, p. 16827–16828

29.1 CAUSES OF SAFETY WITHDRAWALS

It would be comforting to be able to state that the causes of postmarketing withdrawals from drugs were substantially different from those of the failure of drugs in clinical trials. While the last few years (refer back to Table 29.2) are seemingly somewhat different from those in the past (carcinogenic is no longer necessarily a big marketed drugs problem), the historic causes for the modern era (the last 40 years) are lead off by hepatic toxicity—also the primary cause for safety faced failure in early clinical trials (see Table 29.4).

TABLE 29.4 Characteristics of Drug Safety Withdrawals (1960–August 2001)

| | Drugs | % of Total |
|--|-------|------------|
| A. Most common classes | | |
| NSAIDs | 16 | 13 |
| Nonnarcotic analgesics | 10 | 8 |
| Antidepressants | 9 | 7 |
| Vasodilators | 7 | 6 |
| Anorexiant | 5 | 4 |
| CNS stimulants | 5 | 4 |
| Barbiturates | 5 | 4 |
| Anesthetics | 4 | 3 |
| Antihistamines | 4 | 3 |
| Antibiotics | 3 | 2 |
| B. Most common causes of withdrawal | | |
| Hepatic toxicity | | 26 |
| Hematologic toxicity (bone marrow suppression) | | 10 |
| Cardiovascular toxicity | | 6 |
| Carcinogenicity | | 6 |
| Renal toxicity | | 5 |
| Drug interactions | | 4 |
| Neurotoxicity | | 4 |
| Behavioral effects | | 4 |
| Abuse potential | | 4 |

While the proximal wake-up call was Tambocor in 1993, Fung et al. (2001) have done an extensive assembly and analysis of safety withdrawal data through 1999, and Ajayi et al. (2000) have also analyzed factors increasing likelihood of safety problems. Table 29.4 presents this author's extension of their work through of the time of this writing (late in the third quarter of 2001), which changes the results but a little. It should be noted that the rank order of these two lists is different than the rank orders based on the numbers of AEs (see Holland and DeGruz, 1997). AEs can have a wide range of causes which may not even be done to unanticipated effects of a drug but due to a medication error or something as mundane as the discrepancies between doses recommended in *The Physician's Desk Reference* and those recommended or reported in the medical literature (Cohen, 2001).

In 2005, Schuster et al. published an analysis that shared that cardiovascular toxicity was accounting for 40% of current (2000–2005) safety withdrawals and hepatotoxicity for 27% (leaving all others to share 33% of the total) (Schuster et al., 2005). Failure to identify these largely predictable causes of failures in new therapeutic entities largely reflects both a continuing lack of recognition of the actual patient populations utilizing drugs with their existing pathophysiological characteristics (Table 29.5) and the limitations of the currently employed clinical trial scheme (Table 29.6). While such efforts as mandatory assessment of safety

TABLE 29.5 Factors That Increase Patient Risk for Adverse Drug Interactions

| Factor | Group/Disorder |
|---|---|
| Age | Neonates, elderly |
| Gender | Women |
| Genetic phenotype | Slow metabolizers |
| Chronic disease | Moderate/severe renal or hepatic impairment, CHF, cirrhosis |
| Acute illness | Pneumonia, influenza |
| Metabolic disturbances | Hypothyroidism, hypoxia |
| Multiple drug use | Elderly, HIV patients |
| Multiple prescribing physicians | Elderly |
| Use of drugs with a low therapeutic index | |
| Use of drugs that are enzyme inhibitors or inducers | |

TABLE 29.6 Limitations of the FDA's Current Clinical Trials

| | |
|---------------|--|
| 1. Too few | Prior to approval, most drugs are administered to 2000–3000 patients. (To obtain an 80% probability of detecting an adverse drug event that occurs in one out of every 10,000 recipients, 16,000 patients must receive the drug) |
| 2. Too simple | Premarketing trials often exclude patients with complicated medical histories or medication regimens. It is easier to demonstrate efficacy without including these complex patients |
| 3. Too median | Most premarketing trials exclude patient populations such as pediatric, geriatric, and lactating and pregnant patients |
| 4. Too narrow | Premarketing trials are generally intended to investigate a drug for a single indication. After release to the market, the drug may be used to treat other conditions in different populations with varying medical histories |
| 5. Too brief | Adverse drug events that occur only with chronic use will not be detected in the relatively short clinical trial |

Source: Data from Rogers (1987, pp. 915–920).

pharmacology features will serve to improve the situation, for the foreseeable future it remains vital to ensure that our PV systems identify problems as soon as possible.

29.2 REGULATORY REQUIREMENTS

A “case” is a basic unit of drug safety surveillance. It is used to assure, to the greatest extent possible, the safety of approved drug products that are still in use. The basic unit of all postmarketing safety submissions is the adverse

drug experience “case,” which is an individual adverse drug experience.

FDA has explicit requirements for reporting of AE cases for drugs. A postmarketing adverse drug experience source can be categorized into the several sample categories: clinical trial, nonclinical trial/regulatory authority, nonclinical trial/literature, and nonclinical trial/all other. This chapter only deals with spontaneous experiences—nonclinical trial adverse drug experiences reported to the industry any time after a marketed drug product achieves marketing approval from FDA (Adams et al., 1997).

When one looks at a typical case folder, he or she notices that certain types of information are on the folder:

- The outside of a folder is identified by a (alpha) numeric code.
- There is an “initial” report.
- There is either at least one letter requesting additional information regarding the initial report or documentation reflecting the failed attempts to obtain additional information.
- There is at least one “follow-up” report.
- The spontaneous report event is categorized as serious or nonserious, expected or unexpected.
- The source is either literature, regulatory authority, or spontaneous.
- There is at least one MedWatch form or Council for International Organization of Medical Sciences (CIOMS) I form for each report.

Everything in a given case folder is present because of an FDA regulation requirement or a related company-written standard operating procedure.

The number on the outside of the case is required to be numeric or alphanumeric as opposed to the name of the patient. Patient names are not permitted to be publicly disclosed in the context of a MedWatch report per 21 CFR 21.63(f). The initial report is the first reported information received by the company about an individuals’ adverse drug experience. There must be a “prompt” attempt to obtain follow-up information about each initial report. The attempt(s) are made per company’s written procedures. If the written safety procedures are not followed, the safety reports are not appropriately submitted, or the safety records are not appropriately kept, the FDA has the authority under section 80 of Part 315 to withdraw the market NDA. The follow-up report is the format for submitting additional information about an experience. Each case regards only one individual unless the experience is both temporally and clinically unrelated to a second event experienced by the same person taking the same drug product.

Table 29.7 summarizes FDA reporting requirements of spontaneous reports in terms of how the case event is first

TABLE 29.7 How a Spontaneous Drug Case is First Submitted to the Food and Drug Administration

| Case Source/Case Type | Report Submitted |
|---|---------------------------|
| Foreign literature/not both serious and unexpected | Not 15 days, not periodic |
| Foreign literature/serious and unexpected | 15 days |
| US consumer/not both serious and unexpected | Periodic |
| US consumer/serious, unexpected | 15 days |
| Foreign consumer/not both serious and unexpected | Not 15 days, not periodic |
| Foreign consumer/serious, unexpected | 15 days |
| FDA, initial/serious and unexpected | 15 days |
| FDA, initial/serious and unexpected | Periodic, not 15 days |
| International regulatory authority/serious and unexpected | 15 days |
| International regulatory authority/not serious and unexpected | Not 15 days, not periodic |

submitted to the agency. The definitions of serious, unexpected, etc., are in 21 CFR 314.80(a) (CFR, 1994; FDA, 2005).

29.2.1 The 15-Day Report versus the US Periodic Report

Postmarketing adverse drug experiences are reported to a drug company by the public via regulatory authorities, literature, attorneys, consumers, and health professionals. Sometimes a company receives a report of an adverse experience someone had after taking its drug product not from the public but from FDA because instead of submitting the report to the company, the report was submitted directly to FDA. When FDA sends the applicant an initial MedWatch report, the information does not have to be resubmitted to FDA in an initial 15-day report if the information is serious and unexpected. This is because FDA already has knowledge of the report. However, MedWatch and its information are incorporated into the next periodic report of the product. Follow-up to such an FDA MedWatch, obtained from a non-FDA source, would be submitted as a follow-up, expedited 15-day report (and should reference the source of the initial report). When FDA sends the applicant an initial MedWatch that is not both serious and unexpected, the applicant incorporates the information into the next periodic report, under the normal procedure for submitting follow-up information.

If an initial 15-day report was submitted, and the first follow-up information reflects that the event is no longer classified as a 15-day report (never was serious and unexpected), the first follow-up report describes the change in the report classification but is a (first) follow-up 15-day report. Subsequent additional information is not submitted in the form of a 15-day report.

A periodic report contains certain information, such as the event terms submitted during the period, the dates that events of the period were submitted, an event term count by body system, and labeling changes made due to the period's adverse experiences. In addition to (and prior to) being incorporated into a periodic report, 15-day reports are submitted within 15 calendar days of the date the applicant received the data. All 15-day reports contain serious, unexpected events. Non-15-day reports are submitted periodically in FDA periodic reports.

If on a given day a serious, unexpected domestic report is received, it is submitted first on an FDA Form 3500A, within 15 calendar days of receipt, via the 15-day report and subsequently is incorporated (not in the form of an FDA Form 3500A) into a periodic report. If a report is received, that is, domestic but not both serious and unexpected, it is not submitted in a 15-day report but rather in the US periodic report. A US periodic report is submitted quarterly for the first 3 years after the date the product was approved by the FDA for marketing (21 CFR 314.80). However, the March 2001 FDA guidance allows an applicant to request a 21 CFR 314.90 waiver of the US periodic report reporting period and base the report not on the date of FDA marketing approval but instead on the international birthdate (the first date the product was approved in the international community). The request for such a waiver should be submitted to the Director, Office of Postmarketing Drug Risk Assessment, CDER, FDA, 5600 Fishers Lane, HFD-400, Rockville, MD 20857. The request should include the product's name, the date of FDA marketing approval, and the product's approved application number. In addition, an applicant may request a 21 CFR 314.90 waiver of the 21 CFR 314.80©(2)(ii) format of the periodic report submitted. If the waiver is granted, the ICH E2C Periodic Safety Update Report format may be used provided that the content of the Section 80 ©(2)(ii) information that is not in the body of the ICH E2C periodic report is found in appendices, that is, certain reports from consumers that are not in the body of the ICH E2C periodic report submission. 21 CFR 314.90 states that, among other things, the applicant may request FDA to waive any of the postmarketing requirements under 21 CFR 314.80.

ICH E2C and the FDA March 2001 draft guidance *Postmarketing Safety Reporting for Human Drug and Biological Products Including Vaccines* are available at <http://www.fda.gov/RegulatoryInformation/Guidances/default.htm#browse>; "case" requirements are accessible in 21 CFR 314.80.

The major change over the past few years has been a significant attempt to harmonize regulations under the aegis of the ICH, and this is covered in some depth later. The latest response to the ICH in its three participating regions (Europe, the United States, and Japan) is also described together with an update on the current UK regulations. The ICH

potentially offers real advantages to the pharmaceutical industry, but the process takes time, and countries have adopted and implemented the guidance in slightly different ways and at different times. National regulations and guidelines are therefore bound to change in the near future as each country embraces the ICH.

The ICH of Technical Requirements for Registration of Pharmaceuticals for Human Use has brought together as equal partners the regulatory authorities of Europe, Japan, and the United States and experts from the pharmaceutical industry in these regions to discuss scientific and technical aspects of product registration. The WHO, European Free Trade Area (EFTA), and Canada are observers, and the International Federation of Pharmaceutical Manufacturers Association (IFPMA) ensures contact with the research-based industry outside the ICH regions.

The aim of the ICH is to achieve greater harmonization in the interpretation and application of technical guidelines and requirements for product registration and reduce or eliminate duplicate testing. This should result in better use of resources and eliminate unnecessary delay in the global development and availability of new medicines while maintaining safety guards on quality, safety, and efficacy.

There are four broad topic areas within the ICH:

- S—Safety (animal toxicology and pharmacology)
- Q—Quality (pharmaceutical and analytical)
- E—Efficacy (clinical)
- M—Multidisciplinary topics

Timely, complete reporting of ADRs and medical device problems is essential to an effective national system of postmarketing surveillance (PMS). Pharmaceutical manufacturers are required by federal regulations to report all ADRs of which they are aware to the FDA. However, many healthcare professionals do not think to report AEs either to the manufacturers or to FDA. To encourage and facilitate the reporting of serious AEs, FDA launched the MedWatch reporting program in June 1993. The MedWatch reporting form is used by healthcare professionals to voluntarily report ADRs and other problems with all FDA-regulated products used in medical therapy (drugs, biologics, medical devices, and special nutritional agents). The reporting of ADRs associated with vaccine products is the only exception, since reporting of those ADRs is mandatory. The form used for vaccines is the joint FDA/Centers for Disease Control and Prevention (CDC) VAERS form. For drugs and therapeutic biologics, the MedWatch (3500A) form replaces the 1639 reporting form.

FDA does not want reports on every AE observed; that would not be practical for reporters or FDA because of the sheer volume of AE reports already being sent to the agency each year (about 130,000 in 1994). While 80–85% of these

reports are submitted by the manufacturer, 10–15% are received by MedWatch directly from physicians, pharmacists, other healthcare professionals, and consumers. MedWatch encourages reporters to be selective by limiting their reports to events for which the outcome was serious. This enables FDA to focus on those events with potentially the largest public health impact. Reporters are encouraged to fill out the reporting form as completely and accurately as possible.

From 1978 through 1990, the CDC and the FDA divided the responsibility for PMS of vaccines in the United States. The FDA received reports of AEs after vaccines were administered in the private sector; events occurring after the administration of vaccines purchased with public funds were reported to the monitoring system for AEs following immunization.

The monitoring system was a stimulated passive surveillance system. In other words, when vaccines purchased with federal funds were administered in the public sector, “important information” forms were given to recipients or their parents or guardians instructing them to report any illnesses requiring medical attention that occurred within four weeks of vaccination. System coordinators at each immunization project/grantee site and the state health department completed standardized forms that were reviewed for consistency and completeness and then forwarded to the CDC for data entry and analysis.

In response to the National Childhood Vaccine Injury Act of 1988, which required health workers to report vaccine AEs, the CDC and the FDA collaborated in 1990 to implement the VAERS to monitor the safety of vaccines in both sectors. Healthcare professionals and parents/caretakers are *encouraged* to report all clinically significant vaccine AEs. Narrative diagnostic reports are reviewed and assigned standard codes using Coding Symbols for a Thesaurus of Adverse Reaction Terms. The source of the vaccines (public vs. private provider) is recorded on the form.

The WHO system, created in response to the thalidomide disaster, seeks to capture adverse worldwide AEs and identifies problems (WHO, 1975, 2010; Olsson, 1998). It is proposed that all such gathered reports should first be analyzed for mortality effects and trends (Rose and Elnis, 2000) as such would identify the most critical trends and be easiest to evaluate.

29.3 MANAGEMENT OF ADR AND ADE DATA

In monitoring the safety of products, pharmaceutical companies need to comply with worldwide regulations as well as the primary requirement of helping doctors to prescribe safely. It is not intended to provide a comprehensive review in this chapter but to provide an insight into the methods of managing ADR data.

29.3.1 Sources of Data

There can be an enormous variation in the nature and quality of data depending upon the source, and this must be considered when the data is processed, computerized, and analyzed. Safety data may come from any of the sources described later.

29.3.2 Clinical Trials

In Phase I studies, good documentation and additional investigations should be a standard practice. Serious reactions are pretty unusual in these studies, which will detect only very common ADRs, in particular those that are pharmacologically mediated (e.g., bradycardia with beta-adrenergic receptor antagonists).

Good documentation and follow-up should be possible in Phase II studies, but rare reactions will not be identified due to the small numbers of patients involved. The larger numbers in Phase III trials can pose problems, but these can be minimized by careful choice of investigators, good case report form design, and procedures for follow-up. Phase IV studies are designed to test the efficacy and safety of the drug in clinical practice and often share the same constraints in patient numbers as premarketing trials.

29.3.3 Postmarketing Surveillance Studies

Any surveillance of safety of a drug after marketing is PMS (now often referred to as a postauthorization safety study). In practice the distinction between Phase IV studies and PMS is blurred (e.g., German drug experience studies).

29.3.4 Spontaneous Reports

Spontaneous reports are the most effective means of identifying rare, serious adverse reactions (usually idiosyncratic or Type B) after marketing despite the underreporting that exists. Spontaneous reports are an unsolicited communication to a company, regulatory authority, or other organization that describes an AE in a patient given one or more medical products. These reports do not originate from a study or from any organized data collection scheme. Unless indicated otherwise by the reporter, all spontaneous AEs are assumed to be possible ADRs. The quality and completeness of spontaneous reports are often inadequate. Pharmaceutical companies or regulatory authorities can only achieve good case documentation through effective data collection, detailed follow-up, and use of field workers for complex cases. The quality of spontaneous reports also varies from country to country. Some countries do not have a regulatory reporting form for ADRs. There are differences among countries in publicity of drug safety issues, and drug regulations differ regarding the format, content, and submission time frames for ADR reporting.

Reports received by companies via regulatory authorities are often edited and poorly documented, but they cannot be ignored and should be handled alongside reports received directly. The FDA implemented the Medical Products Reporting Program (MedWatch) in 1993, which encourages healthcare providers to regard reporting as a fundamental professional and public health responsibility and submit serious AE reports directly to the FDA on the FDA 3500A form. The FDA forwards these reports to the manufacturer, who is obliged to follow up with those reporting such events and submit any relevant information obtained to the FDA and other regulatory agencies worldwide as required. It is currently proposed that regulatory agencies (FDA) should take a more direct hand in these activities (Snidermann, 2000).

29.3.5 Literature

The publication of case reports in medical and scientific journals is an important primary source of information on ADRs. Many ADRs are noted in medical and scientific journals before they become well known. For example, the association of thalidomide with birth defects was first noted in a letter to the *Lancet* in 1961. The quality of ADR reports in the published literature can be variable and has been the subject of much criticism and correspondence, though guidelines have been promulgated for these (Jones, 1982).

Despite the anecdotal nature and sometimes poor documentation, publication of case reports in journals remains one of the most useful primary sources of information on ADRs. ADR reports in the literature can be identified in several different ways. Prepublication manuscripts describing a spontaneous case report or an event from a clinical trial are sometimes provided by authors to the manufacturer of the drug and the regulatory authority in that country. Pharmaceutical companies are required to be aware of the literature as to the safety of their approved therapeutic products and are assumed (by law) to be cognizant of such.

29.3.6 Searching for ADRs in the Literature

With the increasing number of scientific and biomedical journals, there are more sources of ADR data on many drugs. Conversely, for some drugs, particularly those recently marketed, there is a scarcity of clinical publications, and frequently there is an inadequate account of the adverse reaction profile. Searching for ADRs in the literature may be assisted by online databases such as MEDLINE (Index Medicus), EMBASE (Excerpta Medica), and secondary sources such as SEDBASE (Meyler's side effects of drugs) and ADIS online services such as REACTIONS. Many journals contain relevant information, but some specific ADR-related journals may assist in the search for information. Increasingly, the use of high-capacity storage systems such as compact disks (CD-ROM) has led to stand-alone systems for storage

and search of the literature other than online systems. Integrated dictionaries have allowed the development of user-friendly information; however, due to the anecdotal nature of these reports, pharmaceutical companies should have a clear policy on how to handle them.

29.3.7 Information Required for Reports

In order to draw a conclusion about the possible relationship between a drug and an AE, certain minimal information elements are required. Points considered essential for literature reports have been proposed (Jones, 1982), and some journals issue guidelines or checklists for potential authors. These can be adapted as a potential checklist for information that should be included in any ADR report as follows:

- Patient demography—age, sex, body weight, height, race, and pregnancy.
- Medical history—previous medical history and concurrent conditions, known allergies (including ADRs with similar drugs), and previous experience with drug.
- Timing—duration of treatment with the suspect drug before AE.
- Concurrent medications—details of other drugs including formulation, dose, and duration.
- Dechallenge—action taken with the suspect drug (stopped, continued, and dose reduction).
- Outcome—outcomes of the AEs.
- Alternative causes—what other factors could have accounted for the AE (diet, occupations exposure) and which were excluded?
- Rechallenge—was the patient rechallenged, and, if so, what was the result?
- Relevant additional data—blood levels, laboratory data, biopsy data, and, where relevant, postmortem findings.

29.3.8 Adverse Drug Reaction Forms and Form Design

Many forms are used by different organizations to collect ADR information. Most regulatory authorities have their own form (see Figure 29.2—the FDA 3500A). Although the content of these forms is similar, little attempt has been made to standardize the design other than by CIOMS.

In order to design the best form for their needs, users must first define what data they wish to collect and which factors are of the greatest importance. In addition, all the usual factors in form design need to be considered (e.g., size, layout, color, print type, spacing, flow of questions, boxes, language, and instructions). A pilot to test the form should be carried out before formal introduction and use.

MEDWATCH

THE FDA MEDICAL PRODUCTS REPORTING PROGRAM

Approved by FDA on 3/27

| |
|------------------|
| Mfr report # |
| UF/Dist report # |
| FDA Use Only |

| A. Patient information | | | |
|--|--|---|--|
| 1. Patient identifier in confidence | 2. Age at time of event: or _____ Date of birth: | 3. Sex <input type="checkbox"/> female <input type="checkbox"/> male | 4. Weight _____ lbs or _____ kgs |
| B. Adverse event or product problem | | | |
| 1. <input type="checkbox"/> Adverse event and/or <input type="checkbox"/> Product problem (e.g., defects/malfunctions) | | | |
| 2. Outcomes attributed to adverse event (check all that apply) | | | |
| death _____ (mo/day/yr) | | <input type="checkbox"/> disability | |
| <input type="checkbox"/> life threatening | | <input type="checkbox"/> congenital anomaly | |
| <input type="checkbox"/> hospitalization—initial or prolonged | | <input type="checkbox"/> required intervention to prevent permanent impairment/damage | |
| <input type="checkbox"/> other: _____ | | | |
| 3. Date of event (mo/day/yr) | | 4. Date of this report (mo/day/yr) | |
| 5. Describe event or problem | | | |
| 6. Relevant tests/laboratory data, including dates | | | |
| 7. Other relevant history, including preexisting medical conditions (e.g., allergies, race, pregnancy, smoking and alcohol use, hepatic/renal dysfunction, etc.) | | | |

FDA

Domain Facsimile of FDA
Form 3500A

Submission of a report does not constitute an admission that medical personnel, user facility, distributor, manufacturer or product caused or contributed to the event

| C. Suspect medications(s) | | | |
|---|---------------------------|--|--|
| 1. Name (give labeled strength & mfr/labeler, if known) | | | |
| #1 _____ | | | |
| #2 _____ | | | |
| 2. Dose, frequency & route used | | 3. Therapy date (if unknown, give duration) from/to (or best estimate) | |
| #1 _____ | | #1 _____ | |
| #2 _____ | | #2 _____ | |
| 4. Diagnosis for use(indication) | | 5. Event abated after use stopped or dose reduced | |
| #1 _____ | | #1 <input type="checkbox"/> yes <input type="checkbox"/> no <input type="checkbox"/> doesn't apply | |
| #2 _____ | | #2 <input type="checkbox"/> yes <input type="checkbox"/> no <input type="checkbox"/> doesn't apply | |
| 6. Lot # (if known) | 7. Ex. date (if known) | 8. Event reappeared after reintroduction | |
| #1 _____ | #1 _____ | #1 <input type="checkbox"/> yes <input type="checkbox"/> no <input type="checkbox"/> doesn't apply | |
| #2 _____ | #2 _____ | #2 <input type="checkbox"/> yes <input type="checkbox"/> no <input type="checkbox"/> doesn't apply | |
| 9. NDC #-for product problems only (if known) | | | |
| #1 _____ #2 _____ | | | |
| 10. Concomitant medical products and therapy dates (exclude treatment of event) NI | | | |
| G. All manufacturers | | | |
| 1. Contact office – name/address (& mfring site for devices) | | 2. Phone number | |
| 4. Date received by manufacturer (mo/day/yr) | | 3. Report Source (check all that apply) | |
| 6. If IND, protocol # | | <input type="checkbox"/> foreign | |
| 7. Type of report (check all that apply) | | <input type="checkbox"/> study | |
| <input type="checkbox"/> 5-day <input type="checkbox"/> 15-day | | <input type="checkbox"/> literature | |
| <input type="checkbox"/> 10-day <input type="checkbox"/> periodic | | <input type="checkbox"/> consumer | |
| <input type="checkbox"/> initial <input type="checkbox"/> follow-up | | <input type="checkbox"/> health professional | |
| | | <input type="checkbox"/> user facility | |
| | | <input type="checkbox"/> company representative | |
| | | <input type="checkbox"/> distributor | |
| | | <input type="checkbox"/> other: | |
| 9. Mfr. report number | | 8. Adverse event term(s) | |
| E. Initial reporter | | | |
| 1. Name, address & phone # | | | |
| 2. Health professional? <input type="checkbox"/> yes <input type="checkbox"/> no | | 3. Occupation | |
| | | 4. Initial reported also sent report to FDA <input type="checkbox"/> yes <input type="checkbox"/> no <input type="checkbox"/> unk | |

FIGURE 29.2 FDA form 3500A.

Consideration should be given to what happens to the form once it is returned. Form design will be affected depending upon whether it is intended to serve as a direct entry document (i.e., the data elements closely match the

data entry screens) or whether a transcription document will be used.

The key factor in ADR form design is the compatibility with other forms required for output, most importantly

regulatory authority forms. The FDA, for example, required ADR reports to be submitted on an FDA3500A (Figure 29.2). If the pharmaceutical company does not wish to collect data on an FDA3500A but must submit reports to the FDA, it will need to design a form that collects the same information. AE report forms generally collect the basic data elements outlined in the following text:

- Patient demography
- Relevant medical history and allergies
- Suspect and concurrent drugs, route, and indication
- AE(s)
- Treatment and management of AE
- Dechallenge, rechallenge, and outcome
- Relevant laboratory data
- Reporter's opinion of causality
- Report's source of information

The form can be printed as a folding postage prepaid envelope for domestic use to encourage a reply. The pharmaceutical company must be able to demonstrate due diligence in seeking relevant follow-up information on each AE report.

Within the next 2 years, several key regulatory authorities including the MCA and FDA will require electronic data submission by companies for both expedited and nonexpedited case reports. The compatibility between the company and the regulatory authority's databases with regard to content and format of the key data elements for transmission is a critical factor to success of these initiatives. The adoption of internationally sanctioned standards such as a dictionary of medical terms, various code lists (e.g., countries, routes, and units), file formats, and periodic safety update reports is essential to enable efficient and accurate transmission. The ICH guideline (ICH E2B) defines data elements for transmission of ICSRs. The guidelines aim to standardize the data elements for all ICSRs regardless of source and destination and cover reports for both preapproval and postapproval periods. It also defines the minimum information for a report and the requirements for proper processing of the report. The medium for electronic submissions will be electronic data interchange (EDI)-encrypted transmissions over the internet.

29.3.9 Computerization of Drug Safety Data: Data Collection and Input

"Rubbish in, rubbish out" applies to safety data as to any other computerized data. The enforced control of terms at entry can be linked to checking of data, which should form part of the quality control procedures. Such controls should be driven by the business so that clinical trial data, free from

all errors and needed for statistical analysis, will probably involve double data entry, whereas single data entry is generally considered adequate for AE databases used for signal generation and regulatory reporting.

Data are still generally typed into a database rather than electronically loaded from other systems. The first step of any data entry process should involve a check for duplicate cases. The need for decision making at the data entry stage will depend upon the type of database design. In all cases, there should be clear rules on how data should be entered into each field to ensure consistency and aid subsequent searching and outputting. This is particularly important when there are multiple users distributed over a number of international sites. Use of electronically available field-specific lists of value and well-defined coding conventions will help with this.

In the future, data will increasingly be captured electronically. Image processing and developments in optical character recognition are already proving useful. Electronic data capture (using fax- or pen-based methods) is used to collect data in some clinical trial.

With the increase in licensing agreements between pharmaceutical companies, safety data frequently needs to be exchanged between one or more parties. If the case volume is sufficient, it is worth considering electronic data exchange between the databases involved. In addition to preventing rekeying of data, this minimizes discrepancies between the data sets. With the adoption of proposed ICH standards in the future, this will become a much simpler process.

29.3.10 Medical and Drug Terminology

Medical and drug terminology is at the heart of the ADR systems. Accurate and consistent input of terms is critical for retrieval and analysis of ADR information. An integrated dictionary allows the capture of original text, which is autoencoded against the dictionary to retrieve the correct code for that piece of text. Coded information allows easy retrieval and analysis. The dictionary structure should allow different ways of grouping and analyzing data encompassing body systems at the highest level to specific reporter's wording that should meet the following needs:

- Acceptable to all users of the system.
- New terms can be easily added.
- Specificity of the reported term preserved.
- Hierarchical structure to group terms at various levels of specificity.
- Logical groupings so similar terms are not scattered.
- A default grouping for each term.
- Unambiguous to enable autoencoding on input.

29.3.11 Dictionaries

This section compares commonly used dictionaries in monitoring drug safety. As electronic exchange of ADR data between industry and regulatory authorities in different countries increases, so does the need for standardization of terminology (Benichou et al., 1991). Medical Dictionary for Regulatory Activities (MedDRA) has completed development with version 4.0 just being available and is discussed later in this chapter. Table 29.5 presents a summary of its structure. (Gruchalla, 1995; Brown et al., 1999).

29.3.12 Medical Term Coding Dictionaries

It is logical to deal with AEs, indications, diseases, surgeries, and procedures using one system for the following reasons:

- ADRs frequently mimic spontaneously occurring diseases; hence, the same diagnosis or symptom could appear as an AE or disease.
- In the identification of new ADRs, it is important not to separate a possible side effect from a disease.
- Separate classifications can lead to confusion and add a layer of complexity when developing ADR systems.

Meaningful codes may or may not be needed for modern dictionaries. For example, the new Adverse Drug Reactions Online Information Tracking (ADROIT) dictionaries do not use meaningful codes but rely on linkage of related terms and effective text processing. Where codes are considered necessary, they should be as short as possible (Westland, 1991). Whenever a system is used for AEs from the literature, spontaneous reports, clinical trials, or a combination of these, the needs of the users of the system will influence the selection of the dictionary.

29.3.13 Medical Dictionary for Regulatory Activities

MedDRA is a medical dictionary encompassing terms relevant to pre- and postmarketing phases of the regulatory process. It was developed by the MCA to support its information systems and has subsequently been further developed by the MCA to support its information systems and medical terminology working group. The objective is to harmonize on standards for electronic submissions among regulatory authorities, between authorities, and industry within and across regions. The aims of the dictionary are:

- To address pre- and postmarketing AE reporting
- To cover multiple medical product areas
- To be available in multiple languages
- To be available in multiple formats and platforms
- To be well maintained

The guiding principles are:

- To build from existing terminologies to maximize compatibility
- To focus on the international community need rather than optimizing on individual countries
- To ensure worldwide use through collaboration and participation in development
- To ensure mechanisms and structures are in place for translation into many languages
- To ensure long-term maintenance

The scope of MedDRA is as follows:

- Disease
- Diagnoses

There is a dual classification for some terms (e.g., 573.1 “Hepatitis in viral diseases classified elsewhere”), but this is not extensive. The dictionaries are very comprehensive with the exception of symptoms, which tend to be scattered. They have been widely used in coding patient histories and hospital charts.

ICD-9 CM is a clinical modification of ICD-9 and offers some advantages, particularly the inclusion of synonyms, but is constrained by systems that have used the older versions of ICD-9. ICD-10 is more comprehensive than any ICD revision to date (see websites). It extends well beyond the traditional causes of death and causes of hospitalization. The content has been expanded to include symptoms, signs, abnormal findings, factors related to lifestyle, and other factors causing contact with health services:

- Signs and symptoms
- Therapeutic indications
- Investigation names and qualitative results
- Medical and surgical procedures
- Medical, social, and family history
- Terms from COSTART, WHO-ART, ICD-9, HARTS, and J-ART

The current structure of MedDRA is defined in Table 29.8. There will be a central maintenance organization responsible for development, user support, implementation, and communication as well as an international user group. A management board will oversee the activities of the central maintenance organization with direction provided by the ICH steering committee. A standard medical dictionary will facilitate electronic data exchange between industry and regulatory authorities worldwide, as recommended by the ICH.

TABLE 29.8 MedDRA Structure

| Level of Hierarchy | Approximate Number of Terms | Definition | Example |
|-----------------------|-----------------------------|---|-----------------------------------|
| System organ class | 26 | Broadest collection of concepts for retrieval; grouped by anatomy or physiology | Cardiac disorders |
| High-level group term | 333 | Broad concepts for linking clinically related terms; can be linked to one or more SOC's | Cardiac rhythm disorders |
| High-level term | 1685 | Groups of preferred terms related by anatomy, pathology, physiology, etiology, or function; can be linked to one or more high-level groups terms or SOC's | Tachyarrhythmia |
| Preferred term | 14,287 | International level of information exchange; single, unambiguous clinical concept | Ventricular tachycardia |
| Lowest-level term | 51,083 | Synonyms and quasisynonyms; help define scope of preferred terms | Praxysmal ventricular tachycardia |

29.3.13.1 FDA Under the March 2001 draft guidance *Postmarketing Safety Reporting for Human Drug and Biological Products including Vaccines* (FDA, 2001), FDA will accept SAEs coded with either MedDRA, COSTART, or WHO-ART. MedDRA has been implemented for SAE coding in FDA's Adverse Event Reporting System (AERS) program. While FDA encourages companies to use MedDRA, the deadline for full MedDRA implementation is still pending.

29.3.13.2 European Union The European Agency for the Evaluation of Medical Products (EMA) established January 2002 as the deadline for all electronically filed single case reports to be coded in MedDRA. All ADR drug reporting must be coded in MedDRA by January 2003.

29.3.13.3 Japan MedDRA/J, the Japanese version of MedDRA, officially was issued on December 28, 1999 and the Ministry of Health highly recommended its use for ADR reporting beginning at the end of March 2000. However, J-ART terms are still applicable and upon submission to the ministry are being converted to MedDRA/J terms. No firm deadline for full implementation has been issued.

MedDRA is available in English and Japanese only. The Maintenance and Support Services Organization (MSSO) is working on translations in French, Portuguese, German, Greek, and Spanish. With their annual dues, subscribers can get MedDRA in English and one other European language (when available). Japanese or additional European languages will need to be purchased separately. The German and Portuguese translations were recently submitted to MSSO for review.

The most current version of MedDRA, 11.0, was released in March 2008. The cost of the dictionary depends on the type of organization and annual revenue. An annual subscription provides a company with all versions of MedDRA released during the year. All regulators are provided MedDRA free of charge. Otherwise, costs are as follows.

29.3.14 Periodic Reports

Many regulatory authorities require detailed summary reports on groups of cases on a regular basis. The FDA requires annual progress reports for investigational compounds and periodic reports for marketed drugs either quarterly or annually depending upon the length of time the product has been on the US market. CIOMS II guidelines recommend submission of line listings of serious, unlabelled spontaneous cases in conjunction with a summary of the drug safety profile on a 6-monthly basis. These reports are well defined in format, content, and submission time frame. Most major pharmaceutical companies produce them electronically.

The regulatory requirements, particularly regarding frequency of submission and content, differ in the three regions (Europe, Japan, and the United States). In order to avoid duplication of effort and to ensure that important data are submitted with consistency to regulatory authorities worldwide, the ICH3 Topic E2C Guideline on the format and content for comprehensive periodic safety update reports (PSUR) (1996) of marketed medicine products has been developed. The general principles of this guideline include:

- One report is submitted for one active substance. All dosage forms as well as indications for a given active substance should be covered in one PSUR.
- The focus is on ADRs, which include all spontaneous reports and all drug-related clinical trial and literature reports.
- An international birthdate and frequency of review and reporting are defined. The international birthdate is the date of the first marketing authorization for the product granted to any company in any country in the world. Preparation of PSURs should be based on data sets of 6 months or multiples thereof. The PSUR should be submitted within 60 days of the data lock point.

- The reference safety information is the company core data sheet to determine whether an ADR is listed or unlisted.
- ADR data are presented in line listings and/or summary tabulations.

29.4 CAUSALITY ASSESSMENT

Decisions have to be made by pharmaceutical companies and regulatory authorities about whether a drug can cause a particular AE so that an appropriate action can be taken. What does “can cause” mean? Does it imply certainty? In many cases to wait for “certainty” before taking action would entail many patients suffering unnecessarily. The degree of certainty or “probability” required will vary according to the situation.

There are, nearly always, many factors other than the administration of a drug that can cause an AE and will determine whether the AE will occur in a particular patient. The drug may be “the last straw that broke the camel’s back.” If an AE would not have occurred as and when it did but for the drug, then the drug “caused” the AE (Hutchinson, 1992). So with an adverse drug interaction, both drugs “caused” the AE. Using this definition the drug may only be a minor factor.

Certainty is rarely obtainable; perhaps an AE with a positive rechallenge where there is objective evidence and an absence of confounders in an individual case would be considered as certainty due to the drug. In the majority of cases, action is needed before there is absolute certainty that a drug can cause an AE. This lack of certainty in individual cases has been described using rather vague terms such as “almost certain,” “probably,” “possible,” “unlikely,” etc. These terms have also been defined, but each author has a slightly different definition (Venulet et al., 1982; Stephens et al., 1998).

Again, in epidemiological studies or clinical trials, there is nearly always a degree of uncertainty due to bias, chance, and confounders. In these studies uncertainty is measured in terms of *p* values, odd ratios, relative risks, etc.

The differential diagnosis of AEs associated with a drug or drug(s) is an everyday part of a practicing clinician’s life (Rogers, 1987). However, the term “causality assessment” is reserved for a similar process performed at one or more stages removed from the patient and with some important differences. Clinicians do not necessarily need to find out whether a drug caused an AE in order to satisfy themselves and their patients. They will be more interested in resolving the event as quickly as possible. If there is a possibility that the event might be an ADR, it may be resolved by either reducing the dose or stopping the drug or by treating the ADR while waiting for tolerance to develop, or it may resolve if any of the underlying factors are altered. The resolution of the AE might be because the event has been

caused by the drug or it may have been a transient natural occurrence; either way the patient and doctor will welcome its disappearance. If, however, the doctor is interested in knowing whether it was an ADR, further investigations can be undertaken, as long as the patient is willing, until it is established or refuted.

When causality assessment is undertaken by a regulatory authority or a scientist/physician in industry, it is unlikely that the full details known to the clinician treating the patient will be reported, even after further inquiry is made. The only way to obtain all available data is usually by visiting the physician and, with permission, reading the notes and discussing the case with him or her.

29.4.1 Aims of Causality Assessment

Of the many similar events on an AE database, only a few have sufficient and relevant data to enable the assessor to decide that the AE was more likely caused by the drug than by any other cause or vice versa. A preliminary assessment (sometimes referred to as a “triage”) can be made by placing the event into a category (e.g., probably, possible or unlikely, or using the EEC classification of A, B, or O) (Meyboom and Royer, 1992). This will enable the company to extract the probable cases at regular intervals in order to consider whether there is a “signal.” The possible and unlikely cases will probably not contribute much to this signal.

This preliminary assessment will need to be updated as and when further information becomes available. It should favor sensitivity over specificity so that a borderline possible/probable case is classified as probable rather than possible to make certain that the case is not lost when at a later stage the probable cases are picked out as a signal. A full assessment when all the information is available can then rectify any misclassifications.

29.5 COURSES OF CORRECTIVE ACTION

Identification of a safety issue with a marketed drug does not necessarily (or even usually) lead to the withdrawal of that therapeutic agent from the market. As noted at the beginning of this chapter, there is a range of possible actions. These include:

- Change in dosage or dose form (reformation)
- Change in labeling (warnings)
- Restriction of situation of use (from open prescription to either clinician administration or hospital use only)
- Monitoring of patients during use
- Restriction on use (that is, of patients allowed to use)
- Withdrawal from the market (usually a permanent step but not always)

TABLE 29.9 Criteria for Negligence—D + L + F + C = N

| | |
|----------------------------|--|
| D: duty of care | Owed to the claimant; easy to establish in the case of the supplier/manufacturer <i>vis-à-vis</i> the patient who uses the product |
| L: lack of reasonable care | Evidenced by a failure to conduct operations according to accepted standards applicable at the time—that is, a breach of regulatory requirements or possibly failure to take account of or apply (industry) guidelines |
| F: foreseeable injury | Of the type likely to occur following failure (e.g., side effect of the drug) |
| C: causation | The lack of reasonable care must have caused/contributed to the injury; if a label would not have been read by the patient in any event, an omission from it might not have caused injury |

Which action(s) are taken depends on severity and incidence rates of the adverse response, technical details, the existence of alternative therapies, and the benefit of the use of the drug.

29.6 LEGAL CONSEQUENCES OF SAFETY WITHDRAWAL

Although in the context of personal injury claims an HMO and other parties (e.g., doctors, hospital) may all be the target of proceedings, it is usually the pharmaceutical company, perceived as having “deep pockets,” that is the prime target for claimants. Claims for negligence based upon a failure to act with reasonable care (e.g., to obtain or act upon PV data) and/or the supply of a product that is “defective” in legal terms (e.g., because its labeling was not amended, pursuant to the receipt and review of PV data so as to give adequate warnings and precautions) are always possible.

Tables 29.9 and 29.10 set out in very simple terms the necessary “ingredients” for establishing product liability, either in negligence or under statute: so-called strict liability.

All of the elements of each of these legal wrongs must be present in a given situation for liability to be established. In negligence therefore, where the claim is made against the person alleged to owe the duty of care (in the context of this chapter, this will be the company putting the product on the market), proof of causation, without a lack of reasonable care having occurred, will not afford the claimant a remedy. However, the chief distinction between negligent liability and so-called strict liability is that in the case of the latter, fault is not required to be shown. To establish strict liability, the claimant must establish against the “producer” (manufacturer/importer) that the product was “defective” (for the purpose of the law, this could refer to shortcomings in its presentation, design, or manufacture) and that it caused the injury suffered.

It would not be at all unusual for claimants in personal injury actions to look for a regulatory compliance failure on the part of a company defendant. The demonstration of a regulatory breach will significantly assist the plaintiff in establishing lack of reasonable care (i.e., conduct falling below acceptable standards). In fact, whether the failure is alleged to be directly relevant to the injury or not, it can

TABLE 29.10 Criteria for Strict Liability—D + D + C = SL

| | |
|--------------|---|
| D: defect | Widely defined—product design defect, manufacturing error (so that the product is less safe than persons generally would be entitled to expect), deficiency in “presentation” |
| D: damage | To person or property flowing from the defect |
| C: causation | See Table 26.7 |

be used to demonstrate a general lack of care in the operation of corporate systems with prejudicial effect. Failure to warn is a common element of many pharmaceutical product liability cases, where the pleadings (of negligence and strict liability) might be expected to assert that had the labeling accurately dealt with contraindications, precautions, and/or warnings, the patient would have avoided the injury allegedly suffered, either because the product would not have been used/administered at all, or the patient would have been monitored, advised (by the treating doctor), or managed differently so as to avert injury.

In a case where PV omissions are identified that can be said to lead to no/or an insufficient response being adopted by the manufacturer (especially where the regulatory authorities have taken some form of action or simply criticized a company), the plaintiff is a significant way toward establishing a case for lack of reasonable care in negligence, or that the product was defective in strict liability terms, because it was not presented accurately and was therefore less safe than persons were entitled to expect, given the content of the labeling.

29.6.1 FDA Tools for Risk Management

The FDA Amendments Act of 2007 (PDUFA IV) gives FDA an array of new regulatory tools to exert tighter controls over prescription drugs in the postmarketing setting (McCaughan, 2008).

The centerpiece of the new authorities is included in the section of the law allowing FDA to impose risk management plans on new drugs, under the new acronym Risk Evaluation and Mitigation Strategies (REMS). But the spirit of the REMS lives throughout the new law, providing the FDA with new authority providing direct onset of consumer ads,

a proposed active surveillance system, mandatory postmarketing study provisions, and new safety labeling procedures. Fundamentally, these arise from the belief that product labeling alone is not sufficient to ensure the safe, appropriate use of prescription drugs.

This belief that top FDA officials have expressed publicly for a decade grew out of the agency's frustration over a series of earlier product withdrawals like the type 2 diabetes drug troglitazone (Rezulin) and the cholesterol agent cerivastatin (*Baycol*).

In both those cases (and several less prominent examples), FDA officials believed that the product could be safely marketed if only prescribers followed appropriate use instructions already included in labeling. But such labeling actions did not and the drugs ultimately were taken off the market for safety reason:

Such concerns gave rise to the voluntary risk management plans that serve as the models for the REMS. The amendment makes REMS the new *de facto* baseline for drug approvals, setting a formal mechanism for FDA to impose the programs.

The standard for FDA to impose REMS is quite low: it can do so if a program "is necessary to ensure the benefits of a drug outweigh the risks." That, of course, is the standard FDA already applies for any new drug approval decisions.

That does not mean FDA will impose REMS for every new drug but rather that it can. So sponsors need to anticipate that every new drug review will involve a discussion of whether REMS is necessary, in essence forcing sponsors to work with FDA to define a program up front or prove that one is not necessary for a new drug.

Not all REMS are created equal. Instead, the new law allows FDA to impose different degrees of control depending on the profile of the product under consideration.

29.6.1.1 The Regulatory Pyramid FDA's new regulatory tools can be envisioned as forming a pyramid, with the base representing a relatively small increase in the regulatory burdens already imposed on sponsors in the postmarketing setting and the tip of the pyramid representing the strictest distribution controls that essentially require patients to be identified and screened individually before the medicine is administered.

Though it is not explicitly defined in law, each tier of the pyramid layers new regulatory controls on top of those used in the lower tiers. The result: products regulated on the lower tiers can reach a larger share of the potential patient population for the medicine, while products regulated on higher tiers are limited to narrower and narrower slices.

Here is one approach to arranging FDA's new regulatory authorities that may be useful in drug development planning, along with some examples of currently marketed products that already have these kinds of restrictions in place.

29.6.2 Tier 1: Mandatory Studies

The lowest tier in the new regulatory system will be products for which FDA does not impose REMS at all but does mandate postmarketing studies. FDA's authority to mandate studies is independent of the new REMS provisions in FDAAA, and it is broadly crafted. FDA can mandate studies (i) to assess a known safety risk identified prior to marketing, (ii) to investigate a safety signal about a product, or (iii) to identify "an unexpected serious risk when available data indicates the potential for a serious risk."

Those criteria suggest that FDA will have the ability to argue for mandatory studies for essentially any drug.

The mandatory study authority also interacts with the separate provisions of the law directing FDA to create an active surveillance system via a public/private partnership.

The law stipulates that FDA is not to impose a clinical trial requirement if the active surveillance system can answer a safety question. On the other hand, the active surveillance network will be certain to generate plenty of safety signals that could serve as justification for FDA to impose Phase IV trials.

Still, products in the lowest tier of FDA's new regulatory authority will have no special limits on their distribution or the potential access to patients. Indeed, there is nothing in the law that prevents sponsors from using studies of new indications—trials that could expand use of their medicines—to answer safety questions about existing uses.

Any product marketed with a voluntary *Phase IV* commitment can serve as an example for this tier and, in particular, products approved under FDA's accelerated approval mechanism, which face the threat of a streamlined withdrawal process if the sponsor fails to complete a postmarketing trial or if the trial fails to prove clinical benefit.

But even accelerated approval medicines are not perfect examples of the new regulatory reality, since FDA has not so far used its enhanced power to pull a product from the market for failure to complete trials. There is one pending case where FDA may be considering using that authority: the agency has written to generic drug manufacturers who market versions of Shire PLC's orthostatic hypertension therapy midodrine (*ProAmatine*), pointing out that the required postmarketing studies are not yet complete and inviting comments on proposals from some of the generic manufacturers to do the trials themselves.

One of the first high-profile examples of a sponsor preparing for the new reality of mandatory postmarketing commitments for already marketed products is Amgen Inc., which will be further refining its postmarketing research program for darbepoetin alfa (*Aranesp*).

29.6.3 Tier 2: Labeling and Assessment

The next tier up the pyramid is the first layer of the new REMS authority. The law says that the minimal requirement of REMS is a periodic assessment of a product's postmarketing

safety profile. In other words, everything about the drug would be the same as a non-REMS product, but there would be predefined timelines to assess how effective the labeling is working to ensure safe use.

FDA's separate authority to dictate labeling changes will also come into play in this tier. After 18 months on the market (the first predefined assessment point), FDA and the sponsor would review AE reports and other new data (eventually including the active surveillance system). The new information would be added to labeling, and warning statements might be revised. But in the event of a disagreement about the interpretation of the data, FDA now has the upper hand—the agency can force a change (subject to due process provisions in the law).

In principle, there is nothing new about the idea that FDA will be analyzing postmarketing safety reports for new drugs. But the deadline for a formal rereview makes a profound difference in the impact of those analyses.

FDA is already pilot testing this authority via a series of planned drug safety “report cards” for new molecular entities. The agency published its first review in September, looking at spontaneous AE reports for oral chelating agent deferasirox (Exjade). The review prompted a new round of media coverage about acute renal failure and cytopenia associated with the drug; Novartis alerted prescribers to the issue in May.

Going forward, sponsors will have to be cognizant of the deadlines for reassessment and wary of FDA's ability to dictate new labeling.

At a minimum, sponsors will need to enhance their own postmarketing monitoring of products regulated under this tier to prevent any surprises in their interactions with FDA. And in some cases sponsors may decide to take voluntary steps on their own to further restrict use if they believe the alternative will be facing an unworkable warning in labeling.

29.6.4 Tier 3: Enhanced Communication

The REMS authority allows FDA to require consumer medication guides and enhanced communications for doctors or other tools to communicate appropriate use information more effectively. Current examples of risk management tools that would fall under this authority might include special stickers to be affixed to prescriptions, agreements to limit sampling, enhanced patient compliance programs, or voluntary agreements not to advertise to consumers.

FDA's new direct-to-consumer ad review user fee program and its separate authority to require disclaimers in broadcast ads also should be considered as part of this regulatory tier. Though not part of the REMS, the DTC provisions will work in the same way by giving FDA a stronger hand in shaping the final version of safety messages included in broadcast ads.

The enhanced communication provisions obviously place a higher regulatory burden on the sponsor. Sponsors operating

under an enhanced communication plan may find themselves in jeopardy if there is significant, measurable use of the product in a setting that FDA considers inappropriate.

The agency will consider imposing more restrictive REMS in that situation; it could also seek to hold the company accountable for the failure of the program. The new law makes REMS mandatory with potentially significant penalties for violations. A sponsor that willfully promotes a product in a manner that runs counter to the enhanced communication plan is obviously in jeopardy. But FDA may also view the failure of the plan as suspicious in itself, triggering a deeper investigation of a company's marketing activities.

There are quite a few examples of products marketed with enhanced communication plans. Amylin Pharmaceuticals Inc.'s type 2 diabetes therapy pramlintide (*Symlin*) is noteworthy as an early example of a manufacturer voluntarily agreeing not to advertise directly to consumers.

Another instructive example is AstraZeneca PLC's rosuvastatin (*Crestor*). Developed as a “superstatin,” *Crestor* had the misfortune to be pending at FDA at the time of the *Baycol* withdrawal. In response to concerns about potential toxicity from high-potency statins, AstraZeneca agreed not to sample the high dose of *Crestor*. The product also underwent an unplanned safety reassessment after approval thanks in part to a petition filed by a public citizen seeking *Crestor*'s withdrawal. The result was a relabeling with stronger warnings, followed by the resumption of direct-to-consumer advertising and the emergence of *Crestor* as a brand approaching \$3 billion in annual sales.

Two recent Pfizer Inc. approvals also used enhanced communication plans. In the case of the inhaled insulin brand *Exubera*, the program was part of an effective regulatory strategy to resolve safety issues that long delayed its launch. (See “Pfizer's *Exubera*: Breathing New Life into Inhaled Insulin,” *IN VIVO*, October 2005). But getting the product to market turned out to be the easy part: after dismal uptake, Pfizer is giving up on the brands and writing off almost \$3 billion as a result.

A happier example is the smoking cessation agent varenicline (*Chantix*). Pfizer markets the drug in conjunction with a voluntary patient support program and also delayed the launch of DTC advertising as part of its commitment to principles established by the Pharmaceutical Research and Manufacturers of America. Now the brand is one of the rising stars of the company's commercial portfolio, with sales jumping almost 10-fold to \$241 million in the third quarter of 2007.

29.6.5 Tier 4: Safe Use Restriction Defined by Provider

The next level up the pyramid is when the REMS really start to pinch market sizes. The law gives FDA the authority to set distribution restrictions on a drug if “necessary to assure safe

use of the drug because of its inherent toxicity or potential harmfulness.” FDA is supposed to apply the restrictions only when a drug could not be marketed (or would be withdrawn) without them or if less restrictive REMS fail to assure appropriate use.

The safe use authority is one section of the law, but it really represents two tiers of regulation.

The lower tier is a restriction based on the providers. That restriction could be to allow use only by providers who complete a special training program—a burden to the sponsor, to be sure, but one with a clear upside in promoting a new medicine.

Or the restriction could be defined by physician specialty or practice setting. FDA could allow use of a medicine only in hospitals, for example, as was the case for Praecis Pharmaceuticals’ abarelix (*Plenaxis*). The restriction essentially killed the product; Praecis withdrew it and was itself acquired by GlaxoSmithKline PLE.

In other cases, such as with Genentech Inc.’s omalizumab (*Xolair*), the distribution restriction can serve more as an opportunity. *Xolair*’s labeling stipulates that the asthma drug must be administered in a physician’s office to monitor for acute adverse reactions. The product did not initially meet Genentech’s expectations but is now emerging as a prime example of the “minibuster” model that will be critical to the industry in the future.

Still, restricted access provisions place a significant burden on sponsors, typically requiring specialized distribution systems and close coordination with providers. In other words, the primary care blockbuster model that built the modern Big Pharma business does not apply.

29.6.6 Tier 5: Safe Use Restriction Defined by Patient

The highest tier on the pyramid involves the other elements of restricted distribution programs, ones that define the limits by patient-specific criteria.

The law allows FDA to require distribution only to patients with a documented lab test, or to require special monitoring of each patient given the drug, or to mandate patient registries for recipients. This tier represents the tip of the regulatory pyramid, where FDA’s authority dictates the conditions for administering the drug to individual patients.

But even at this layer of regulation, commercial success is possible. Biogen Idec Inc.’s *Tysabri* is marketed under the most restrictive distribution plan approved by FDA. The product is not at the level the sponsor hoped for when it was first launched for MS—but the company believes it could still be a billion dollar brand.

Or consider Celgene Corp.’s thalidomide (*Thalomid*), the drug that ushered in the modern era of risk management

plans when it was approved in 1998. The program has been so successful that it has essentially defined Celgene’s business model—an approach that many other companies are likely to want to imitate post-FDAAA.

There is an even earlier example: Novartis’ atypical anti-psychotic clozapine (*Clozaril*) which was sold with a requirement for blood monitoring, initially tied to a single lab test provider (Caremark). *Clozaril* has been surpassed by many newer atypicals, but it is worth remembering that one of the biggest blockbuster classes of the past two decades began as a prototype for the era of restrictive distribution plans.

REFERENCES

- Adams DG, Cooper RM, Kahan JS. (1997) *Fundamentals of Law and Regulation*, Vol. II. FDLI, Washington, DC.
- Ajayi FO, Sun H, Perry J. (2000) Adverse drug reactions: A review of relevant factors. *J. Clin. Pharmacol.* 40:1093–1101.
- American Society of Hospital Pharmacy. (1995) ASHP guidelines on adverse drug reaction monitoring and reporting. *Am. J. Health Syst. Pharm.* 52:417–419.
- Antonow JA, Smith AB, Silver MP. (2000) Medication error reporting: a survey of nursing staff. *J. Nurs. Care Qual.* 15:42–48.
- Benichou C, Danan G, Solal-Celigny P. (1991) Standardization of International databases: definitions of drug-induced blood cytopenias. *Int. J. Clin. Pharmacol. Toxicol.* 29:75–81.
- Brewer T, Colditz GA. (1999) Postmarketing surveillance and adverse drug reactions: current perspectives and future needs. *J. Am. Med. Assoc.* 281:824–829.
- Brown EG, Wood L, Wood S. (1999) The medical dictionary for regulatory activities (Med DRA). *Drug Saf.* 20:109–117.
- Bukowski JA, Wartenberg D. (1996) Comparison of adverse drug reaction reporting in veterinary and human medicine. *J. Am. Vet. Med. Assoc.* 209:40–45.
- CFR (1994) *Postmarketing Reporting of Adverse Drug Experiences*. CFR 314.80. US Government Publishing Office, Washington, DC.
- Chyka PA. (1999) Role of US poison centers in adverse drug reactions monitoring. *Vet. Hum. Toxicol.* 41:400–402.
- Chyka PA, McEommon SW. (2000) Reporting of adverse drug reactions by poison control centers in the United States. *Drug Saf.* 23:87–93.
- Cohen JS. (2001) Dose discrepancies between the physicians desk reference and the medical literature and their possible roles in the high incidence of dose-related adverse drug events. *Arch. Intern. Med.* 161:957–964.
- FDA (2001) *Postmarketing Safety Reporting for Human Drug and Biological Products including Vaccines*. Draft Guidance, Center for Drug Evaluation and Research (CDER) and Center for Biologics Evaluation and Research (CBER), Rockville, MD, March, 2001.

- FDA (2005) *Guidance for Industry: Good Pharmacovigilance Practices and Pharmacoepidemiologic Assessment*. Available at: <http://www.fda.gov/downloads/RegulatoryInformation/Guidances/UCM126834.pdf> (accessed March 31, 2016).
- Figueras A, Tato F, Fontainas J, Gestal-Otero JJ. (1999) Influence of physician's attitudes on reporting adverse drug events: a case-control study. *Med. Care.* 37:809–814.
- Fung M, Thorton A, Mazbeck K, Wu JH, Hornbuckle K, Muniz E. (2001) Evaluation of the characteristics of safety withdrawal of premarketing prescription drugs from worldwide pharmaceutical markets—1960 to 1999. *Drug Inf. J.* 35:293–317.
- Goldman SA. (1998) Limitations and strengths of spontaneous reports data. *Clin. Ther.* 201(Supplement C):CAO-58.
- Gruchalla RS. (1995) A one-year perspective on MedWatch: the Food and Drug Administration's new medical products reporting program. *J. Allergy Clin. Immunol.* 95:1153–1157.
- Holland EG, DeGruiz FV. (1997) Drug-induced disorders. *Am. Fam. Physician.* 56:1781–1788.
- Hunziker T, Kunyi UP, Braunschweig S, Zehnder D, Hoigne R. (1977) Comprehensive hospital drug monitoring (CHDM): adverse skin reactions, a 20-year survey. *Allergy.* 52:388–393.
- Hutchinson TA. (1992) Causality assessments of suspected adverse drug reactions. In: *Detection of New Adverse Drug Reactions* (Stephens MBD, Ed.). Macmillan Press, London.
- Jones JK. (1982) Criteria for journal reports of suspected adverse drug reactions. *Clin. Pharm.* 1:554–555.
- Keller WC, Bataller N, Oeller DS. (1998) Processing and evaluation of adverse drug experience reports at the Food and Drug Administration Center for Veterinary Medicine. *J. Am. Vet. Med. Assoc.* 213:208–211.
- Kennedy DL, Goldman SA. (1997) Monitoring for adverse drug events. *Am. Fam. Physician.* 56:1748–1788.
- LaCalamita S. (1995) Top 10 reasons for not reporting adverse drug reactions. *Hosp. Pharm.* 30:245–246.
- Mann R, Andrews E. (2007) *Pharmacovigilance*. John Wiley & Sons, Inc, Hoboken, NJ.
- Martin RM, Kapoor KV, Wilton LV, Mann RO. (1998) Underreporting of suspected adverse drug reactions to newly marketed ("black triangle") drugs in general practice: observational study. *Br. J. Med.* 317:119–120.
- McCaughan M (2008) *Pyramid Schemes and Pipeline Dreams: Adapting to a Tiered Regulatory System*, RTM Report. Informa, London.
- Meyboom RHB, Royer RJ. (1992) Causality classification at pharmacovigilance centers in the European Community. *Pharmacoepidemiol. Drug Saf.* 1:87–97.
- Moor TJ. (1995) *Deadly Medicine*. Simon & Schuster, New York.
- Niu MT, Rhodes P, Saline M, Livelz T, Davis DM, Black S, Shinefield H, Chen RT, Ellenberg SS, the VAERS and VSD Working Groups. (1998) Comparative safety of two recombinant hepatitis B vaccines in children: data from the Vaccine Adverse Event Reporting System (VAERS) and Vaccine Safety Data links (VSD). *J. Clin. Epidemiol.* 51:503–510.
- Niu MT, Saline ME, Ellenberg SS. (1999) Neonatal deaths after hepatitis B. Vaccine: The Vaccine Adverse Event Reporting System, 1991–1998. *Arch. Pediatr. Adolesc. Med.* 153:1279–1282.
- Olsson S. (1998) The role of the WHO programme on International Drug Monitoring in coordinating worldwide drug safety efforts. *Drug Saf.* 19:1–10.
- Piazza-Hepp TD, Kennedy DL. (1995) Reporting of adverse events to MedWatch. *Am. J. Health Syst. Pharm.* 52:1436–1439.
- Rogers AS. (1987) Adverse drug events: identification and attribution. *Drug Intell. Clin. Pharm.* 21:915–920.
- Rose JC, Elnis AS. (2000) A mortality index for postmarketing surveillance of new medications. *Am. J. Emerg. Med.* 18(2):176–179.
- Schuster D, Laggner C, Langer T. (2005) Why drugs fail—a study on side effects in new chemical entities. *Curr. Pharm. Des.* 11:3545–3559.
- Sharpe R (1998) MedWatch system comes under fire. *Wall Street J.* June:B5.
- Snidermann OA. (2000) The need for greater involvement of regulatory agencies in assessing adverse drug reactions. *CMAJ.* 162:209–210.
- Stephens MDB, Talbot JCC, Routledge PA. (1998) *Detection of New Adverse Drug Reactions*, 4th Ed. Grove Dictionaries, New York.
- Strom B, Kimmel S, Hennessy S. (Eds.) (2012) *Pharmacoepidemiology*, 5th Ed. John Wiley & Sons, Ltd, Chichester.
- Varrinchio F. (1998) The vaccine adverse event reporting system. *Clin. Toxicol.* 36:765–768.
- Venulet J, Blattner R, Von Bülow J, Bernecker GC. (1982) How good are articles on ADR? *Br. Med. J.* 284:252–254.
- Westland MM. (1991) Coding: the mortar in the bricks of data analysis. *Drug Inf. J.* 25:197–200.
- White GG, Love L. (1998) The MedWatch program. *Clin. Toxicol.* 36:145–149.
- WHO (1975) *Requirements for Adverse Drug Reaction Reporting*. WHO, Geneva.
- WHO (2010) *Minimum Requirements for a Functional Pharmacovigilance System*. Available at: <http://www.who-umc.org/graphics/24733.pdf> (accessed March 31, 2016).

STATISTICS IN PHARMACEUTICAL SAFETY ASSESSMENT

30.1 INTRODUCTION

Modern nonclinical safety assessment is heavily dependent on the use of multiple tools, one of which is statistical analysis. This chapter is meant to serve the practicing toxicologist and pathologist, as a practical guide to the common statistical problems encountered in drug safety assessment and the methodologies that are available to solve them. The chapter also includes discussions of why a particular procedure or interpretation is recommended, by the clear enumeration of the assumptions that are necessary for a procedure to be valid and by discussion of problems drawn from the actual practice of toxicology and toxicological pathology. Studies are frequently over- or misused in the place of knowledgeable and critical integration and analysis of data.

Studies continue to be designed and executed to generate increased amounts of data. The resulting problems of data analysis have then become more complex and toxicology has drawn more deeply from the well of available statistical techniques. Statistics has also been very active and growing during the last 35 years—to some extent, at least, because of the parallel growth of toxicology. These simultaneous changes have led to an increasing complexity of data and, unfortunately, to the introduction of numerous confounding factors which severely limit the utility of the resulting data in all too many cases.

A major difficulty is that there is a very real necessity to understand the biological realities and implications of a problem, as well as to understand the peculiarities of toxicological data before procedures are selected and employed for analysis. These characteristics include the following:

1. The need to work with a relatively small sample set of data collected from the members of a population (laboratory animals) which is not actually our

populations of interest (i.e., humans or a target animal population).

2. Dealing frequently with data resulting from a sample which was censored on a basis other than by the investigators as design. By censoring, of course, we mean that not all data points were collected as might be desired. This censoring could be the result of either a biological factor (the test animal being dead or too debilitated to manipulate) or a logistic factor (equipment being inoperative or a tissue being missed in necropsy).
3. The conditions under which our experiments are conducted are extremely varied. In pharmacology (the closest cousin to at least classical toxicology), the possible conditions of interaction of a chemical or physical agent with a person are limited to a small range of doses via a single route over a short course of treatment to a defined patient population. In toxicology however, all these variables (dose, route, time span, and subject population) are determined by the investigator.
4. The time frames available to solve our problems are limited by practical and economic factors. This frequently means that there is no time to repeat a critical study if the first attempt fails. So a true iterative approach is not possible.

The training of most pathologists in statistics remains limited to a single introductory course which concentrates on some theoretical basics. As a result, the armamentarium of statistical techniques of most toxicologists is limited, and the tools that are usually present (*t*-tests, chi-square, analysis of variance (ANOVA), and linear regression) are neither fully developed nor well understood. It is hoped that this chapter will help change this situation.

TABLE 30.1 Sample Size Required to Obtain a Specified Sensitivity at $p < 0.05$

| Background Tumor Incidence | P^a | Treatment Group Incidence | | | | | | | | | |
|----------------------------|-------|---------------------------|------|------|------|------|------|------|------|------|------|
| | | 0.95 | 0.90 | 0.80 | 0.70 | 0.60 | 0.50 | 0.40 | 0.30 | 0.20 | 0.10 |
| 0.30 | 0.90 | 10 | 12 | 18 | 31 | 46 | 102 | 389 | | | |
| | 0.50 | 6 | 6 | 9 | 12 | 22 | 32 | 123 | | | |
| 0.20 | 0.90 | 8 | 10 | 12 | 18 | 30 | 42 | 88 | 320 | | |
| | 0.50 | 5 | 5 | 6 | 9 | 12 | 19 | 28 | 101 | | |
| 0.10 | 0.90 | 6 | 8 | 10 | 12 | 17 | 25 | 33 | 65 | 214 | |
| | 0.50 | 3 | 3 | 5 | 6 | 9 | 11 | 17 | 31 | 68 | |
| 0.05 | 0.90 | 5 | 6 | 8 | 10 | 13 | 18 | 25 | 35 | 76 | 464 |
| | 0.50 | 3 | 3 | 5 | 6 | 7 | 9 | 12 | 19 | 24 | 147 |
| 0.01 | 0.90 | 5 | 5 | 7 | 8 | 10 | 13 | 19 | 27 | 46 | 114 |
| | 0.50 | 3 | 3 | 5 | 5 | 6 | 8 | 10 | 13 | 25 | 56 |

^a P = Power for each comparison of treatment group with background tumor incidence.

As a point of departure toward this objective, it is essential that any analysis of study results be interpreted by a professional who firmly understands three concepts: the difference between biological significance and statistical significance, the nature and value of different types of data, and causality.

For the first concept, we should consider the four possible combinations of these two different types of significance, for which we find the relationship shown in the succeeding text.

| | | Statistical Significance | |
|-------------------------|-----|--------------------------|---------|
| | | No | Yes |
| Biological Significance | No | Case I | Case II |
| | Yes | Case III | Case IV |

Cases I and IV give us no problems, for the answers are the same statistically and biologically. But cases II and III present problems. In case II (the “false positive”), we have a circumstance where there is a statistical significance in the measured difference between treated and control groups, but there is no true biological significance to the finding. This is not an uncommon happening, for example, in the case of clinical chemistry parameters. This is called type I error by statisticians, and the probability of this happening is called the alpha (α) level. In case III (the “false negative”), we have no statistical significance, but the differences between groups are biologically/toxicologically significant. This is called type II error by statisticians, and the probability of such an error happening by random chance is called the beta (β) level. An example of this second situation is when we see a few of a very rare tumor type in treated animals. In both of these latter cases, numerical analysis, no matter how well done, is no substitute for professional judgment. Along with this, however, one must have a feeling for the different types of data and for the value or relative merit of each. Note that

the two error types interact and in determining sample size we need to specify both α and β levels. Table 30.1 demonstrates this interaction in the case of tumor or specific lesion incidence.

The reasons that biological and statistical significance are not identical are multiple, but a central one is certainly causality. Through our consideration of statistics, we should keep in mind that just because a treatment and a change in an observed organism are seemingly or actually associated with each other does not “prove” that the former caused the latter. Though this fact is now widely appreciated for correlation (e.g., the fact that the number of storks’ nests found each year in England is correlated with the number of human births that year does not mean that storks bring babies), it is just as true in the general case of significance. Timely establishment and proof that treatment causes an effect require an understanding of the underlying mechanism and proof of its validity. At the same time, it is important that we realize that not finding a good correlation or suitable significance associated with a treatment and an effect likewise does not prove that the two are not associated—that a treatment does not cause an effect. At best, it gives us a certain level of confidence that under the conditions of the current test, these items are not associated.

These points will be discussed in greater detail in “Assumptions and Limitations” for each method, along with other common pitfalls and shortcomings associated with the method. To help in better understanding the chapters to come, terms frequently used in discussion throughout this book should first be considered. These are presented in Table 30.2 later.

Each measurement we make—each individual piece of experimental information we gather—is called a datum. However, we gather and analyze multiple pieces at one time, the resulting collection being called data.

Data are collected on the basis of their association with a treatment (intended or otherwise) as an effect (a property)

TABLE 30.2 Some Frequently Used Terms and Their General Meanings (Marriott, 1991)

| Term | Meaning |
|---------------------------------|---|
| 95% CI | A range of values (above, below, or above and below the sample) (mean, median, mode, etc.) have a 95% chance of containing the true value of the population (mean, median, and mode). Also called the fiducial limit equivalent to the $P < 0.05$ |
| Bias | Systemic error as opposed to a sampling error. For example, selection bias may occur when each member of the population does not have an equal chance of being selected for the sample |
| df | The number of independent deviations, usually abbreviated df |
| Independent variables | Also known as predictors or explanatory variables |
| P value | Another name for significance level, usually 0.005 |
| Power | The effect of the experimental conditions on the dependent variable relative to sampling fluctuation. When the effect is maximized, the experiment is more powerful. Power can also be defined as the probability that there will not be a type II error ($1 - \beta$). Conventionally, power should be at least 0.07 |
| Random | Each individual member of the population has the same chance of being selected for the sample |
| Robust | Having inferences or conclusions little effected by departure from assumptions |
| Sensitivity | The number of subjects experiencing each experimental condition divided by the variance of scores in the sample |
| Significance level | The probability that a difference has been erroneously declared to be significant, typically 0.005 and 0.001 corresponding to 5 and 1% chance of error |
| Type I error (false positives) | Concluding that there is an effect when there really is not an effect. Its probability is the α level |
| Type II error (false negatives) | Concluding there is an effect when there really is an effect. Its probability is the β level |

that is measured in the experimental subjects of a study, such as body weights. These identifiers (i.e., treatment and effect) are termed variables. Our treatment variables (those that the researcher or nature control and that can be directly controlled) are termed independent variables, while our effect variables (such as weight, life span, and number of neoplasms) are termed dependent variables—their outcome is believed to be dependent on the “treatment” being studied.

All the possible measures of a given set of variables in all the possible subjects that exist are termed the population for those variables. Such a population of variables cannot be truly measured—for example, one would have to obtain, treat, and measure the weights of all the Fischer 344 rats that were, are, or ever will be. Instead, we deal with a representative group—a sample. If our sample of data is appropriately collected and of sufficient size, it serves to provide good estimates of the characteristics of the parent population from which it was drawn.

30.1.1 Bias and Chance

Any toxicological study aims to determine whether a treatment elicits a response. An observed difference in response between a treated and control group need not necessarily be a result of treatment. There are, in principle, two other possible explanations—*bias*, or systematic differences other than treatment between the groups, and *chance*, or random differences. A major objective of both experimental design and analysis is to try to avoid bias. Wherever possible, treated and control groups to be compared should be alike in respect of all other factors. Where differences remain, these should be corrected for in the statistical analysis. Chance cannot be wholly

excluded, since identically treated animals will not respond identically. While even the most extreme difference might in theory be due to chance, a proper statistical analysis will allow the experimenter to assess this possibility. The smaller the probability of a “false positive,” the more confident the experimenter can be that the effect is real. Good experimental design improves the chance of picking up a true effect with confidence by maximizing the ratio between “signal” and “noise.”

30.1.2 Hypothesis Testing and Probability (p) Values

A relationship of treatment to some toxicological end point is often stated to be “statistically significant ($p < 0.05$).” What does this really mean? A number of points have to be made. *First*, statistical significance need not necessarily imply biological importance, if the end point under study is not relevant to the animal’s well-being. *Second*, the statement will usually be based only on the data from the study in question and will not take into account prior knowledge. In some situations, for example, when one or two of a very rare tumor type are seen in treated animals, statistical significance may not be achieved, but the finding may be biologically extremely important, especially if a similar treatment was previously found to elicit a similar response. *Third*, the p value does not describe the probability that a true effect of treatment exists. Rather, it describes the probability of the observed response, or one more extreme, occurring on the assumption that treatment actually had no effect whatsoever. A p value that is not significant is consistent with a treatment having a small effect, not detected with sufficient certainty in this study. *Fourth*, there are two types of p value. A “one-tailed

(or one-sided) p value” is the probability of getting by chance a treatment effect in a specified direction as great as or greater than that observed. A “two-tailed p value” is the probability of getting, by chance alone, a treatment difference in either direction which is as great as or greater than that observed. By convention p values are assumed to be two-tailed unless the contrary is stated. Where, which is unusual, one can rule out in advance the possibility of a treatment effect except in one direction, a one-tailed p value should be used. Often, however, two-tailed tests are to be preferred, and it is certainly not recommended to use one-tailed tests and *not* report large differences in the other direction. In any event, it is important to make it absolutely clear whether one- or two-tailed tests have been used.

It is a great mistake, when presenting results of statistical analyses, to mark, as do some laboratories, results simply as significant or not significant at one defined probability level (usually $p < 0.05$). This poor practice does not allow the reader any real chance to judge whether or not the effect is a true one. Some statisticians present the actual p value for every comparison made. While this gives precise information, it can make it difficult to assimilate results from many variables. One practice I recommend is to mark p values routinely using plus signs to indicate positive differences (and minus signs to indicate negative differences) as follows: $+++$, $p = 0.001$, $++$, $0.001 \leq p < 0.01$, $+$, $0.01 \leq p < 0.05$, \pm , $0.05 \leq p < 0.1$. This highlights significant results more clearly and also allows the reader to judge the whole range from “virtually certain treatment effect” to “some suspicion.” Note that using two-tailed tests, bracketed plus signs indicate findings that would be significant at the conventional $p < 0.05$ level using one-tailed tests but are not significant at this level using two-tailed tests. This “fiducial limit” ($p < 0.05$) implies a false-positive incidence of 1 in 20 and, though now imbedded in regulation, practice, and convention, was somewhat an arbitrary choice to begin with. In interpreting p values it is important to realize they are only an aid to judgment to be used in conjunction with other available information. One might validly consider a $p < 0.01$ increase as chance when it was unexpected, occurred only at a low-dose level with no such effect seen at higher doses, and was evident in only one subset of the data. In contrast, a $p < 0.05$ increase might be convincing if it occurred in the top dose and was for an end point one might have expected to be increased from known properties of the chemical or closely related chemicals.

30.1.3 Multiple Comparisons

When a p value is stated to be less than 0.05, this implies that, for that particular test, the difference could have occurred by chance less than 1 time in 20. Toxicological studies frequently involve making treatment–control comparisons for a large number of variables and, in some situations, also for

various subsets of animals. Some statisticians worry that the larger the number of tests, the greater is the chance of picking up statistically significant findings that do not represent true treatment effects. For this reason, an alternative “multiple comparison” procedure has been proposed in which, if the treatment was totally without effect, 19 times out of 20 *all* the tests should show nonsignificance when testing at the 95% confidence level. Automatic use of this approach cannot be recommended. Not only does it make it much more difficult to pick up any real effects, but also there is something inherently unsatisfactory about a situation where the relationship between a treatment and a particular response depends arbitrarily on which other responses happened to be investigated at the same time. It is accepted that in any study involving multiple end points that there will inevitably be a gray area between those showing highly significant effects and those showing no significant effects, where there is a problem distinguishing chance and true effects. However, changing the methodology so that the gray areas all come up as nonsignificant can hardly be the answer.

30.1.4 Estimating the Size of the Effect

It should be clearly understood that a p value does not give direct information about the size of any effect that has occurred. A compound may elicit an increase in response by a given amount, but whether a study finds this increase to be statistically significant will depend on the size of the study and the variability of the data. In a small study, a large and important effect may be missed, especially if the end point is imprecisely measured. In a large study, on the other hand, a small and unimportant effect may emerge as statistically significant.

Hypothesis testing tells us whether an observed increase can or cannot be reasonably attributed to chance, but not how large it is. Although much statistical theory relates to hypothesis testing, current trends in medical statistics are toward confidence interval (CI) estimation with differences between test and control groups expressed in the form of a best estimate, coupled with the 95% CI. Thus, if one states that treatment increases response by an estimated 10 units (95% CI 3–17 units), this would imply that there is a 95% chance that the indicated interval includes the true difference. If the lower 95% confidence limit exceeds zero, this implies the increase is statistically significant at $p < 0.05$ using a two-tailed test. One can also calculate, for example, 99 or 99.9% confidence limits, corresponding to testing for significance at $p < 0.01$ or $p < 0.001$.

In screening studies of standard design, the tendency has been to concentrate mainly on hypothesis testing. However, presentation of the results in the form of estimates with CI can be a useful adjunct for some analyses and is very important in studies aimed specifically at quantifying the size of an effect.

Two terms refer to the quality and reproducibility of our measurements of variables. The first, accuracy, is an expression of the closeness of a measured or computed value to its actual or “true” value in nature. The second, precision, reflects the closeness or reproducibility of a series of repeated measurements of the same quantity.

If we arrange all of our measurements of a particular variable in order as a point on an axis marked as to the values of that variable, and if our sample were large enough, the pattern of distribution of the data in the sample would begin to become apparent. This pattern is a representation of the frequency distribution of a given population of data—that is, of the incidence of different measurements, their central tendency, and dispersion.

The most common frequency distribution—and one we will talk about throughout this book—is the normal (or Gaussian) distribution. The normal distribution is such that two-thirds of all values are within one standard deviation (SD) (to be defined in Section 30.2) of the mean (or average value for the entire population), and 95% are within 1.96 SD of the mean. Symbols used are μ for the mean and σ for the SD. Other common frequency distributions include the binomial, Poisson, and chi-square.

In all areas of biological research, optimal design and appropriate interpretation of experiments require that the researcher understand both the biological and technological underpinnings of the system being studied and of the data being generated. From the point of view of the statistician, it is vitally important that the experimenter both know and be able to communicate the nature of the data and understand its limitations. One classification of data types is presented in Table 30.3.

The nature of the data collected is determined by three considerations. These are the biological source of the data (the system being studied), the instrumentation and techniques being used to make measurements, and the design of the experiment. The researcher has some degree of control

over each of these—least over the biological system (he or she normally has a choice of only one of several models to study) and most over the design of the experiment or study. Such choices, in fact, dictate the type of data generated by a study.

Statistical methods are based on specific assumptions. Parametric statistics—those that are most familiar to the majority of scientists—have more stringent underlying assumptions than do nonparametric statistics. Among the underlying assumptions for many parametric statistical methods (such as the ANOVA) is that the data are continuous. The nature of the data associated with a variable (as described in the preceding text) imparts a “value” to that data, the value being the power of the statistical tests which can be employed.

Continuous variables are those which can at least theoretically assume any of an infinite number of values between any two fixed points (such as measurements of body weight between 2.0 and 3.0 kg). Discontinuous variables, meanwhile, are those which can have only certain fixed values, with no possible intermediate values (such as counts of five and six dead animals, respectively).

Limitations on our ability to measure constrain the extent to which the real-world situation approaches the theoretical, but many of the variables studied in toxicology are in fact continuous. Examples of these are lengths, weights, concentrations, temperatures, periods of time, and percentages. For these continuous variables, we may describe the character of a sample with measures of central tendency and dispersion that we are most familiar with—the mean, denoted by the symbol \bar{X} and also called the arithmetic average, and the SD, which is denoted by the symbol σ and is calculated as being equal to

$$\sqrt{\frac{\sum X^2 - (\sum X)^2 / N}{N - 1}}$$

where X is the individual datum and N is the total number of data in the group.

Contrasted with these continuous data, however, we have discontinuous (or discrete) data, which can only assume certain fixed numerical values. In these cases our choice of statistical tools or tests is, as we will find later, more limited.

TABLE 30.3 Types of Variables (Data) and Examples of Each Type

| Classified By | Type | Example |
|-------------------------------|-----------|---|
| <i>Scale</i> | | |
| Continuous | Scalar | Body weight |
| | Ranked | Severity of a lesion |
| Discontinuous | Scalar | Weeks until the first observation of a tumor in a carcinogenicity study |
| | Ranked | Clinical observations in animals |
| | Attribute | Eye colors in fruit flies |
| | Quantal | Dead/alive or present/absent |
| <i>Frequency Distribution</i> | | |
| | Normal | Body weights |
| | Bimodal | Some clinical chemistry parameters |
| | Others | Measures of time-to-incapacitation |

30.1.5 Functions of Statistics

Statistical methods may serve to do any combination of three possible tasks. The one we are most familiar with is hypothesis testing—that is, determining if two (or more) groups of data differ from each other at a predetermined level of confidence. A second function is the construction and use of models which may be used to predict future outcomes of chemical–biological interactions. This is most commonly

seen in linear regression or in the derivation of some form of correlation coefficient. Model fitting allows us to relate one variable (typically a treatment or “independent” variable) to another. The third function, reduction of dimensionality, continues to be less commonly utilized than the first two. This final category includes methods for reducing the number of variables in a system while only minimally reducing the amount of information, therefore making a problem easier to visualize and to understand. Examples of such techniques are factor analysis and cluster analysis. A subset of this last function, discussed later under descriptive statistics, is the reduction of raw data to single expressions of central tendency and variability (such as the mean and SD).

There is also a special subset of statistical techniques which is part of both the second and third functions of statistics. This is data transformation, which includes such things as the conversion of numbers to log or probit values.

30.1.6 Descriptive Statistics

Descriptive statistics are used to summarize the general nature of a data set. As such, the parameters describing any single group of data have two components. One of these describes the location of the data, while the other gives a measure of the dispersion of the data in and about this location. Often overlooked is the fact that the choice of which parameters are used to give these pieces of information implies a particular type of distribution for the data.

Most commonly, location is described by giving the (arithmetic) mean and dispersion by giving the SD or the standard error of the mean (SEM). The calculation of the first two of these has already been described. If we again denote the total number of data in a group as N , then the SEM would be calculated as

$$\text{SEM} = \frac{\text{SD}}{\sqrt{N}}$$

The use of the mean with either the SD or SEM implies, however, that we have reason to believe that the sample of data being summarized is from a population which is at least approximately normally distributed. If this is not the case, then we should rather use a set of statistical descriptions which do not require a normal distribution. These are the median, for location, and the semiquartile distance, for a measure of dispersion. These somewhat less familiar parameters are characterized as follows.

Median When all the numbers in a group are arranged in a ranked order (i.e., from smallest to largest), the median is the middle value. If there is an odd number of values in a group, then the middle value is obvious (in the case of 13 values, e.g., the seventh largest is the median). When the number of values in the sample is even, the median is calculated as the

midpoint between the $(N/2)$ th and the $((N/2)+1)$ th number. For example, in the series of numbers 7, 12, 13, and 19, the median value would be the midpoint between 12 and 13, which is 12.5.

The SD and the SEM are related to each other but yet are quite different.

The SEM is quite a bit smaller than the SD, making it very attractive to use in reporting data. This size difference is because the SEM actually is an estimate of the error (or variability) involved in measuring the means of samples, and not an estimate of the error (or variability) involved in measuring the data from which means are calculated. This is implied by the *central limit theorem*, which tells us three major things:

1. The distribution of sample means will be approximately normal regardless of the distribution of values in the original population from which the samples were drawn.
2. The mean value of the collection.
3. The SD of the collection of all possible means of samples of a given size, called the SEM, depends on both the SD of the original population and the size of the sample.

The SEM should be used only when the uncertainty of the estimate of the mean is of concern—which is almost never the case in toxicology. Rather, we are concerned with an estimate of the variability of the population—for which the SD is appropriate.

Semiquartile Distance When all the data in a group are ranked, a quartile of the data contains one ordered quarter of the values. Typically, we are most interested in the borders of the middle two quartiles Q_1 and Q_3 , which together represent the semiquartile distance and which contain the median as their center. Given that there are N values in an ordered group of data, the upper limit of the j th quartile (Q_j) may be computed as being equal to the $((jN/1)/4)$ th value. Once we have used this formula to calculate the upper limits of Q_1 and Q_3 , we can then compute the semiquartile distance (which is also called the quartile deviation and as such is abbreviated as the QD) with the formula $\text{QD} = (Q_3 - Q_1)/2$.

For example, for the 15-value data set 1, 2, 3, 4, 4, 5, 5, 5, 6, 6, 6, 7, 7, 8, 9, we can calculate the upper limits of Q_1 and Q_3 as

$$Q_1 = \frac{1(15+1)}{4} = \frac{16}{4} = 4$$

$$Q_3 = \frac{3(15+1)}{4} = \frac{48}{4} = 12$$

The 4th and 12th values in this data set are 4 and 7, respectively. The semiquartile distance can then be calculated as

$$\text{QD} = \frac{7-4}{2} = 1.5$$

There are times when it is desired to describe the relative variability of one or more sets of data. The most common way of doing this is to compute the coefficient of variation (CV), which is calculated simply as the ratio of the SD to the mean or

$$CV = \frac{SD}{\bar{X}}$$

A CV of 0.2 or 20% thus means that the SD is 20% of the mean. In toxicology the CV is frequently between 20 and 50% and may at times exceed 100%.

30.2 EXPERIMENTAL DESIGN

Toxicological experiments generally have a twofold purpose. The first question is whether or not an agent results in an effect on a biological system. The second question, never far behind, is how much of an effect is present. It has become increasingly desirable that the results and conclusions of studies aimed at assessing the effects of environmental agents be as clear and unequivocal as possible. It is essential that every experiment and every study yield as much information as possible and that the results of each study have the greatest possible chance of answering the questions it was conducted to address. The statistical aspects of such efforts, so far as they are aimed at structuring experiments to maximize the possibilities of success, are called experimental design (Cochran and Cox, 1975).

The four basic statistical principles of experimental design are replication, randomization, concurrent ("local") control, and balance. In abbreviated form, these may be summarized as follows.

Replication Any treatment must be applied to more than one experimental unit (animal, plate of cells, litter of offspring, etc.). This provides more accuracy in the measurement of a response than can be obtained from a single observation, since underlying experimental errors tend to cancel each other out. It also supplies an estimate of the experimental error derived from the variability among each of the measurements taken (or "replicates"). In practice, this means that an experiment should have enough experimental units in each treatment group (i.e., a large enough "*N*") so that reasonably sensitive statistical analysis of data can be performed. The estimation of sample size is addressed in detail later in this chapter.

Randomization This is practiced to ensure that every treatment shall have its fair share of extreme high and extreme low values. It also serves to allow the toxicologist to proceed as if the assumption of "independence" is valid. That is, there is no avoidable (known) systematic bias in how one obtains data.

Concurrent Control Comparisons between treatments should be made to the maximum extent possible between

experimental units from the same closely defined population. Therefore, animals used as a "control" group should come from the same source, lot, age, and so on as test group animals. Except for the treatment being evaluated, test and control animals should be maintained and handled in exactly the same manner.

Balance If the effect of several different factors is being evaluated simultaneously, the experiment should be laid out in such a way that the contributions of the different factors can be separately distinguished and estimated. There are several ways of accomplishing this using one of several different forms of design, as will be discussed in the succeeding text.

There are 10 facets of any study which may affect its ability to detect an effect of a treatment. The first six concern minimizing the role of chance and the last four relate to avoidance of bias.

30.2.1 Choice of Species and Strain

Ideally, the responses of interest should be rare in untreated control animals but should be reasonably readily evoked by appropriate treatments. Some species or specific strains, perhaps because of inappropriate diets (Roe, 1989; Gad, 2015), have high background tumor incidences which make increases both difficult to detect and difficult to interpret when detected.

30.2.2 Sampling

Sampling—the selection of which individual data points will be collected, whether in the form of selecting which animals to collect blood from or to remove a portion of a diet mix from for analysis—is an essential step upon which all other efforts toward a good experiment or study are based.

There are three assumptions about sampling which are common to most of the statistical analysis techniques that are used in toxicology. These are that the sample is collected without bias, that each member of a sample is collected independently of the others, and that members of a sample are collected with replacements. Precluding bias, both intentional and unintentional, means that at the time of selection of a sample to measure, each portion of the population from which that selection is to be made has an equal chance of being selected. Ways of precluding bias are discussed in detail in the section on experimental design, Section 30.2.

Independence means that the selection of any portion of the sample is not affected by and does not affect the selection or measurement of any other portion.

Finally, sampling with replacement means that in theory, after each portion is selected and measured, it is returned to the total sample pool and thus has the opportunity to be selected again. This is a corollary of the assumption of independence. Violation of this assumption (which is almost

always the case in toxicology and all the life sciences) does not have serious consequences if the total pool from which samples are sufficiently large (say, 20 or greater) is such that the chance of reselecting that portion is small anyway.

There are four major types of sampling methods—random, stratified, systematic, and cluster. Random is by far the most commonly employed method in toxicology. It stresses the fulfillment of the assumption of avoiding bias. When the entire pool of possibilities is mixed or randomized (procedures for randomization are presented in a later chapter), then the members of the group are selected in the order that is drawn from the pool.

Stratified sampling is performed by first dividing the entire pool into subsets or strata and then doing randomized sampling from each stratum. This method is employed when the total pool contains subsets which are distinctly different but in which each subset contains similar members. An example is a large batch of a powdered pesticide in which it is desired to determine the nature of the particle-size distribution. Larger pieces or particles are on the top, while progressively smaller particles have settled lower in the container, and at the very bottom, the material has been packed and compressed into aggregates. To determine a timely representative answer, proportionally sized subsets from each layer or stratum should be selected, mixed, and randomly sampled. This method is used more commonly in diet studies.

In systematic sampling, a sample is taken at set intervals (such as every fifth container of reagent or taking a sample of water from a fixed sample point in a flowing stream every hour). This is most commonly employed in quality assurance or (in the clinical chemistry lab) in quality control.

In cluster sampling, the pool is already divided into numerous separate groups (such as bottles of tablets), and we select small sets of groups (such as several bottles of tablets) and then select a few members from each set. What one gets then is a cluster of measures. Again, this is a method most commonly used in quality control or in environmental studies when the effort and expense of physically collecting a small group of units are significant.

In classical toxicology studies sampling arises in a practical sense in a limited number of situations. The most common of these are as follows:

1. Selecting a subset of animals or test systems from a study to make some measurement (which either destroys or stresses the measured system or is expensive) at an interval during a study. This may include such cases as doing interim necropsies in a chronic study or collecting and analyzing blood samples from some animals during a subchronic study.
2. Analyzing inhalation chamber atmospheres to characterize aerosol distributions with a new generation system.

3. Analyzing diet in which test material has been incorporated.
4. Performing quality control on an analytical chemistry operation by having duplicate analyses performed on some materials.
5. Selecting data to audit for quality assurance purposes.

30.2.3 Dose Levels

This is a very important and controversial area. In screening studies aimed at hazard identification, it is normal, in order to avoid requiring huge numbers of animals, to test at dose levels higher than those to which man will be exposed, but not so high that marked toxicity occurs. A range of doses are usually tested to guard against the possibility of a misjudgment of an appropriate high dose and that the metabolic pathways at the high doses differ markedly from those at lower doses and, perhaps, to ensure no large effects occur at dose levels in the range to be used by man. In studies aimed more at risk estimation, more and lower doses may be tested to obtain fuller information on the shape of the dose-response curve.

30.2.4 Number of Animals

This is obviously an important determinant of the precision of the findings. The calculation of the appropriate number depends on (i) the critical difference, that is, the size of the effect it is desired to detect; (ii) the false-positive rate, that is, the probability of an effect being detected when none exists (equivalent to the “ α level” or “type I error”); (iii) the false-negative rate, that is, the probability of no effect being detected when one of exactly the critical size exists (equivalent to the “ β level” or “type II error”); and (iv) some measure of the variability in the material.

Tables relating numbers of animals required to obtain values of critical size, α and β are given in Kraemer and Thiemann (1987) and Gad (2005) and software is also available for this purpose. As a rule of thumb, to reduce the critical difference by a factor n for a given α and β , the number of animals required will have to increase by a factor n^2 .

30.2.5 Duration of the Study

It is obviously important not to terminate the study too early for fatal conditions, which are normally strongly age-related. Less obviously, going on for too long in a study can be a mistake, partly because the last few weeks or months may produce relatively few extra data at a disproportionate cost and partly because diseases of extreme old age may obscure the detection of tumors and other conditions of more interest. For nonfatal conditions, the ideal is to sacrifice the animals when the average prevalence is around 50%.

30.2.6 Stratification

To detect a treatment difference with accuracy, it is important that the groups being compared are as homogeneous as possible with respect to other known causes of the response. In particular, suppose that there is another known important cause of the response for which the animals vary, so that the animals are a mixture of hyper- and hyporesponders from this cause. If the treated group has a higher proportion of hyperresponders, it will tend to have a higher response even if treatment has no effect. Even if the proportion of hyperresponders is the same as in the controls, it will be more difficult to detect an effect of treatment because of the increased between-animal variability.

Given that this other factor is known, it will be sensible to take it into account in both the design and analysis of the study. In the design, it can be used as a “blocking factor” so that animals at each level are allocated equally (or in the correct proportion) to control and treated groups. In the analysis, the factor should be treated as a stratifying variable, with separate treatment–control comparisons made at each level, and the comparisons combined for an overall test of difference. This is discussed later, where we refer to the factorial design as one example of the more complex designs that can be used to investigate the separate effect of multiple treatments.

30.2.7 Randomization

Random allocation of animals to treatment groups is a prerequisite of good experimental design. If not carried out, one can never be sure whether treatment–control differences are due to treatment or to “confounding” by other relevant factors. The ability to randomize easily is a major advantage animal experiments have over epidemiology.

While randomization eliminates bias (as least in expectation), simple randomization of all animals may not be the optimal technique for producing a sensitive test. If there is another major source of variation (e.g., sex or batch of animals), it will be better to carry out stratified randomization (i.e., carry out separate randomizations within each level of the stratifying variable).

The need for randomization applies not only to the allocation of the animals to the treatment but also to anything that can materially affect the recorded response. The same random number that is used to apply animals to treatment group can be used to determine cage position, order of weighing, order of bleeding for clinical chemistry, order of sacrifice at terminations, and so on.

30.2.8 Adequacy of Control Group

While historical control data can, on occasion, be useful, a properly designed study demands that a relevant concurrent

control group be included with which results for the test group can be compared. The principle that like should be compared with like, apart from treatment, demands that control animals should be randomized from the same source as treatment animals. Careful consideration should also be given to the appropriateness of the control group. Thus, in an experiment involving treatment of a compound in a solvent, it would often be inappropriate to include only an untreated control group, as any differences observed could only be attributed to the treatment–solvent combination. To determine the specific effects of the compound, a comparison group given the solvent only, by the same route of administration, would be required.

It is not always generally realized that the position of the animal in the room in which it is kept may affect the animal’s response. An example is the strong relationship between incidence of retinal atrophy in albino rats and closeness to the lighting source. Systematic differences in cage position should be avoided, preferably via randomization.

We have now become accustomed to developing exhaustively detailed protocols for an experiment or study prior to its conduct. *A priori* selection of statistical methodology (as opposed to the *post hoc* approach) is as significant a portion of the process of protocol development and experimental design as any other and can measurably enhance the value of the experiment or study. Prior selection of statistical methodologies is essential for proper design of other portions of a protocol such as the number of animals per group or the sampling intervals for body weight. Implied in such a selection is the notion that the toxicologist has both an in-depth knowledge of the area of investigation and an understanding of the general principles of experimental design, for the analysis of any set of data is dictated to a large extent by the manner in which the data are obtained.

A second concept and its understanding are essential to the design of experiments in toxicology, that of censoring. Censoring is the exclusion of measurements from certain experimental units, or indeed of the experimental units themselves, from consideration in data analysis or inclusion in the experiment at all. Censoring may occur either prior to initiation of an experiment (where, in modern toxicology, this is almost always a planned procedure), during the course of an experiment (when they are almost universally unplanned, such as resulting from the death of animals on test), or after the conclusion of an experiment (when usually data are excluded because of being identified as some form of outlier).

In practice, *a priori* censoring in toxicology studies occurs in the assignment of experimental units (such as animals) to test groups. The most familiar example is in the common practice of assignment of test animals to acute, subacute, subchronic, and chronic studies, where the results of otherwise random assignments are evaluated for body weights of the assigned members. If the mean weights are

found not to be comparable by some preestablished criterion (such as a 90% probability of difference by ANOVA), then members are reassigned (censored) to achieve comparability in terms of starting body weights. Such a procedure of animal assignment to groups is known as a *censored randomization*.

The first precise or calculable aspect of experimental design encountered is determining sufficient test and control group sizes to allow one to have an adequate level of confidence in the results of a study (i.e., in the ability of the study design with the statistical tests used to detect a true difference—or effect—when it is present). The statistical test contributes a level of power to such detection. Remember that the power of a statistical test is the probability that a test results in rejection of a hypothesis, H_0 , say, when some other hypothesis, H , say, is valid. This is termed the power of the test “with respect to the (alternative) hypothesis H .”

If there is a set of possible alternative hypotheses, the power, regarded as a function of H , is termed the *power function* of the test. When the alternatives are indexed by a single parameter θ , simple graphic presentation is possible. If the parameter is a vector θ , one can visualize a *power surface*.

If the power function is denoted by $\beta(\theta)$ and H_0 specifies $\theta = \theta_0$, then the value of $\beta(\Pi)$ —the probability of rejecting H_0 when it is in fact valid—is the significance level. A test’s power is greatest when the probability of a type II error is the least. Specified powers can be calculated for tests in any specific or general situation.

Some general rules to keep in mind are:

- The more stringent the significance level, the greater the necessary sample size. More subjects are needed for a 1% level test than for a 5% level test.
- Two-tailed tests require larger sample sizes than one-tailed tests. Assessing two directions at the same time requires a greater investment.
- The smaller the critical effect size, the larger the necessary sample size. Subtle effects require greater efforts.
- Any difference can be significant if the sample size is large enough.
- The larger the power required, the larger the necessary sample size. Greater protection from failure requires greater effort. The smaller the sample size, the smaller the power, that is, the greater the chance of failure.
- The requirements and means of calculating necessary sample size depend on the desired (or practical) comparative sizes of test and control groups.

This number (N) can be calculated, for example, for equal-sized test and control groups, using the formula

$$N = \frac{(t_1 + t_2)^2}{d^2} S$$

where t_1 is the one-tailed t value with $N - 1$ degrees of freedom (df) corresponding to the desired level of confidence, t_2 is the one-tailed t value with $N - 1$ df corresponding to the probability that the sample size will be adequate to achieve the desired precision, variable “ d ” is the meaningful change in variable, and S is the sample SD, derived typically from historical data and calculated as

$$S = \sqrt{\frac{1}{N-1} \sum (V_1 - V_2)^2}$$

V_1 and V_2 are the highest and lowest historical data points. There are a number of aspects of experimental design which are specific to the practice of toxicology. Before we look at a suggestion for step-by-step development of experimental designs, these aspects should first be considered as follows:

1. Frequently, the data gathered from specific measurements of animal characteristics are such that there is wide variability in the data. Often, such wide variability is not present in a control or low-dose group, but in an intermediate dosage group, variance inflation may occur. That is, there may be a large SD associated with the measurements from this intermediate group. In the face of such a set of data, the conclusion that there is no biological effect based on a finding of no statistically significant effect might well be erroneous.
2. In designing experiments, one should keep in mind that potential effect of involuntary censoring on sample size. In other words, though a study might start with five dogs per group, this provides no margin should any die before the study is ended, and blood samples are collected and analyzed. Just enough experimental units per group frequently leave too few at the end to allow meaningful statistical analysis, and allowances should be made accordingly in establishing group sizes.
3. It is certainly possible to pool the data from several identical toxicological studies. One approach to this is meta-analysis, considered in detail later in this chapter. For example, after first having performed an acute inhalation study where only three treatment group animals survived to the point at which a critical measure (such as analysis of blood samples) was performed, we would not have enough data to perform a meaningful statistical analysis. We could then repeat the protocol with new control and treatment group animals from the same source. At the end, after assuring ourselves that the two sets of data are comparable, we could combine (or pool) the data from survivors of the second study with those from the first. The costs of this approach, however, would then be both a greater degree of effort expended (than if we had performed a

single study with larger groups) and increased variability in the pooled samples (decreasing the power of our statistical methods).

4. Another frequently overlooked design option in toxicology is the use of an unbalanced design—that is, of different group sizes for different levels of treatment.

There is no requirement that each group in a study (control, low dose, intermediate dose, and high dose) have an equal number of experimental units assigned to it. Indeed, there are frequently good reasons to assign more experimental units to one group than to others, and, as we shall see later in this book, all the major statistical methodologies have provisions to adjust for such inequalities, within certain limits. The two most common uses of the unbalanced design have larger groups assigned to either the highest dose, to compensate for losses due to possible deaths during the study, or to the lowest dose to give more sensitivity in detecting effects at levels close to an effect threshold—or more confidence to the assertion that no effect exists.

5. We are frequently confronted with the situation where an undesired variable is influencing our experimental results in a nonrandom fashion. Such a variable is called a confounding variable—its presence, as discussed earlier, makes the clear attribution and analysis of effects at best difficult and at worst impossible. Sometimes such confounding variables are the result of conscious design or management decisions, such as the use of different instruments, personnel, facilities, or procedures for different test groups within the same study. Occasionally, however, such confounding variables are the result of unintentional factors or actions, in which it is called a lurking variable. Examples of such variables are almost always the result of standard operating procedures being violated—water not being connected to a rack of animals over a weekend, a set of racks not being cleaned as frequently as others, or a contaminated batch of feed being used.
6. Finally, some thought must be given to the clear definition of what is meant by experimental unit and concurrent control.

The experimental unit in toxicology encompasses a wide variety of possibilities. It may be cells, plates of microorganisms, individual animals, litters of animals, and so on. The importance of clearly defining the experimental unit is that the number of such units per group is the “*N*” which is used in statistical calculations or analyses and critically affects such calculations. The experimental unit is the unit which receives treatments and yields a response which is measured and becomes a datum.

A true concurrent control is one that is identical in every manner with the treatment groups except for

the treatment being evaluated. This means that all manipulations, including gavaging with equivalent volumes of vehicle or exposing to equivalent rates of air exchanges in an inhalation chamber, should be duplicated in control groups just as they occur in treatment groups.

The goal of the four principles of experimental design is statistical efficiency and the economizing of resources. The single most important initial step in achieving such an outcome is to clearly define the objective of the study—get a clear statement of what questions are being asked.

For the reader who would like to further explore experimental design, there are a number of more detailed texts available which include more extensive treatments of the statistical aspects of experimental design (Federer, 1955; Myers, 1972; Cochran and Cox, 1975; Diamond, 1981; Hicks, 1982; Kraemer and Thiemann, 1987).

There are four basic experimental design types used in toxicology. These are the randomized block, Latin square, factorial design, and nested design. Other designs that are used are really combinations of these basic designs and are very rarely employed in toxicology. Before examining these four basic types, however, we must first examine the basic concept of blocking.

Blocking is, simply put, the arrangement or sorting of the members of a population (such as all of an available group of test animals) into groups based on certain characteristics which may (but are not sure to) alter an experimental outcome. Such characteristics which may cause a treatment to give a differential effect include genetic background, age, sex, overall activity levels, and so on. The process of blocking then acts (or attempts to act), so that each experimental group (or block) is assigned its fair share of the members of each of these subgroups.

We should now recall that randomization is aimed at spreading out the effect of undetectable or unsuspected characteristics in a population of animals or some portion of this population. The merging of the two concepts of randomization and blocking leads to the first basic experimental design, the randomized block. This type of design requires that each treatment group have at least one member of each recognized group (such as age), the exact members of each block being assigned in an unbiased (or random) fashion.

The second type of experimental design assumes that we can characterize treatments (whether intended or otherwise) as belonging clearly to separate sets. In the simplest case, these categories are arranged into two sets which may be thought of as rows (for, say, source litter of test animal, with the first litter as row 1, the next as row 2, etc.) and the secondary set of categories as columns (for, say, our ages of test animals, with 6–8 weeks as column 1, 8–10 weeks as column 2, etc.). Experimental units are then assigned so that each major treatment (control, low dose, intermediate dose,

etc.) appears once and only once in each row and each column. If we denote our test groups as A (control), B (low), C (intermediate), and D (high), such an assignment would appear as in the table in the succeeding text:

| Source Litter | Age | | | |
|------------------|--------------|---------------|----------------|----------------|
| | 6–8 Weeks | 8–10 Weeks | 10–12 Weeks | 12–14 Weeks |
| 1 | A | B | C | D |
| 2 | B | C | D | A |
| 3 | C | D | A | B |
| 4 | D | A | B | C |

The third type of experimental design is the factorial design, in which there are two or more clearly understood treatments, such as exposure level to test chemical, animal age, or temperature. The classical approach to this situation (and to that described under the Latin square) is to hold all but one of the treatments constant and at any one time to vary just that one factor. Instead, in the factorial design all levels of a given factor are combined with all levels of every other factor in the experiment. When a change in one factor produces a different change in the response variable at one level of a factor than at other levels of this factor, there is an interaction between these two factors which can then be analyzed as an interaction effect.

The last of the major varieties of experimental design are the nested designs, where the levels of one factor are nested within (or are subsamples of) another factor. That is, each subfactor is evaluated only within the limits of its single larger factor.

30.3 DATA RECORDING

Two distinct sources of systematic bias may occur in data recording. One is that awareness of treatment may, consciously or subconsciously, affect the values recorded by the measurer. This can be avoided by organizing data recording so that observations are made blind. The second is that there is a systematic shift in the standard of measurement with time, coupled with a tendency for different groups to be measured at different time points. This is particularly important when a pathologist grades a lesion for severity and when the control and high-dose animals are read before the intermediate-dose animals. In some situations it may be necessary to reread all the slides blind and in random order to be sure that diagnostic drift is avoided (Gad and Taulbee, 1996).

Valid analysis cannot be conducted unless one can distinguish animals which were examined and did not have the relevant response and animals which were not examined. It can also be important clearly to identify why data are missing. Table 30.4 identifies some basic rules for effective design of data collection forms.

TABLE 30.4 Rules for Form Design and Preparation

1. Forms should be used when some form of repetitive data must be collected. They may be either paper or electronic
2. If only a few (two or three) pieces of data are to be collected, they should be entered into a notebook and not onto a form. This assumes that the few pieces are not a daily event, with the aggregate total of weeks/months/years ending up as lots of data to be pooled for analysis
3. Forms should be self-contained, but should not try to repeat the content of the SOPs or method descriptions
4. Column headings on forms should always specify the units of measurement and other details of entries to be made. The form should be arranged so that sequential entries proceed down a page, not across. Each column should be clearly labeled with a heading that identifies what is to be entered in the column. Any fixed part of entries (such as °C) should be in the column header
5. Columns should be arranged from left to right so that there is a logical sequential order to the contents of an entry as it is made. An example would be date/time/animal number/bodyweight/name of the recorder. The last item for each entry should be the name or unique initials of the individual who made the data entry
6. Standard conditions that apply to all the data elements to be recorded on a form or the columns of the form should be listed as footnotes at the bottom of the form
7. Entries of data on the form should not use more digits than are appropriate for the precision of the data being recorded
8. Each form should be clearly titled to indicate its purpose and use. If multiple types of forms are being used, each should have a unique title or number
9. Before designing the form, carefully consider the purpose for which it is intended. What data will be collected, how often, with what instrument, and by whom? Each of these considerations should be reflected in some manner on the form. Before designing the form, carefully consider the purpose for which it is intended. What data will be collected, how often, with what instrument, and by whom? Each of these considerations should be reflected in some manner on the form. Before designing the form, carefully consider the purpose for which it is intended. What data will be collected, how often, with what instrument, and by whom? Each of these considerations should be reflected in some manner on the form
10. Those things which are common/standard for all entries on the form should be stated as such once. These could include such things as instrument used, scale of measurement (°C, F, or K), or the location where the recording is made

30.4 GENERALIZED METHODOLOGY SELECTION

One approach for the selection of appropriate techniques to employ in a particular situation is to use a decision tree method. Figure 30.1 is a decision tree that leads to the choice of one of three other trees to assist in technique selection, with each of the subsequent trees addressing one of the three functions of statistics that were defined earlier in this chapter. Figure 30.2 is for the selection of hypothesis-testing procedures, Figure 30.3 for modeling procedures, and Figure 30.4 for reduction of dimensionality procedures. For the vast majority of situations, these trees will guide the user into the choice of the proper technique. The tests and terms in these trees will be explained subsequently.

30.5 STATISTICAL ANALYSIS: GENERAL CONSIDERATIONS

30.5.1 Variables to Be Analyzed

Although some pathologists still regard their discipline as providing qualitative rather than quantitative data, it is abundantly clear that pathology, when applied to routine screening of animal toxicity and carcinogenicity studies, has to be quantitative to at least some degree so that statistical inferences and statements can be made about possible treatment effects. Inevitably, there will be some descriptive text which will not be appropriate for statistical analysis. However, the main objective of the pathologist should be to provide

information on the presence or absence (with severity grade or size where appropriate) of a list of conditions, consistently recorded from animal to animal and classified by well-defined criteria, which can be validly used in a statistical assessment.

Given that statistical analysis is worth doing and data are available that would be analyzed, should one then analyze all the end points recorded? Some arguments have been put forward against analyzing all the end-point studies, none of which really hold water.

One argument is that some end points are not of interest. Perhaps the study is essentially a carcinogenicity study, so that nonneoplastic end points are not considered to be “background pathology” and almost *per se* unrelated to treatment. However, if the pathologist has gone to the trouble of recording the data, then surely, in general, they ought to be analyzed. The costs of the statistical analysis are much less than those of doing the study and the pathology. While one might justify failure to analyze nonneoplastic data where tumor analysis has already shown that the compound is clearly carcinogenic and no longer of market potential, the general rule ought to be to analyze everything that has been specifically investigated.

Another argument put forward against doing multiple analyses is that it may yield many chance significant p values that have to be considered and evaluated for biological significance in the context of the entire set of available data. The whole context of dose response, as summarized in Table 30.5, must be kept in mind. A detailed look at the data can only aid interpretation, provided that one is not hide-bound by the false argument that statistical significance

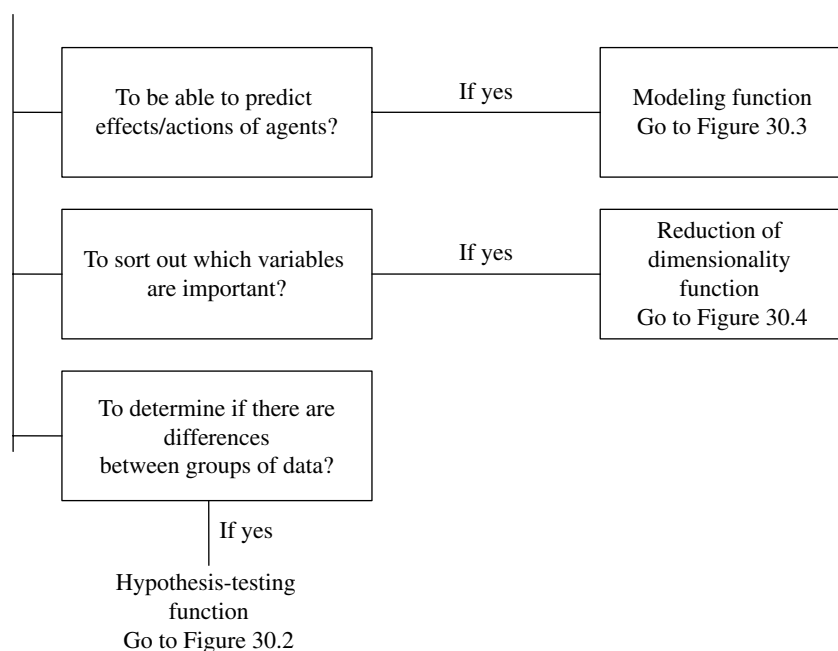


FIGURE 30.1 Overall decision tree for selecting statistical procedures.

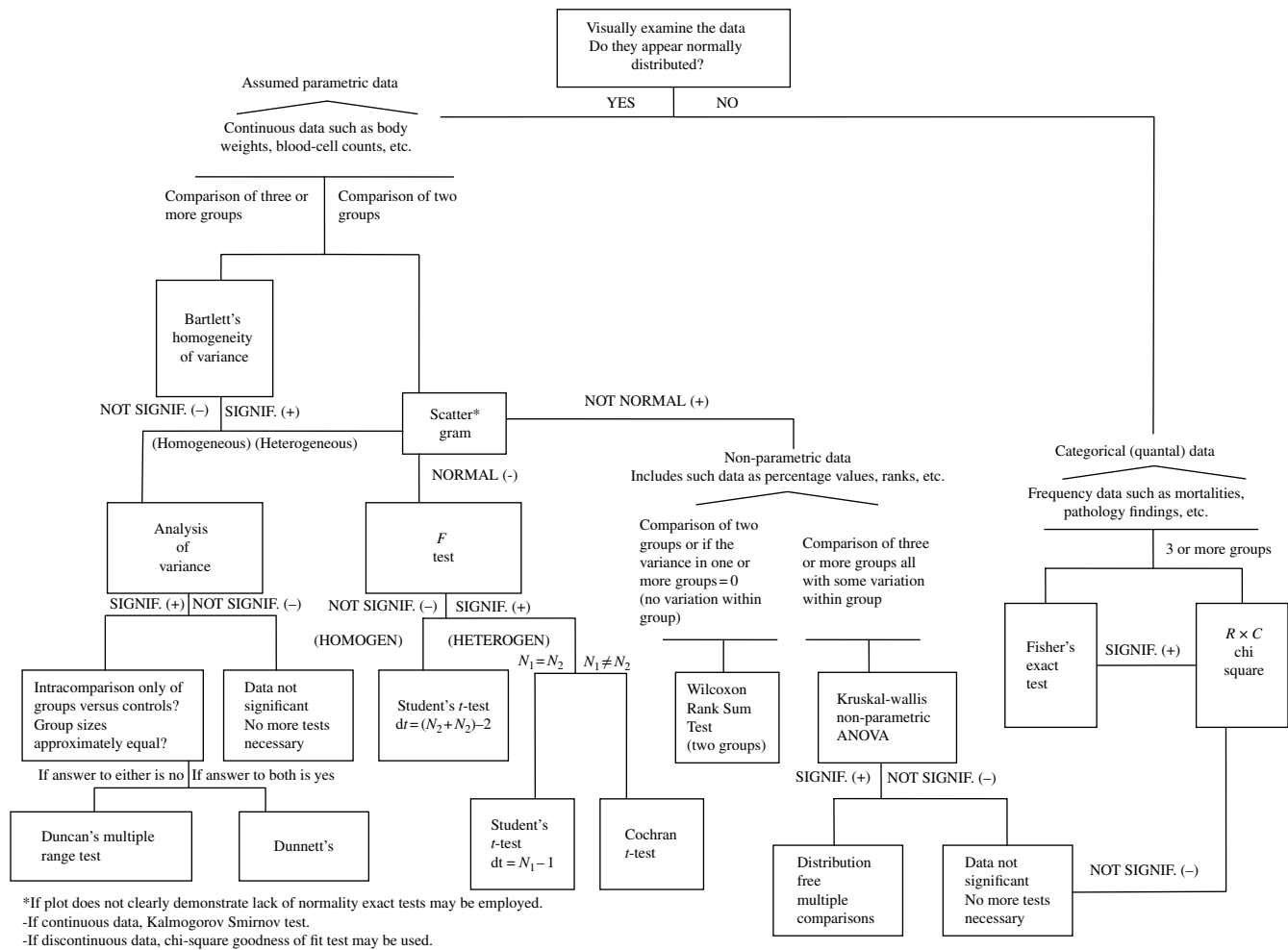


FIGURE 30.2 Decision tree for selecting hypothesis-testing procedures.

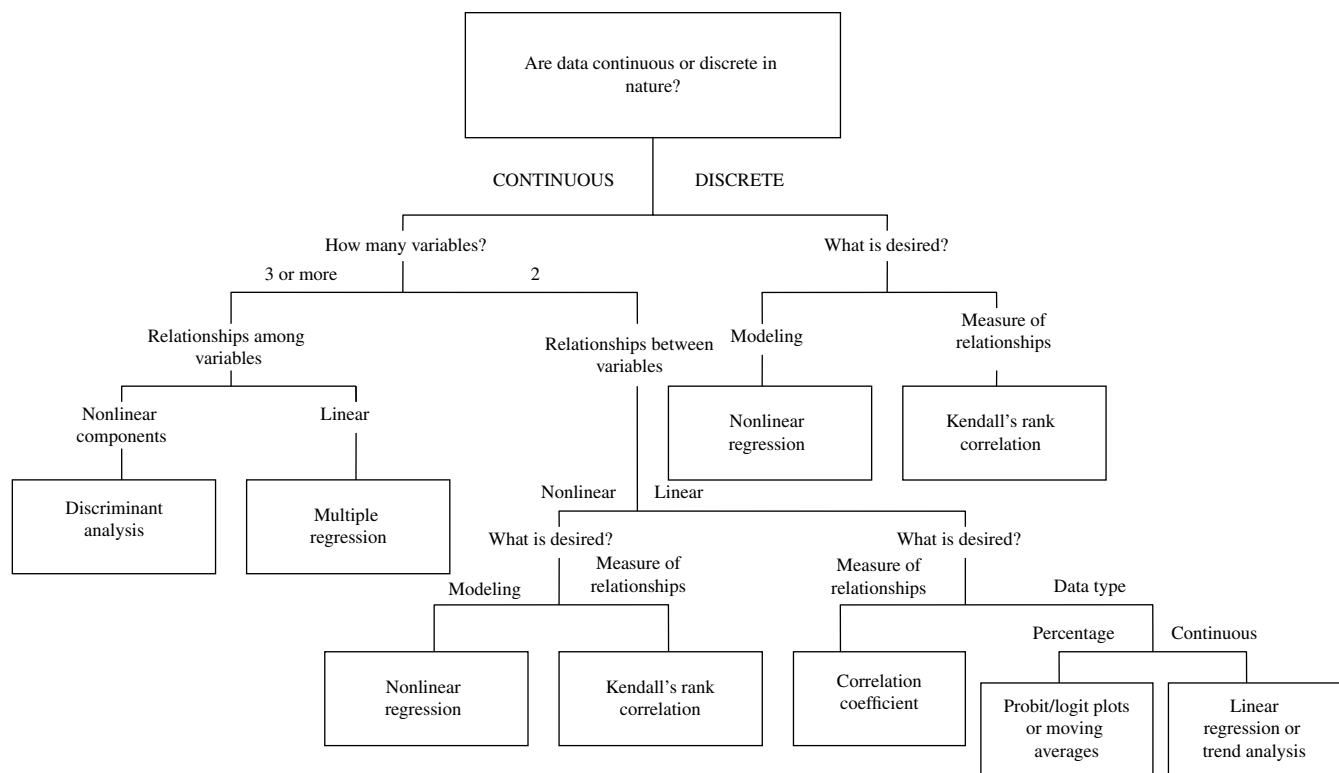


FIGURE 30.3 Decision tree for selecting modeling procedures.

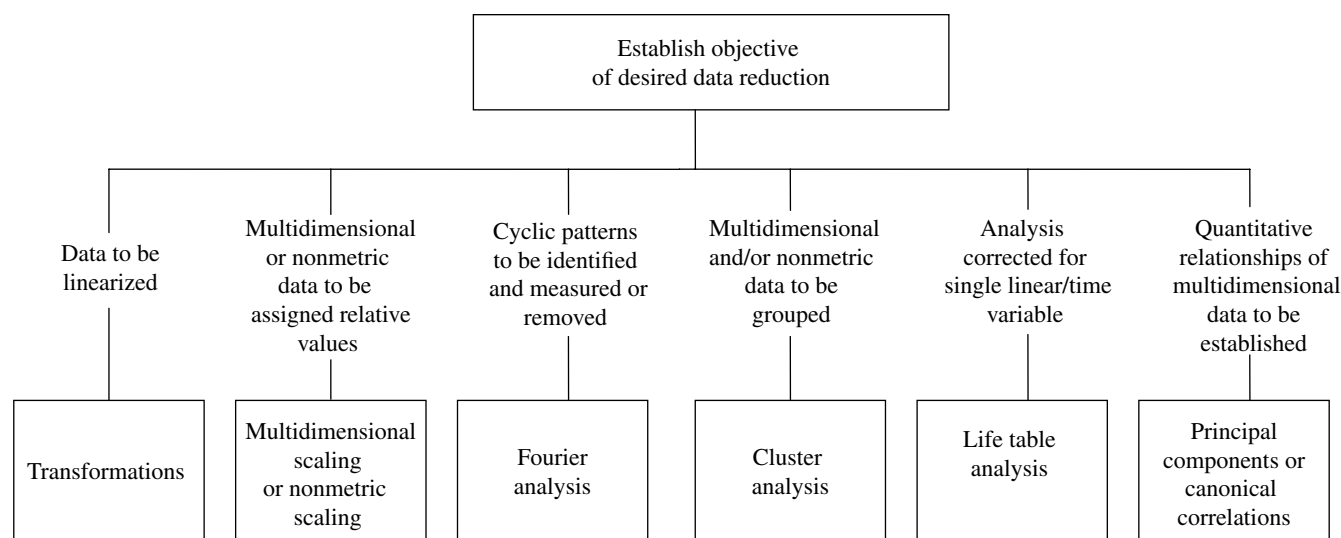


FIGURE 30.4 Decision tree for selection of reduction of dimensionality procedures.

TABLE 30.5 The Three Dimensions of Dose Response

As dose increases:

- Incidence of responders in an exposed population increases
- Severity of response in effected individuals increases
- Time to occurrence of response or of progressive stage of response decreases

necessarily equates with biological importance and definitely indicates a true effect of treatment.

Finally, some end points occur only very rarely. One must then be clear what “very rarely” is. For a typical study with a control and three dose groups of equal size, one would get a significant trend statistics if all three cases occurred at the top-dose level or in the control group (two-tailed $p \approx 0.03$), so a total of three cases will normally be enough for statistical analysis. End points occurring once or twice only are not worth analyzing formally, although, if only seen in the top-dose group, they may be worth noting in the report. This is especially true if they are lesions that are rarely reported.

30.5.2 Combination of Observations (Such as Pathological Conditions)

There are four main situations where one might consider combining pathological conditions in a statistical analysis.

The first is when essentially the same pathological condition has been recorded under two or more different names or even under the same name in different places. Here failure to combine these conditions in the analysis may severely limit the chances of detecting a true treatment effect. It should be noted, however, that grouping together conditions which are actually different may also result in the masking of a true treatment effect, particularly if the treatment has a very specific effect.

The second is when separately recorded pathological conditions form successive steps on the pathway of the same process. The most important example of this is for the incidence of related types of malignant tumor, benign tumor, and focal hyperplasia. It will normally be appropriate to carry out analyses of (i) incidence of malignant tumor, (ii) incidence of benign or malignant tumor, and, where appropriate, (iii) incidence of focal hyperplasia and benign or malignant tumor. It will not normally be appropriate to carry out analyses of benign tumor incidence only or of the incidence of hyperplasia only.

The third situation for combining is when the same pathological condition appears in different organs as a result of the same underlying process. Examples of this are the multicentric tumors (such as myeloid leukemia, reticulum cell sarcoma, and lymphosarcoma) or certain nonneoplastic conditions (such as arteritis/periarteritis and amyloid degeneration). Here analysis will normally be carried out only of incidence at any site, although in some situations site-specific analyses might be worth carrying out.

The final situation where an analysis of combined pathological conditions is normal is for analyses of overall incidence of malignant tumor at any site, of benign or malignant tumor at any site, or of multiple tumor incidence. While analyses of tumor incidence at specific sites are normally more meaningful, since treatments often affect only a few specific sites, these additional analyses are usually required to guard against the possibility that treatment had some weak but general tumor-enhancing effect not otherwise evident.

In some situations, one might also envisage analyses of other combinations of specific tumors, such as tumors at related sites (e.g., endocrine organs if the compound had a hormonal effect) or of similar histological type.

30.5.3 Taking Severity into Account

The same line of argument that suggests that if the pathologist records data they should be analyzed also suggests that if the pathologist chooses to grade a condition for severity, the grade should be taken into account in the analysis. There are two ways to carry out analysis when the grade has to be taken into account. In one, analyses are carried out not only of whether or not the animal has a condition but also of whether or not the condition is at least grade 2, at least grade 3, and so on. In the other approach, nonparametric (rank) methods are used. The latter approach is more powerful, as it uses all the information in one analysis, although the output may not be so easily understood by those without some statistical training.

Note that the analyses based on grade can be carried out only if grading has been consistently applied throughout. If a condition has been scored only as present/absent for some animals, but has been graded for others, it is not possible to carry out graded analyses unless the pathologist is willing to go back and grade the specific animals showing the condition.

30.5.4 Using Simple Methods Which Avoid Complex Assumptions

Different methods for statistical analysis can vary considerably in their complexity and in the number of assumptions they make. Although the use of statistical models has its place, more so for effect estimation than for hypothesis testing, and more so in studies of complex design than in those of simple design, there are advantages in using, wherever possible, statistical methods that are simple and robust and make as few assumptions as possible. There are three reasons for this. First, such methods are more generally understandable to the toxicologist. Second, there are hardly ever extensive enough data in practice to validate any given formal model fully. Third, even if a particular model were known to be appropriate, the loss of efficiency in using appropriate simpler methods is often only very small.

The methods we advocate for routine use for the analysis of tumor incidence tend, therefore, not to be based on the use of formal parametric statistical models. For example, when studying the relationship of treatment to incidence of a pathological condition and wishing to adjust for other factors (in particular, age at death) that might otherwise bias the comparison, methods involving “stratification” are recommended, rather than a multiple regression approach or time-to-tumor models. ANOVA methods can be useful in the case of continuously distributed data for estimating treatment effects. However, they involve underlying assumptions (normally distributed variables and variability equal in each group). If these assumptions are violated, nonparametric methods based on the rank of observations, rather than their actual value, may be preferable for hypothesis testing.

30.5.5 Using All the Data

Often information is available about the relationship between treatment and a condition of interest for groups of animals differing systematically in respect to some other factor. Obvious examples are males and females, differing times of sacrifice, and differing secondary treatments. While it will be necessary, in general, to look at the relationship within levels of this other factor, it will also be advisable to try to come to some assessment of the relationship over all levels of the other factor, and where a combined inference is not sensible, but in far more situations, this is not the case, and using all the data in one analysis allows a more powerful test of the relationship under study. Some scientists consider that conclusions for males and females should always be drawn separately, but there are strong statistical arguments for a joint analysis.

30.5.6 Combining, Pooling, and Stratification

Suppose, in a hypothetical study of a toxic agent which induces tumors that do not shorten the lives of tumor-bearing animals, the data regarding the number of animals with tumor out of the number examined are as follows:

| | Control | Exposed | Combined |
|--------------|--------------|--------------|--------------|
| Early deaths | 1/20 (5%) | 18/90 (20%) | 19/110 (17%) |
| Late deaths | 24/80 (30%) | 7/10 (70%) | 31/90 (34%) |
| Total | 25/100 (25%) | 25/100 (25%) | 50/200 (25%) |

It can be seen that if the time of death is ignored and the *pooled* data are studied, the incidence of tumors is the same in each group, resulting in the *false* conclusion that treatment had no effect. Looking within each time of death, however, an increased incidence in the exposed group can be seen. An appropriate statistical method would *combine* a measure of difference between the groups based on the early deaths and a measure of difference based on the late deaths and conclude *correctly* that incidence, after adjustment for time of death, is greater in the exposed groups.

In this example, time of death is the stratifying variable, with two strata—early deaths and late deaths. The essence of the methodology is to make comparisons only within strata (so that one is always comparing like with like except in respect to treatment) and then to combine the differences over strata. Stratification can be used to adjust for any variable or indeed combinations of variables.

Some studies are of factorial design, in which combinations of treatments are tested. The simplest such design is one in which four equal-sized groups of animals receive (i) no treatment, (ii) treatment A only, (iii) treatment B only, and (iv) treatments A and B. If one is prepared to assume that any effects of the two treatments are independent, one can

use stratification to enable more powerful tests to be conducted of the possible individual treatment effects. Thus, to test for effects of treatment A, for example, one conducts comparisons in two strata, the first consisting of groups 1 and 2 not given treatment B and the second consisting of groups 3 and 4 given treatment B. A combination of results from the two strata is based on twice as many animals and is therefore markedly more likely to detect possible effects of treatment A than a simple comparison of groups 1 and 2. There is also the possibility of identifying interactions, such as synergism and antagonism, between the two treatments.

In some routine long-term screening studies, the study design involved five groups of (usually) 50 animals of each sex, three of which are treated with successive doses of a compound and two of which are untreated controls. Assuming that there is no systematic difference between the control groups (e.g., the second control group in a different room or from a different batch of animals), it will be normal to carry out the main analyses with the control groups treated as a single group of 100 animals. It will usually be a sensible preliminary precaution to carry out additional analyses comparing incidences in the two control groups.

30.5.7 Trend Analysis, Low-Dose Extrapolation, and NOEL Estimation

While comparisons of individual treated groups with the control group are important, a more powerful test of a possible effect of treatment will be to carry out a test for a dose-related trend. This is because most true effects of treatment tend to result in a response which increases (or decreases) with increasing dose and because trend tests take into account all the data in a single analysis. In interpreting the results of trend tests, it should be noted that a significant trend does not necessarily imply an increased risk at lower doses nor, conversely, does a lack of increase at lower doses necessarily indicate evidence of a threshold (i.e., a dose below which no increase occurs).

Note that the testing for trend is seen as a more sensitive way of picking up a possible treatment effect than simple pairwise comparisons of treated and control groups. Attempting to estimate the magnitude of effects at low doses, typically below the lowest positive dose tested in the study, is a much more complex procedure and is heavily dependent on the assumed functional form of the dose-response relationship.

Deterministic trend models are based on the assumption that the trend of a time series can be approximated closely by simple mathematical functions of time over the entire span of the series. The most common representation of a deterministic trend is by means of polynomials or of transcendental functions. The time series from which the trend is to be identified is assumed to be generated by a nonstationary process where the nonstationarity results from a deterministic trend. A classical model is the regression

or error model (Anderson, 1971) where the observed series is treated as the sum of a systematic part or trend and a random part or irregular. This model can be written as

$$Z_i = Y_i + U_i'$$

where U_i is a purely random process, that is, $U_i \sim \text{i.i.d. } (0, \sigma_u^2)$ (independent and identically distributed with expected value 0 and variance σ_u^2).

Trend tests are generally described as “ k -sample tests of the null hypothesis of identical distribution against an alternative of linear order,” that is, if sample i has distribution function F_i , $i=1$, then the null hypothesis

$$H_0 : F_1 = F_2 = \cdots = F_k$$

is tested against the alternative

$$H_1 : F_1 \geq F_2 \geq \cdots = F_k$$

(or its reverse); at least one of the inequalities is strict. These tests can be thought of as special cases of tests of regression or correlation in which association is sought between the observations and its ordered sample index. They are also related to ANOVA except that the tests are tailored to be powerful against the subset of alternatives H_1 , instead of the more general set $\{F_i \neq F_j, \text{ some } i \neq j\}$.

Different tests arise from requiring power against specific elements or subsets of this rather extensive set of alternatives.

The most popular trend test in toxicology is currently that presented by Tarone in 1975 because it is used by the National Cancer Institute in the analysis of carcinogenicity data. A simple, but efficient alternative is the Cox and Stuart (1955) test which is a modification of the sign test. For each point at which we have a measure (such as the incidence of animals observed with tumors), we form a pair of observations—one from each of the groups we wish to compare. In a traditional NCI bioassay, this would mean pairing control with low dose and low dose with high dose (to explore a dose-related trend) or each time period observation in a dose group (except the first) with its predecessor (to evaluate time-related trend). When the second observation in a pair exceeds the earlier observation, we record a plus sign for that pair. When the first observation is greater than the second, we record a minus sign for that pair. A preponderance of plus signs suggests a downward trend, while an excess of minus signs suggests an upward trend. A formal test at a preselected confidence level can then be performed.

More formally put, after having defined what trend we want to test for, we first match pairs as $(X_1, X_{1+c}), (X_2, X_{2+c}), \dots, (X_{n'-c}, X_{n'})$ where $c=n'/2$ when n' is even and $c=(n'+1)/2$ when n' is odd (where n' is the number of observations in a set).

EXAMPLE 30.1

In a chronic feeding study in rats, we tested the hypothesis that, in the second year of the study, there was a dose-responsive increase in tumor incidence associated with the test compound. We utilize, in the succeeding text, a Cox–Stuart test for trend to address this question. All groups start the second year with an equal number of animals.

| Control | | | Low Doses | | | High Doses | | |
|----------------|------------------------------|----------------------|------------------------------|----------------------|-------------------------------|------------------------------|----------------------|-------------------------------|
| Month of Study | Total X, Animals with Tumors | Change (X_{A-B}) | Total Y, Animals with Tumors | Change (Y_{A-B}) | Compared to Control ($Y-X$) | Total Z, Animals with Tumors | Change (Z_{a-b}) | Compared to Control ($Z-X$) |
| 12 (A) | 1 | NA | 0 | NA | NA | 5 | NA | NA |
| 13 (B) | 1 | 0 | 0 | 0 | 0 | 7 | 2 | (+)2 |
| 14 (C) | 3 | 2 | 1 | 1 | (-)1 | 11 | 4 | (+)2 |
| 15 (D) | 3 | 0 | 1 | 0 | 0 | 11 | 0 | 0 |
| 16 (E) | 4 | 1 | 1 | 0 | (-)1 | 13 | 2 | (+)1 |
| 17 (F) | 5 | 1 | 3 | 2 | (+)1 | 14 | 1 | 0 |
| 18 (G) | 5 | 0 | 3 | 0 | 0 | 15 | 1 | (+)1 |
| 19 (H) | 5 | 0 | 5 | 2 | (+)2 | 18 | 3 | (+)3 |
| 20 (I) | 6 | 1 | 6 | 1 | 0 | 19 | 1 | 0 |
| 21 (J) | 8 | 2 | 7 | 1 | (-)1 | 22 | 3 | (+)1 |
| 22 (K) | 12 | 4 | 9 | 2 | (-)2 | 26 | 4 | 0 |
| 23 (L) | 14 | 2 | 12 | 3 | (+)1 | 28 | 2 | 0 |
| 24(M) | 18 | 4 | 17 | 5 | (+)1 | 31 | 3 | (-)1 |
| Sum of Signs | | | | | (+)4 | (+)6 | | |
| | | | | | (-)4 | (-)1 | | |
| Y-X=0 | | | | | Z-X=+5 | | | |
| NO TREND | | | | | | | | |

The hypothesis is then tested by comparing the resulting number of excess positive or negative signs against a sign test table such as those found in Beyer (1976).

We can, of course, combine a number of observations to allow ourselves to actively test for a set of trends, such as the existence of a trend of increasing difference between two groups of animals over a period of time. This is demonstrated in Example 30.1.

ASSUMPTIONS AND LIMITATIONS

1. Trend tests seek to evaluate whether there is a monotonic tendency in response to a change in treatment. That is, the dose–response direction is absolute—as dose goes up, the incidence of tumors increases. Thus the test loses power rapidly in response to the occurrences of “reversals”—for example, a low-dose group with a decreased tumor incidence. There are methods (Dykstra and Robertson, 1983) which “smooth the bumps” of reversals in long data series. In toxicology, however, most data series are short (i.e., there are only a few dose levels).

Tarone’s trend test is most powerful at detecting dose-related trends when tumor onset hazard functions are proportional to each other. For more power against other dose-related group differences, weighted versions of the statistic are also available (Breslow, 1984; Crowley and Breslow, 1984).

In 1985, the US *Federal Register* recommended that the analysis of tumor incidence data be carried out with a Cochran–Armitage (Cochran, 1954; Armitage, 1955) trend test. The test statistic of the Cochran–Armitage test is defined as this term:

$$T_{CA} = \sqrt{\frac{N}{(N-r)r}} \cdot \frac{\sum_{i=0}^k (R_i - (n_i/N)r) d_i}{\sqrt{\sum_{i=0}^k (n_i/N) d_i^2 - \left(\sum_{i=0}^k (n_i/N) d_i \right)^2}}$$

with dose scores d_i . Armitage’s test statistic is the square of this term (T_{CA}^2). The actual “test” is of a contingency table form. R = row number, in increments from i to N . As one-sided tests are carried out for an increase in tumor rates, the square is not considered. Instead, the aforementioned test statistic which is presented by Portier and Hoel (1984) is

used. This test statistic is asymptotically standard normal distributed. The Cochran–Armitage test is asymptotically efficient for all monotone alternatives (Tarone, 1975), but this result only holds asymptotically. And tumors are rare events, so the binominal proportions are small. In this situation approximations may become unreliable.

Therefore, exact tests which can be performed using two different approaches: conditional and unconditional are considered. In the first case, the total number of tumors r is regarded as fixed. As a result the null distribution of the test statistic is independent of the common probability p . The exact conditional null distribution is a multivariate hypergeometric distribution.

The unconditional model treats the sum of all tumors as a random variable. Then the exact unconditional null distribution is a multivariate binomial distribution. The distribution depends on the unknown probability.

Such low-dose extrapolation is typically only conducted for tumors believed to be caused by a genotoxic effect which some, but by no means all, scientists believe has no threshold. For other types of tumors and for many nonneoplastic end points, a threshold cannot be estimated directly from data at a limited number of dose levels; a no-observable-effect level (NOEL) can be estimated by finding the highest dose level at which there is no significant increase in effects.

A useful technique for determining if there is an effect of treatment on any toxicological parameter is the NOSTASOT method (Tukey et al., 1985; Antonello et al., 1993). This test is based on the principle that a possible toxicological effect of interest occurs with a normal dose response, that is, there is an increasing effect with increasing dosage. The data to be analyzed should be examined first to confirm that this principle is not violated. In this method, regression analysis is used to determine if there is an increased or decreased response in a parameter with increasing dosage. This method can be visualized as a plot of response versus dosage in which the analysis determines if the slope of the plotted line deviates significantly from zero.

This method can be used for essentially all parameters. Three analyses are performed—each with different spacing between dosage levels. The spacing in the first analysis is based on the arithmetic values of the dosage levels. The spacing in the second, referred to as the ordinal scaling, has equal spacing between dosage levels, that is, the control through high-dosage levels are assigned values of 0, 1, 2, and 3. In the third analysis, the log of the dosage level is used. Since the log of zero is impractical, the control group is assigned a value based on the spacing between the low- and middle-dosage levels according to a formula that assigns a log scale value to the control such that the ratio of the difference between the control and low-dose groups and the difference between the low- and middle-dose groups is equal both in absolute values and in log scale values. This places the control group at a reasonable distance from the low-dosage group. The lowest p

value among the three analyses—arithmetic, ordinal, and logarithmic—is taken as the p value of the overall analysis based on the assumption that, if there is a dosage-related effect, the method of analysis yielding the lowest value is the best model for that dosage response. A correction for the multiplicity of analyses can be applied. If none of the three analyses are significant at the 0.05 level, the analysis is complete, and the high-dosage level is referred to as the “no statistical significance of trend dose” or the NOSTASOT dose. If there is a significant trend through the high-dosage level, the data from the high-dosage level is deleted and the trend test repeated. This process is repeated until an NOSTASOT dose is determined. Effects at dosage levels above the NOSTASOT dose are then considered to be statistically significant.

There are two major benefits of the NOSTASOT method. One is that spurious statistically significant results only at the low- and/or middle-dosage levels are eliminated, resulting in a reduction in false positives. A second benefit is that in some cases there may be real effects at multiple dosage levels that at any single-dosage level are not statistically significant but will nevertheless result in a significant trend, thus providing increased sensitivity and reducing false negatives.

30.5.8 Need for Age Adjustment

Where there are marked differences in survival between treated groups, it is widely recognized that there is a need for an age adjustment (i.e., an adjustment for age at death or onset). This is illustrated in the example in the preceding text, where, because of the greater number of deaths occurring early in the treated group, the true effect of treatment disappears if no adjustment is made. Thus, a major purpose of age adjustment is to avoid bias.

It is not so generally recognized, however, that, even where there are no survival differences, age adjustment can increase the power to detect between-group differences. This is illustrated in the example in the succeeding text:

| | Control | Exposed |
|---------------|---------|---------|
| Early deaths | 0/20 | 0/20 |
| Middle deaths | 1/10 | 9/10 |
| Late deaths | 20/20 | 20/20 |
| Total | 21/50 | 29/50 |

Here treatment results in a somewhat earlier onset of a condition which occurs eventually in all animals. Failure to age-adjust will result in a comparison of 29/50 with 21/50, which is not statistically significant. Age adjustment will essentially ignore the early and late deaths, which contribute no comparative statistical information, and be based on the comparison of 9/10 with 1/10, which is statistically significant. Here age adjustment sharpens the contrast, rather

than avoiding bias, by avoiding diluting data capable of detecting treatment effects with data that are of little or no value for this purpose.

30.5.9 Need to Take Context of Observation into Account

It is now widely recognized that age adjustment cannot properly be carried out unless the context of observation is taken into account. There are three relevant contexts, the first two relating to the situation where the condition is only observed at death (e.g., an internal tumor) and the third where it can be observed in life (e.g., a skin tumor).

In the first context the condition is assumed to have caused the death of the animal, that is, to be *fatal*. Here the incidence rate for a time interval and a group is calculated by

$$\frac{\text{Number of animals dying in the interval because of the lesion}}{\text{Number of animals alive at the start of the interval}}$$

In the second context, the animal is assumed to have died of another cause, that is, the condition is *incidental*. Here the rate is calculated by:

$$\frac{\text{Number of animals dying in the interval with the lesion}}{\text{Total number of animals dying in the interval}}$$

In the third context, where the condition is *visible*, the rate is calculated by

$$\frac{\text{Number of animals getting the condition in the interval}}{\text{Number of animals without the condition at the start of the interval}}$$

A problem with the method of Peto et al. (1980), which takes context of observation into account, is that some pathologists are unwilling or feel unable to decide whether, in any given case, a condition is fatal or incidental. A number of points should be made here.

First, where there are marked survival differences, it may not be possible to conclude reliably whether a treatment is beneficial or harmful unless such a decision is made. This is well illustrated by the example of Peto et al. (1980), where assuming all pituitary tumors were fatal results in the (false) conclusion that *N*-nitrosodimethylamine (NDMA) was carcinogenic, while assuming they were all incidental resulted in the (false) conclusion that NDMA was protective. Using, correctly, the pathologist's best opinion as to which were, and which were not, likely to be fatal resulted in an analysis which (correctly) concluded NDMA had no effect. If the pathologist, in this case, had been unwilling to make a judgment as to fatality, believing it to be unreliable, no conclusion could have been reached. This state of affairs would, however, be a fact of life and *not* a position reached because an inappropriate statistical method was being used.

Although it will normally be a good routine for the pathologist to ascribe "factors contributory to death" for each animal that was not part of a scheduled sacrifice, it is in fact not strictly necessary to determine the context of observation for all conditions at the outset. An alternative strategy is to analyze under differing assumptions: (i) no cases fatal, (ii) all cases occurring in decedents fatal, and (iii) all cases of same defined severity occurring in decedents fatal, with, under each assumption, other cases incidental.

If the conclusion turns out the same under each assumption, or if the pathologist can say, on general grounds, that one assumption is likely to be a close approximation to the truth, it may not be necessary to know the context of observation for the condition in question for each individual animal. Using the alternative strategy might result in a saving of the pathologist's time by only having to make a judgment for a limited number of conditions where the conclusion seems to hinge on correct knowledge of the context of observation.

Finally, it should be noted that, although many nonneoplastic conditions observed at death are never causes of death, it is, in principle, as necessary to know the context of observation for nonneoplastic conditions as it is for tumors.

30.5.10 Experimental and Observational Units

In many situations, the animal is both the "experimental unit" and the "observational unit," but this is not always so. For determining treatment effects by the methods of the next section, it is important that each experimental unit provides only one item of data for analysis, as the methods all assume that individual data items are statistically independent. In many feeding studies, where the cage is assigned to a treatment, it is the cage, rather than the animal, that is the experimental unit. In histopathology, observations for a tissue are often based on multiple sections per animal, so that the section is the observational unit. Multiple observations per experimental unit should be combined in some suitable way into an overall average for that unit before analysis.

30.5.11 Missing Data

In many types of analysis, animals with missing data are simply removed from the analysis. There are, however, some situations where this can be an inappropriate thing to do. One situation is when carrying out an analysis of a condition that is assumed to have caused the death of the animal. Although an animal dying at week 83 for which the section was unavailable for microscopic examination cannot contribute to the group comparison at week 83, one knows that it did not die because of any condition in previous weeks, so it should contribute to the denominator of the calculations in all previous weeks.

Another situation is when histopathological examination of a tissue is not carried out unless an abnormality is seen post-mortem. In such an experiment one might have the following data for that tissue:

- Control group: 50 animals, 2 abnormal post-mortem, 2 examined microscopically, and 2 with tumor of specific type
- Treated group: 50 animals, 15 abnormal post-mortem, 15 examined microscopically, and 14 with tumor of specific type

Ignoring animals with no microscopic sections, one would compare $2/2=100\%$ with $14/15=93\%$ and conclude treatment nonsignificantly decreased incidence. This is likely to be a false conclusion, and it would be better here to compare the percentages of animals which had a postmortem abnormality which turned out to be a tumor, that is, $2/50=4\%$ with $14/50=28\%$. Unless some aspect of treatment made tumors much easier to detect post-mortem, one could then conclude that treatment did have an effect on tumor incidence.

Particular care has to be taken in studies where the procedures for histopathological examination vary by group. In a number of studies conducted in recent years, the protocol demands full microscopic examination of a given tissue list in decedents in all groups and in terminally killed controls in high-dose animals. In other animals, terminally killed low- and mid-dose animals, microscopic examination of a tissue is only conducted if the tissue is found to be abnormal post-mortem. Such a protocol is designed to save money, but leads to difficulty in comparing the treatment groups validly. Suppose, for example, responses in terminally killed animals are 8/20 in the controls and 3/3 (with 17 unexamined) in the low-dose and 5/6 (with 14 unexamined) in the mid-dose animals. Is one supposed to conclude that treatment at the low and mid doses increased response, based on a comparison of the proportions examined microscopically (40, 100, and 83%), or that it decreased response, based on the proportion of animals in the group (40, 15, and 25%)? It could well be that treatment had no effect but some small tumors were missed post-mortem. In this situation, a valid comparison can only be achieved by ignoring the low- and mid-dose groups when carrying out the comparison for the age stratum "terminal kill." This, of course, seems wasteful of data, but these are data that cannot be usefully used owing to the inappropriate protocol.

30.5.12 Use of Historical Control Data

In some situations, particularly where incidences are low, the results from a single study may suggest an effect of treatment on tumor incidence but be unable to demonstrate it conclusively. The possibility of comparing results in the

treated groups with those of control groups from other studies is then often raised. Thus, a nonsignificant incidence of 2 cases out of 50 in a treated group may seem much more significant if no cases have been seen in, say, 1000 animals representing controls from 20 similar studies. Conversely, a significant incidence of 5 cases out of 50 in a treated group as compared with 0 out of 50 in the study controls may seem far less convincing if many other control groups had incidences around 5 out of 50.

While not understating the importance of looking at historical control data, it must be emphasized that there are a number of reasons why variation between studies may be greater than variation within study. Differences in diet, in duration of the study, in intercurrent mortality, and in who the study pathologist is may all contribute. Statistical techniques that ignore this and carry out simple statistical tests of treatment incidence against a pooled control incidence may well give results that are seriously in error and are likely to overstate statistical significance considerably.

30.5.13 Methods for Data Examination and Preparation

The data from toxicology studies should always be examined before any formal analysis. Such examinations should be directed to determining if the data are suitable for analysis and if so what form the analysis should take (see Figure 30.2). If the data as collected are not suitable for analysis, or if they are only suitable for low-powered analytic techniques, one may wish to use one of the many forms of data transformation to change the data characteristics so that they are more amenable to analysis.

The previous two objectives, data examination and preparation, are the primary focus of this chapter. For data examination, two major techniques are presented—the scattergram and Bartlett's test. Likewise, for data preparation (with the issues of rounding and outliers having been addressed previously) two techniques are presented—randomization (including a test for randomness in a sample of data) and transformation. Exploratory data analysis (EDA) is presented and briefly reviewed later. This is a broad collection of techniques and approaches to "probe" data—that is, to both examine and to perform some initial, flexible analysis of the data.

30.5.14 Scattergram

Two of the major points to be made throughout this volume are (i) the use of the appropriate statistical tests and (ii) the effects of small sample sizes (as is often the case in toxicology) on our selection of statistical techniques. Frequently, simple examination of the nature and distribution of data collected from a study can also suggest patterns and results which were unanticipated and for which the use of

additional or alternative statistical methodology is warranted. It was these three points which caused the author to consider a section on scattergrams and their use essential for toxicologists.

Bartlett's test may be used to determine if the values in groups of data are homogeneous. If they are, this (along with the knowledge that they are from a continuous distribution) demonstrates that parametric methods are applicable.

But, if the values in the (continuous data) groups fail Bartlett's test (i.e., are heterogeneous), we cannot be secure in our belief that parametric methods are appropriate until we gain some confidence that the values are normally distributed. With large groups of data, we can compute parameters of the population (kurtosis and skewness, in particular) and from these parameters determine if the population is normal (with a certain level of confidence). If our concern is especially marked, we can use a chi-square goodness-of-fit test for normality. But when each group of data consists of 25 or fewer values, these measures or tests (kurtosis, skewness, and chi-square goodness-of-fit) are not accurate indicators of normality. Instead, in these cases we should prepare a scattergram of the data and then evaluate the scattergram to estimate if the data are normally distributed. This procedure consists of developing a histogram of the data and then examining the histogram to get a visual appreciation of the location and distribution of the data.

The abscissa (or horizontal scale) should be in the same scale as the values and should be divided so that the entire

range of observed values is covered by the scale of the abscissa. Across such a scale we then simply enter symbols for each of our values. Example 30.2 shows such a plot.

Figure 30.5 is a traditional and rather limited form of scatterplot, but such plots can reveal significant information about the amount and types of association between the two variables, the existence and nature of outliers, the clustering of data, and a number of other two-dimensional factors (Anscombe, 1973; Chambers et al. 1983; Gad, 1985).

Current technology allows us to add significantly more graphic information to scatterplots by means of graphic symbols (letters, faces, different shapes such as squares, colors, etc.) for the plotted data points. One relatively simple example of this approach is shown in Figure 30.5, where the simple case of dose (in a dermal study), dermal irritation, and white blood cell (WBC) count is presented. This graph quite clearly suggests that as dose (variable x) is increased, dermal irritation (variable y) also increases, and as irritation becomes more severe, WBC count (variable z) an indicator of immune system involvement, suggesting infection or persistent inflammation, also increases. There is no direct association of variables x and z , however (Cleveland and McGill, 1984; Cleveland, 1985; Tufte, 1990).

Group 1 can be seen to approximate a normal distribution (bell-shaped curve); we can proceed to perform the appropriate parametric tests with such data. But group 2 clearly does not appear to be normally distributed. In this case, the appropriate nonparametric technique must be used.

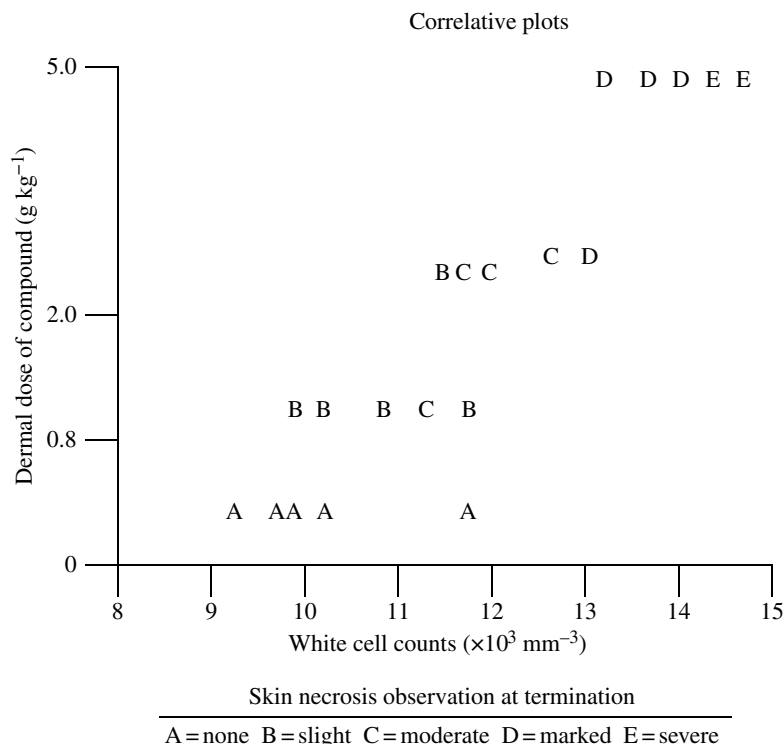


FIGURE 30.5 EDA.

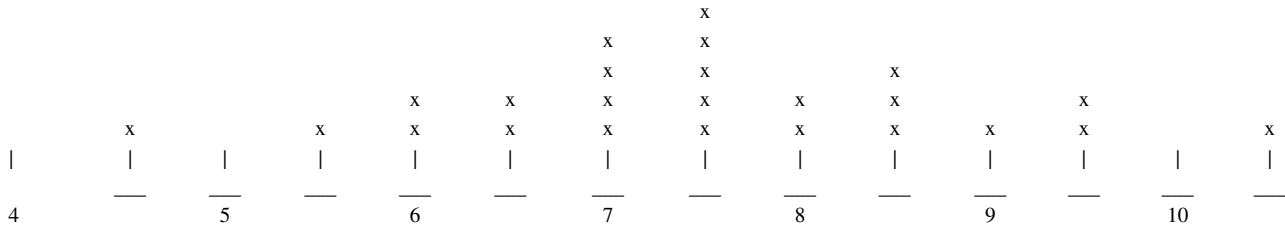
EXAMPLE 30.2

Suppose we have the two data sets in the succeeding text:

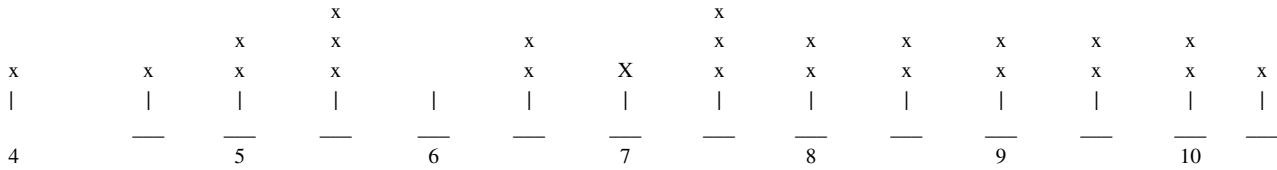
Group 1 4.5, 5.4, 5.9, 6.0, 6.4, 6.5, 6.9, 7.0, 7.1, 7.0, 7.4, 7.5, 7.5, 7.5, 7.6, 8.0, 8.1, 8.4, 8.5, 8.6, 9.0, 9.4, 9.5, and 10.4

Group 2 4.0, 4.5, 5.0, 5.1, 5.4, 5.5, 5.6, 6.5, 6.5, 7.0, 7.4, 7.5, 7.5, 8.0, 8.1, 8.5, 8.5, 9.0, 9.1, 9.5, 9.5, 10.1, 10.0, and 10.4

Both of these groups contain 24 values and cover the same range. From these, we can prepare the following scattergrams: *Group 1*



Group 2



X is individual datum point; χ (chi) is calculated value.

30.5.15 Bartlett's Test for Homogeneity of Variance

Bartlett's test (Sokal and Rohlf, 1994) is used to compare the variances (values reflecting the degree of variability in data sets) among three or more groups of data, where the data in the groups are continuous sets (such as body weights, organ weights, red blood cell (RBC) counts, or diet consumption measurements). It is expected that such data will be suitable for parametric methods (normality of data is assumed) and Bartlett's is frequently used as a test for the assumption of equivalent variances.

Bartlett's is based on the calculation of the corrected χ^2 (chi-square) value by the formula

$$\chi_{\text{cor}}^2 = 2.3026 \frac{\sum \text{df} \left(\log_{10} \left\{ \frac{\sum [\text{df} (S^2)]}{\sum \text{df}} \right\} \right) - \sum [\text{df} (\log_{10} S^2)]}{1 + \left\{ 1 / [3(K-1)] \right\} \left[\sum (1/\text{df}) - (1/\sum \text{df}) \right]}$$

where

$$S^2 = \text{variance} = \left[n \sum X^2 - (\sum X)^2 / n \right] / (n-1)$$

X = individual datum within each group

n = number of data within each group

K = number of groups being compared

df = df for each group = (N - 1).

The corrected χ^2 value yielded by the earlier calculations is compared to the values listed in the chi-square

table according to the number of df (Snedecor and Chocran, 2014).

If the calculated value is smaller than the table value at the selected p level (traditionally 0.05), the groups are accepted to be homogeneous and the use of ANOVA is assumed proper. If the calculated χ^2 is greater than the table value, the groups are heterogeneous and other tests (as indicated in Figure 30.2, the decision tree) are necessary.

ASSUMPTIONS AND LIMITATIONS

1. Bartlett's test does not test for normality but rather homogeneity of variance (also called equality of variances or homoscedasticity).
2. Homoscedasticity is an important assumption for Student's t -test, ANOVA, and analysis of covariance (ANCOVA).
3. The F -test (covered in the next section) is actually a test for the two-sample (i.e., control and one test group) case of homoscedasticity. Bartlett's test is designed for three or more samples.
4. Bartlett's test is very sensitive to departures from normality. As a result, a finding of a significant chi-square value in Bartlett's test may indicate non-normality rather than heteroscedasticity. Such a finding can be brought about by outliers, and the sensitivity to such erroneous findings is extreme with small sample sizes.

30.5.16 Statistical Goodness-of-Fit Tests

A goodness-of-fit test is a statistical procedure for comparing individual measurements to a specified type of statistical distribution. For example, a normal distribution is completely specified by its arithmetic mean and variance (the square of the SD). The null hypothesis, that the data represent a sample from a single normal distribution, can be tested by a statistical goodness-of-fit test. Various goodness-of-fit tests have been devised to determine if the data deviate significantly from a specified distribution. If a significant departure occurs, it indicates only that the specified distribution can be rejected with some assurance. This does not necessarily mean that the true distribution contains two or more subpopulations. The true distribution may be a single distribution, based upon a different mathematical relationship, for example, lognormality. In the latter case, logarithms of the measurement would not be expected to exhibit by a goodness-of-fit test a statistically significant departure from a lognormal distribution.

Everitt and Hand (1981) recommended the use of a sample of 200 or more to conduct a valid analysis of mixtures of populations. Even the maximum likelihood method, the best available method, should be used with extreme caution, or not at all, when separation between the means of the subpopulations is less than 3 SD and sample sizes are less than 300. None of the available methods conclusively establish bimodality, which may, however, occur when separation between the two means (modes) exceeds 2 SD. Conversely, inflections in probits or separations in histograms *less than* 2 SD apart may arise from genetic differences in test subjects.

Mendell et al. (1993) compared eight tests of normality to detect a mixture consisting of two normally distributed components with different means but equal variances. Fisher's skewness statistic was preferable when one component comprised less than 15% of the total distribution. When the two components comprised more nearly equal proportions (35–65%) of the total distribution, the Engelman and Hartigan test (1969) was preferable. For other mixing proportions, the maximum likelihood ratio test was best. Thus, the maximum likelihood ratio test appears to perform very well, with only small loss from optimality, even when it is not the best procedure.

The method of *maximum likelihood* provides estimators which are usually quite satisfactory. They have the desirable properties of being consistent, asymptotically normal, and asymptotically efficient for large samples under quite general conditions. They are often biased, but the bias is frequently removable by a simple adjustment. Other methods of obtaining estimators are also available, but the maximum likelihood method is the most frequently used.

Maximum likelihood estimators also have another desirable property: *invariance*. Let us denote the maximum likelihood estimator of the parameter θ by $\hat{\theta}$. Then, if $f(\theta)$ is

a single-valued function of θ , the maximum likelihood estimator of $f(\theta)$ is $f(\hat{\theta})$. Thus, for example, $\hat{\sigma} = (\hat{\sigma}^2)^{1/2}$.

The principle of maximum likelihood tells us that we should use as our estimate that value which maximizes the likelihood of the observed event.

These maximum likelihood methods can be used to obtain *point estimates* of a parameter, but we must remember that a point estimator is a random variable distributed in some way around the true value of the parameter. The true parameter value may be higher or lower than our estimate. It is often useful therefore to obtain an interval within which we are reasonably confident the true value will lie, and the generally accepted method is to construct what are known as *confidence limits*.

The following procedure will yield upper and lower 95% confidence limits with the property that when we say that these limits include the true value of the parameter, 95% of all such statements will be true and 5% will be incorrect:

1. Choose a (test) statistic involving the unknown parameter and no other unknown parameter.
2. Place the appropriate sample values in the statistic.
3. Obtain an equation for the unknown parameter by equating the test statistic to the upper 2½% point of the relevant distribution.
4. The solution of the equation gives one limit.
5. Repeat the process with the lower 2½% point to obtain the other limit.

One can also construct 95% CI using unequal tails (e.g., using the upper 2% point and the lower 3% point). We usually want our CI to be as short as possible, however, and with a symmetrical distribution such as the normal or t , this is achieved using equal tails. The same procedure very nearly minimizes the CI with other nonsymmetrical distributions (e.g., chi-square) and has the advantage of avoiding rather tedious computations.

When the appropriate statistic involves the square of the unknown parameter, both limits are obtained by equating the statistic to the upper 5% point of the relevant distribution. The use of two tails in this situation would result in a pair of nonintersecting intervals. When two or more parameters are involved, it is possible to construct a region within which we are reasonably confident the true parameter values will lie. Such regions are referred to as confidence regions. The implied interval for p_1 does not form a 95% CI, however. Nor is it true that an 85.7375% confidence region for p_1 , p_2 , and p_3 can be obtained by considering the intersection of the three separate 95% CI, because the statistics used to obtain the individual CI are not independent. This problem is obvious with a multiparameter distribution such as the multinomial, but it even occurs with the normal distribution because the statistic which we use to obtain a CI for the mean and the

statistic which we use to obtain a CI for the variance are not independent. The problem is not likely to be of great concern unless a large number of parameters are involved.

30.5.17 Randomization

Randomization is the act of assigning a number of items (e.g., plates of bacteria or test animals) to groups in such a manner that there is an equal chance for any one item to end up in any one group. This is a control against any possible bias in assignment of subjects to test groups. A variation on this is censored randomization, which ensures that the groups are equivalent in some aspect after the assignment process is complete. The most common example of a censored randomization is one in which it is ensured that the body weights of test animals in each group are not significantly different from those in the other groups. This is done by analyzing group weights both for homogeneity of variance and by ANOVA after animal assignment and then rerandomizing if there is a significant difference at some nominal level, such as $p \leq 0.10$. The process is repeated until there is no significant difference.

There are several methods for actually performing the randomization process. The three most commonly used are card assignment, use of a random number table, and use of a computerized algorithm.

For the card-based method, individual identification numbers for items (e.g., plates or animals) are placed on separate index cards. These cards are then shuffled, placed one at a time in succession into piles corresponding to the required test groups. The results are a random group assignment.

The random number table method requires only that one have unique numbers assigned to test subjects and access to a random number table. One simply sets up a table with a column for each group to which subjects are to be assigned. We start from the head of any one column of numbers in the random table (each time the table is used, a new starting point should be utilized). If our test subjects number less than 100, we utilize only the last two digits in each random number in the table. If they number more than 99 but less than 1000, we use only the last three digits. To generate group assignments, we read down a column, one number at a time. As we come across digits which correspond to a subject number, we assign that subject to a group (enter its identifying number in a column) proceeding to assign subjects to groups from left to right filling one row at a time. After a number is assigned to an animal, any duplication of its unique number is ignored. We use as many successive columns of random numbers as we may need to complete the process.

The third (and now most common) method is to use a random number generator that is built into a calculator or computer program. Procedures for generating these are generally documented in user manuals.

30.5.18 Transformations

If our initial inspection of a data set reveals it to have an unusual or undesired set of characteristics (or to lack a desired set of characteristics), we have a choice of three courses of action. We may proceed to select a method or test appropriate to this new set of conditions, or abandon the entire exercise, or transform the variable(s) under consideration in such a manner that the resulting transformed variates (e.g., X' and Y' , as opposed to the original variates X and Y) meet the assumptions or have the characteristics that are desired.

The key to all this is that the scale of measurement of most (if not all) variables is arbitrary. Although we are most familiar with a linear scale of measurement, there is nothing which makes this the "correct" scale on its own, as opposed to a logarithmic scale (familiar logarithmic measurements are that of pH values or earthquake intensity (Richter scale)). Transforming a set of data (converting X to X') is really as simple as changing a scale of measurement.

There are at least four good reasons to transform data:

1. To normalize the data, making them suitable for analysis by our most common parametric techniques such as ANOVA. A simple test of whether a selected transformation will yield a distribution of data which satisfies the underlying assumptions for ANOVA is to plot the cumulative distribution of samples on probability paper (i.e., a commercially available paper which has the probability function scale as one axis). One can then alter the scale of the second axis (i.e., the axis other than the one which is on a probability scale) from linear to any other (logarithmic, reciprocal, square root, etc.) and see if a previously curved line indicating a skewed distribution becomes linear to indicate normality. The slope of the transformed line gives us an estimate of the SD. And if the slopes of the lines of several samples or groups of data are similar, we accordingly know that the variance of the different groups is homogenous.
2. To linearize the relationship between a paired set of data, such as dose and response. This is the most common use in toxicology for transformations and is demonstrated in the section under probit/logit plots.
3. To adjust data for the influence of another variable. This is an alternative in some situations to the more complicated process of ANCOVA. A ready example of this usage is the calculation of organ weight to body weight ratios in *in vivo* toxicity studies, with the resulting ratios serving as the raw data for an ANOVA performed to identify possible target organs. This use is discussed in detail later in this chapter.

4. Finally, to make the relationships between variables clearer by removing or adjusting for interactions with third, fourth, and so on. Uncontrolled variables influence the pair of variables of interest. This case is discussed in detail under time series analysis.

Common transformations are presented in Table 30.6.

30.5.19 Exploratory Data Analysis

Over the past 20 years, an entirely new approach has been developed to get the most information out of the increasingly larger and more complex data sets that scientists are faced with. This approach involves the use of a very diverse set of fairly simple techniques which comprise EDA. As expounded by Tukey (1977), there are four major ingredients to EDA:

Displays These visually reveal the behavior of the data and suggest a framework for analysis. The scatterplot (presented earlier) is an example of this approach.

Residuals These are what remain of a set of data after a fitted model (such as a linear regression) or some similar level of analysis has been removed.

Reexpressions These involved questions of what scale would serve to best simplify and improve the analysis of the data. Simple transformations, such as those presented earlier in this chapter, are used to simplify data behavior (e.g., linearizing or normalizing) and clarify analysis.

Resistance This is a matter of decreasing the sensitivity of analysis and summary of data to misbehavior, so that the occurrence of a few outliers, for example, will not complicate or invalidate the methods used to analyze the data. For example, in summarizing the location of a set of data, the median (but not the arithmetic mean) is highly resistant.

These four ingredients are utilized in a process falling into two broad phases: an exploratory phase and a confirmatory phase. The exploratory phase isolates patterns in and features of the data and reveals them, allowing an inspection of the data before there is any firm choice of actual hypothesis testing or modeling methods have been made.

Confirmatory analysis allows evaluation of the reproducibility of the patterns or effects. Its role not only is close to that of classical hypothesis testing but also often includes steps such as (i) incorporating information from an analysis of another, closely related set of data and (ii) validating a result by assembling and analyzing additional data. These techniques are in general beyond the scope of this text. However, Velleman and Hoaglin (1981) and Hoaglin and Mosteller (2000) present a clear overview of the more important methods, along with codes for their execution on a microcomputer (they have also now been incorporated into Minitab). A short examination of a single case of the use of these methods, however, is in order.

Toxicology has long recognized that no population—animal or human—is completely uniform in its response to any particular toxicant. Rather, a population is composed of a (presumably normal) distribution of individuals—some resistant to intoxication (hyporesponders), the bulk that responds close to a central value (such as an LD_{50}), and some that are very sensitive to intoxication (hyperresponders). This population distribution can, in fact, result in additional statistical techniques. The sensitivity of techniques such as ANOVA is reduced markedly by the occurrence of outliers (extreme high or low values—including hyper- and hyporesponders) which, in fact, serve to markedly inflate the variance (SD) associated with a sample. Such variance inflation is particularly common in small groups that are exposed or dosed at just over or under a threshold level, causing a small number of individuals in the sample (who are more sensitive than the other members) to respond markedly. Such a situation is displayed in Figure 30.6 which plots the mean and SD of

TABLE 30.6 Common Data Transformations^a

| Transformation | Means of Calculation ^b | Example of Use |
|-------------------------------|---|--|
| Arithmetic | $x' = x/y$ or $x' = x + c$ | Organ weight/body weight |
| Reciprocals | $x' = 1/x$ | Linearizing data, particularly rate phenomena |
| Arcsine (also called angular) | $x' = \arcsin \sqrt{x}$ | Normalizing dominant lethal and mutation rate data |
| Logarithmic | $x' = \log x$ | pH values |
| Probability (probit) | $x' = p(x)$ | Percentage responding |
| Square roots | $x' = \sqrt{x}$ | Surface area of animal from body weights |
| Box-Cox | $x' = (x^v - 1)/v$: for $v \neq 0$ $x' = \ln x$: for $v = 0$ | A family of transforms For use when one has no prior knowledge of the appropriate transformation to use |
| Rank transformations | Depends on nature of samples | As a bridge between parametric and nonparametric statistics (Conover and Inman, 1981) |

^a Plotting a double reciprocal (i.e., $1/x$ vs. $1/y$) will linearize almost any data set. So will plotting the log transforms of a set of variables.

^b x and y are original variables and x' and y' transformed values. “ c ” stands for a constant.

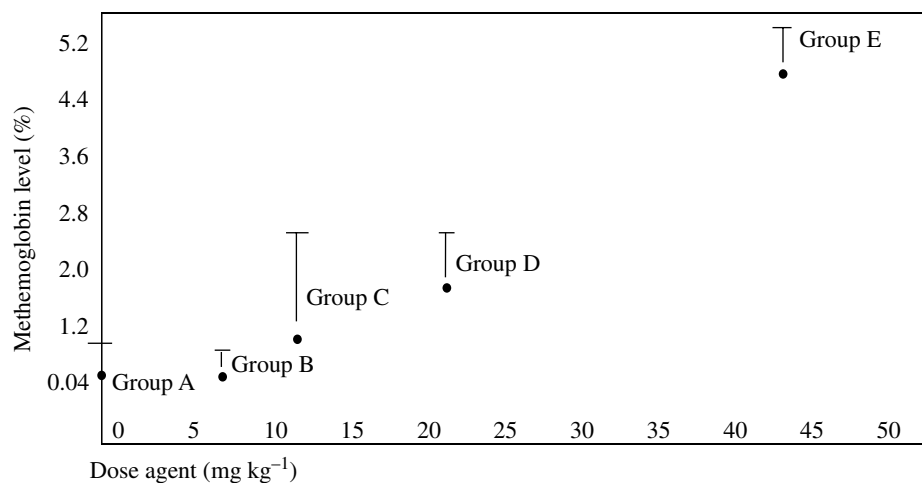


FIGURE 30.6 Variance inflation: points are means minus error bars plus one SD.

methemoglobin levels in a series of groups of animals exposed to successively higher levels of a hemolytic agent.

Though the mean level of methemoglobin in group C is more than double that of the control group (A), no hypothesis test will show this difference to be significant because it has such a large SD associated with it. Yet this “inflated” variance exists because a single individual has such a marked response. The occurrence of the inflation is certainly an indicator that the data need to be examined closely. Indeed, all tabular data in toxicology should be visually inspected for both trend and variance inflation.

A concept related (but not identical) to resistance and EDA is that of robustness. Robustness generally implies insensitivity to departures from assumptions surrounding an underlying model, such as normality.

In summarizing the location of data, the median, though highly resistant, is not extremely robust. But the mean is both nonresistant and nonrobust.

30.6 HYPOTHESIS TESTING OF CATEGORICAL AND RANKED DATA

Categorical (or contingency table) presentations of data can contain any single type of data, but generally the contents are collected and arranged so that they can be classified as belonging to treatment and control groups, with the members of each of these groups then classified as belonging to one of two or more response categories (such as tumor/no tumor or normal/hyperplastic/neoplastic). For these cases, two forms of analysis are presented—Fisher’s exact test (for the 2×2 contingency table) and the $R \times C$ chi-square test (for large tables). It should be noted, however, that there are versions of both of these tests which permit the analysis of any size of contingency table.

The analysis of rank data—what is generally called nonparametric statistical analysis—is an exact parallel of the more traditional (and familiar) parametric methods. There are methods for the single comparison case (just as Student’s *t*-test is used) and for the multiple comparison case (just as ANOVA is used) with appropriate *post hoc* tests for exact identification of the significance with a set of groups. Four tests are presented for evaluating statistical significance in rank data—the Wilcoxon rank-sum test, distribution-free multiple comparisons, Mann–Whitney *U* test, and the Kruskal–Wallis nonparametric ANOVA. For each of these tests, tables of distribution values for the evaluations of results can be found in any of a number of reference volumes (Gad, 2005).

It should be clearly understood that for data which do not fulfill the necessary assumptions for parametric analysis, these nonparametric methods are either as powerful or in fact more powerful than the equivalent parametric test.

30.6.1 Fisher’s Exact Test

Fisher’s exact test should be used to compare two sets of discontinuous, quantal (all or none) data. Small sets of such data can be checked by contingency data tables, such as those of Finney et al. (1963). Larger sets, however, require computation. These include frequency data such as incidences of mortality or certain histopathological findings, and so on. Thus, the data can be expressed as ratios. These data do not fit on a continuous scale of measurement but usually involve a number of responses classified as either negative or positive—that is, contingency table situation (Sokal and Rohlf, 2011).

The analysis is started by setting up a 2×2 contingency table to summarize the number of “positive” and “negative” responses as well as the totals of these, as follows:

| | “Positive” | “Negative” | Total |
|----------|---------------------|---------------------|---|
| Group I | <i>A</i> | <i>B</i> | <i>A</i> + <i>B</i> |
| Group II | <i>C</i> | <i>D</i> | <i>C</i> + <i>D</i> |
| Totals | <i>A</i> + <i>C</i> | <i>B</i> + <i>D</i> | <i>A</i> + <i>B</i> + <i>C</i> + <i>D</i> = <i>N</i> _{total} |

Using the earlier set of symbols, the formula for *P* appears as follows:

$$P = \frac{(A+B)!(C+D)!(A+C)!(B+D)!}{N!A!B!C!D!}_1$$

The exact test produces a probability (*P*) which is the sum of the previous calculation repeated for each possible arrangement of the numbers in the earlier cells (i.e., *A*, *B*, *C*, and *D*) showing an association equal to or stronger than that between the two variables.

The *P* resulting from these computations will be the exact one- or two-tailed probability depending on which of these two approaches is being employed. This value tells us if the groups differ significantly (with a probability <0.05) and the degree of significance.

ASSUMPTIONS AND LIMITATIONS

1. Tables are available which provide individual exact probabilities for small sample size contingency tables (see Zar 1974, pp. 518–542).
2. Fisher’s exact test must be used in preference to the chi-square test when there are small cell numbers.
3. The probability resulting from a two-tailed test is exactly double that of a one-tailed test from the same data.
4. Ghent has developed and proposed a good (though, if performed by hand, laborious) method extending the calculation of exact probabilities to 2×3, 3×3, and *R*×*C* contingency tables (Ghent, 1972).
5. Fisher’s probabilities are not necessarily symmetrical. Although some analysts will double the one-tailed *p* value to obtain the two-tailed result, this method is usually overly conservative.

30.6.2 2×2 Chi-Square

Though Fisher’s exact test is preferable for analysis of most 2×2 contingency tables in toxicology, the chi-square test is still widely used and is preferable in a few unusual situations (particularly if cell sizes are large yet only limited computational support is available).

The formula is simply

$$\chi^2 = \frac{(O_1 - E_1)^2}{E_1} + \frac{(O_2 - E_2)^2}{E_2}$$

$$= \sum \frac{(O_i - E_i)^2}{E_i}$$

where *O* are observed numbers (or counts) and *E* are expected numbers. The common practice in toxicology is for the observed figures to be test or treatment group counts. The expected figure is calculated as

$$E = \frac{(\text{column total})(\text{row total})}{\text{grand total}}$$

for each box or cell in a contingency table.

Our df are (*R*−1)(*C*−1)=(2−1)(2−1)=1. Looking at a chi-square table (provided in Gad, 2005)) for one df, we see that this is greater than the test statistic at 0.05 (3.84) but less than that at 0.01 (6.64) so that 0.05>*p*>0.01.

ASSUMPTIONS AND LIMITATIONS

Assumptions:

1. Data are univariate and categorical.
2. Data are from a multinomial population.
3. Data are collected by random, independent sampling.
4. Groups being compared are of approximately same size, particularly for small group sizes.

When to use:

1. The data are of a categorical (or frequency) nature.
2. The data fit the earlier assumptions.
3. Goodness-to-fit to a known form of distribution is being tested.
4. Cell sizes are large.

When not to use:

1. The data are continuous rather than categorical.
2. Sample sizes are small and very unequal.
3. Sample sizes are too small (e.g., when total *N* is less than 50 or if any expected value is less than 5).
4. Any 2×2 comparison is being performed (use Fisher’s exact test instead).

30.6.3 *R*×*C* Chi-Square

The *R*×*C* chi-square test can be used to analyze discontinuous (frequency) data as in Fisher’s exact or 2×2 chi-square tests. However, in the *R*×*C* test (*R*=row; *C*=column) we

¹A! is A factorial. For 4! as an example this would be (4) (3) (2) (1)=24.

wish to compare three or more sets of data. An example would be comparison of the incidence of tumors among mice on three or more oral dosage levels. We can consider the data as “positive” (tumors) or “negative” (no tumors). The expected frequency for any box = (row total)(column total)/(N_{total}).

As in Fisher’s exact test, the initial step is setting up a table (this time an $R \times C$ contingency table). This table would appear as follows:

| | “Positive” | “Negative” | Total |
|----------|--------------|--------------|--------------------|
| Group I | A_1 | B_1 | $A_1 + B_1 = N_1$ |
| Group II | A_2 | B_2 | $A_2 + B_2 = N_2$ |
| | \downarrow | \downarrow | |
| Group R | A_R | B_R | $A_R + B_R = N_R$ |
| Totals | N_A | N_B | N_{total} |

Using these symbols, the formula for chi-square (χ^2) is

$$\chi^2 = \frac{N_{\text{tot}}^2}{N_A N_B N_K} \left(\frac{A_1^2}{N_1} + \frac{A_2^2}{N_2} + \dots + \frac{A_K^2}{N_K} - \frac{N_A^2}{N_{\text{tot}}} \right)$$

The resulting χ^2 value is compared to table values (as in Snedecor and Cochran (2014) pp. 470–471) according to the number of df, which is equal to $(R-1)(C-1)$. If χ^2 is smaller than the table value at the 0.05 probability level, the groups are not significantly different. If the calculated χ^2 is larger, there is some difference among the groups, and $2 \times R$ chi-square or Fisher’s exact tests will have to be compared to determine which group(s) differs from which other group(s).

ASSUMPTIONS AND LIMITATIONS

1. The test is based on data being organized in a table (such as below) with *cells* (in the following table, A , B , C , and D are “cells”).

| | Columns (C) | | |
|-----------|-------------|---------|-----------------|
| | Control | Treated | Total |
| No Effect | A | B | $A + B$ |
| Effect | C | D | $C + D$ |
| Total | $A + C$ | $B + D$ | $A + B + C + D$ |

2. None of the expected frequency values should be less than 5.0.
3. The chi-square test is always one tailed.
4. Without the use of some form of correction, the test

becomes less accurate as the differences between group sizes increase.

5. The results from each additional column (group) are approximately additive. Due to this characteristic, the chi-square test can be readily used for evaluating any $R \times C$ combination.
6. The results of the chi-square calculation must be a positive number.
7. The test is weak with either small sample sizes or when the expected frequency in any cell is less than 5 (this latter limitation can be overcome by pooling, or combining, cells).
8. Test results are independent of order of cells, unlike Kolmogorov–Smirnov.
9. The test can be used to test the probability of validity of any distribution.

30.6.4 Wilcoxon Rank-Sum Test

The Wilcoxon rank-sum test is commonly used for the comparison of two groups of nonparametric (interval or not normally distributed) data, such as those which are not measured exactly but rather as falling within certain limits (e.g., how many animals died during each hour of an acute study). The test is also used when there is no variability (variance=0) within one or more of the groups we wish to compare (Sokal and Rohlf, 2011).

The data in both groups being compared are initially arranged and listed in order of increasing value. Then each number in the two groups must receive a rank value. Beginning with the smallest number in either group (which is given a rank of 1.0), each number is assigned a rank. If there are duplicate numbers (called “ties”), then each value of equal size will receive the median rank for the entire identically sized group. Thus if the lowest number appears twice, both figures receive a rank of 1.5. This, in turn, means that the ranks of 1.0 and 2.0 have been used and that the next highest number has a rank of 3.0. If the lowest number appears three times, then each is ranked as 2.0 and the next number has a rank of 4.0. Thus, each tied number gets a “median” rank. This process continues until all of the numbers are ranked. Each of the two columns of ranks (one for each group) is totaled giving the “sum of ranks” for each group being compared. As a check, we can calculate the value

$$\frac{(N)(N+1)}{2}$$

where N is the total number of data in both groups. The result should be equal to the sum of the sum of ranks for both groups.

The sum of rank values are compared to table values (Beyer, 1976, pp. 409–413) to determine the degree of significant differences, if any. These tables include two limits (an upper and a lower) that are dependent upon the probability level. If the number of data is the same in both groups ($N_1 \neq N_2$), then the lesser sum of ranks (smaller N) is compared to the table limits to find the degree of significance. Normally the comparison of the two groups ends here and the degree of significant difference can be reported.

30.6.5 Distribution-Free Multiple Comparison

The distribution-free multiple comparison test should be used to compare three or more groups of nonparametric data. These groups are then analyzed two at a time for any significant differences (Hollander and Wolfe, 2013, pp. 124–129). The test can be used for data similar to those compared by the rank-sum test. We often employ this test for reproduction and mutagenicity studies (such as comparing survival rates of offspring of rats fed with various amounts of test materials in the diet).

Two values must be calculated for each pair of groups: the difference in mean ranks and the probability level value against which the difference will be compared. To determine the difference in mean ranks, we must first arrange the data within each of the groups in order of increasing values. Then we must assign rank values, beginning with the smallest overall figure. Note that this ranking is similar to that in the Wilcoxon test except that it applies to more than two groups.

The ranks are then added for each of the groups. As a check, the sum of these should equal

$$\frac{N_{\text{tot}}(N_{\text{tot}} + 1)}{2}$$

where N_{tot} is the total number of figures from all groups. Next we can find the mean rank (R) for each group by dividing the sum of ranks by the numbers in the data (N) in the group. These mean ranks are then taken in those pairs which we want to compare (usually each test group vs. the control) and the differences are found ($|R_1 - R_2|$). This value is expressed as an absolute figure, that is, it is always a positive number.

The second value for each pair of groups (the p value) is calculated from the expression

$$z \left[\frac{a}{K} (K-1) \right] \sqrt{\frac{N_{\text{tot}}(N_{\text{tot}} + 1)}{12}} \sqrt{\frac{1}{N_1} \frac{1}{N_2}}$$

where a is the level of significance for the comparison (usually 0.05, 0.01, 0.001, etc.), K is the total number of groups, and Z is a figure obtained from a normal probability table and determining the corresponding “Z-score.”

The result of the p value calculation for each pair of groups is compared to the corresponding mean difference $|R_1 - R_2|$. If $|R_1 - R_2|$ is smaller, there is no significant difference between the groups. If it is larger, the groups are different and $|R_1 - R_2|$ must be compared to the calculated p values for $a=0.01$ and 0.001 to find the degree of significance.

ASSUMPTIONS AND LIMITATIONS

1. As with the Wilcoxon rank-sum test, too many tied ranks inflate the false positive.
2. Generally, this test should be used as a *post hoc* comparison after Kruskal–Wallis.

30.6.6 Mann–Whitney U Test

This is a nonparametric test in which the data in each group are first ordered from lowest to highest values, and then the entire set (both control and treated values) is ranked, with the average rank being assigned to tied values. The ranks are then summed for each group and U is determined according to

$$U_t = n_c n_t + \frac{n_t(n_t + 1)}{2} - R_t$$

$$U_c = n_c n_t + \frac{n_c(n_c + 1)}{2} - R_c$$

where n_c , n_t = sample size for control and treated groups and R_c , R_t = sum of ranks for the control and treated groups.

For the level of significance for a comparison of the two groups, the larger value of U_c or U_t is used. This is compared to critical values as found in tables (Siegel, 1956).

The Mann–Whitney U test is employed for the count data, but which test should be employed for the percentage variables should be decided on the same grounds as described later under reproduction studies.

ASSUMPTIONS AND LIMITATIONS

1. It does not matter whether the observations are ranked from smallest to largest or vice versa.
2. This test should not be used for paired observations.
3. The test statistics from a Mann–Whitney test are linearly related to those of Wilcoxon. The two tests will always yield the same result. The Mann–Whitney is presented here for historical completeness, as it has been much favored in reproductive and developmental toxicology studies; however, it should be noted that the authors do not include it in the decision tree for method selection (Figure 30.2).

30.6.7 Kruskal–Wallis Nonparametric ANOVA

The Kruskal–Wallis nonparametric one-way ANOVA should be the initial analysis performed when we have three or more groups of data which are by nature nonparametric (either not a normally distributed population, or of a discontinuous nature, or all the groups being analyzed are not from the same population) but not a categorical (or quantal) nature. Commonly these will be either rank-type evaluation data (such as behavioral toxicity observation scores) or reproduction study data. The analysis is initiated (Pollard, 1977, pp. 170–173) by ranking all the observations from the combined groups to be analyzed. Ties are given the average rank of the tied values (i.e., if two values would tie for 12th rank—and therefore would be ranked 12th and 13th—both would be assigned the average rank of 12.5).

The sum of ranks of each group (r_1, r_2, \dots, r_k) is computed by adding all the rank values for each group. The test value H is then computed as

$$H = \frac{12}{n(n+1)} \sum \left(\frac{r_1^2}{n_1} + \frac{r_2^2}{n_2} + \dots + \frac{r_k^2}{n_k} \right) - 3(n+1)$$

where n_1, n_2, \dots, n_k are the number of observations in each group. The test statistic is then compared with a table of H values. If the calculated value of H is greater than the table value for the appropriate number of observations in each group, there is significant difference between the groups, but further testing (using the distribution-free multiple comparison method) is necessary to determine where the difference lies.

ASSUMPTIONS AND LIMITATIONS

1. The test statistic H is used for both small and large samples.
2. When we find a significant difference, we do not know which groups are different. It is not correct to then perform a Mann–Whitney U test on all possible combinations; rather, a multiple comparison method must be used, such as the distribution-free multiple comparisons.
3. Data must be independent for the test to be valid.
4. Too many tied ranks will decrease the power of this test and also lead to increased false-positive levels.
5. When $k=2$, the Kruskal–Wallis chi-square value has 1 degree of freedom. This test is identical to the normal approximation used for the Wilcoxon rank-sum test. As noted in previous sections, a chi-square with 1 degree of freedom can be represented by the square of a standardized normal random variable. In the case of $k=2$, the h -statistic is the square of the Wilcoxon rank-sum Z -test (without the continuity correction).
6. The effect of adjusting for tied ranks is to slightly increase the value of the test statistic, h ; therefore, omission of this adjustment results in a more conservative test.

30.6.8 Log-Rank Test

The log-rank test is a statistical methodology for comparing the distribution of time until the occurrence of the event in independent groups. In toxicology, the most common event of interest is death or occurrence of a tumor, but it could just as well be liver failure, neurotoxicity, or any other event which occurs only once in an individual. The elapsed time from initial treatment or observation until the *event* is the *event time*, often referred to as “survival time,” even when the *event* is not “death.”

The log-rank test provides a method for comparing “risk-adjusted” event rates, useful when test subjects in a study are subject to varying degrees of opportunity to experience the event. Such situations arise frequently in toxicology studies due to the finite duration of the study, early termination of the animal, or interruption of treatment before the event occurs.

Examples where use of the log-rank test might be appropriate include comparing survival times in carcinogenicity bioassay animals which are given a new treatment with those in the control group or comparing times to liver failure for several dose levels of a new NSAID where the animals are treated for 10 weeks or until cured, whichever comes first.

If every animal were followed until the event occurrence, the event times could be compared between two groups using the Wilcoxon rank-sum test. However, some animals may die or complete the study before the event occurs. In such cases, the actual time to the event is unknown since the event does not occur while under study observation. The event times for these animals are based on the last known time of study observation and are called “censored” observations since they represent the lower bound of the true, unknown event times. The Wilcoxon rank-sum test can be highly biased in the presence of the censored data.

The null hypothesis tested by the log-rank test is that of equal event time distributions among groups. Equality of the distributions of event times implies similar event rates among groups not only for the clinical trial as a whole but also for any arbitrary time point during the trial. Rejection of the null hypothesis indicates that the event rates differ among groups at one or more time points during the study.

The idea behind the log-rank test for comparison of two life tables is simple; if there were no difference between the groups, the total deaths occurring at any time should split between the two groups at that time. So if the numbers at risk in the first and second groups in (say) the sixth month were 70 and 30, respectively, and 10 deaths occurred in that month, we would expect

$$10 \times \frac{70}{70+30} = 7$$

of these deaths to have occurred in the first group and

$$10 \times \frac{30}{70 + 30} = 3$$

of these deaths to have occurred in the second group.

A similar calculation can be made at each time of death (in either group). By adding together for the first group the results of all such calculations, we obtain a single number, called the extent of exposure (E_1), which represents the “expected” number of deaths in that group if the two groups had the distribution of survival time. An extent of exposure (E_2) can be obtained for the second group in the same way. Let O_1 and O_2 denote the actual total number of deaths in the two groups. A useful arithmetic check is that the total number of deaths $O_1 + O_2$ must equal the sum $E_1 + E_2$ of the extents of exposure.

The discrepancy between the O s and E s can be measured by the quantity

$$x^2 = \frac{(|O_1 - E_1| - (1/2))^2}{E_1} + \frac{(|O_2 - E_2| - (1/2))^2}{E_2}$$

For rather obscure reasons, x^2 is known as the log-rank statistic. An approximate significance test of the null hypothesis of identical distributions of survival time in the two groups is obtained by referring x^2 to a chi-square distribution on 1 df.

The log-rank test as presented by Peto et al. (1977) uses the product-limit life-table calculations rather than the actuarial estimators shown in the preceding text. The distinction is unlikely to be of practical importance unless the grouping intervals are very coarse.

Peto and Pike (1973) suggest that the approximation in treating the null distribution of χ^2 as a chi-square is conservative, so that it will tend to understate the degree of statistical significance. In the formula for χ^2 , we have used the continuity correction of subtracting 1/2 from $|O_1 - E_1|$ and $|O_2 - E_2|$ before squaring. This is recommended by Peto et al. (1977) when, as in nonrandomized studies, the permutational argument does not apply. Peto et al. (1977) give further details of the log-rank test and its extension to comparisons of more than two treatment groups and to tests that control for categorical confounding factors.

ASSUMPTIONS AND LIMITATIONS

1. The end point of concern is or is defined so that it is “right censored”—once it happens, it does not reoccur. Examples are death or a minimum or maximum value of an enzyme or physiologic function (such as respiration rate).
2. The method makes no assumptions on distribution.

3. Many variations of the log-rank test for comparing survival distributions exist. The most common variant has the form:

$$X^2 = \frac{(O_1 - E_1)^2}{E_1} + \frac{(O_2 - E_2)^2}{E_2}$$

where O_i and E_i are computed for each group, as in the formulas given previously. This statistic also has an approximate chi-square distribution with 1 degree of freedom under H_0 .

A continuity correction can also be used to reducing the numerators by 1/2 before squaring. Use of such a correction leads to even further conservatism and may be omitted when sample sizes are moderate or large.

4. The Wilcoxon rank-sum test could be used to analyze the event times in the absence of censoring. A “generalized Wilcoxon” test, sometimes called the Gehan test, based on an approximate chi-square distribution has been developed for use in the presence of censored observations.

Both the log-rank and the generalized Wilcoxon tests are nonparametric tests and require no assumptions regarding the distribution of event times. When the event rate is greater early in the trial than toward the end, the generalized Wilcoxon test is the more appropriate test since it gives greater weight to the earlier differences.

5. Survival and failure times often follow the exponential distribution. If such a model can be assumed, a more powerful alternative to the log-rank test is the likelihood ratio test.

This parametric test assumes that event probabilities are constant over time. That is, the chance that a patient becomes event-positive at time t given that he or she is event-negative up to time t does not depend on t . A plot of the negative log of the event time distribution showing a linear trend through the origin is consistent with exponential event times.

6. Life tables can be constructed to provide estimates of the event time distributions. Estimates commonly used are known as the Kaplan–Meier estimates.

30.7 HYPOTHESIS TESTING: UNIVARIATE PARAMETRIC TESTS

Univariate case² data from normally distributed populations generally have a higher information value associated with them, but the traditional hypothesis testing techniques (which include all the methods described in this chapter) are

²That is, where each datum is defined by one treatment and one effect variable.

generally neither resistant nor robust. All the data analyzed by these methods are also, effectively, continuous—that is, at least for practical purposes, the data may be represented by any number and each such data number has a measurable relationship to other data numbers. Note that Diem and Lentner (1975) and Pollard (1977) are excellent sources of statistical value tables.

30.7.1 Student's *t*-Test (Unpaired *t*-Test)

Pairs of groups of continuous, randomly distributed data are compared via this test. We can use this test to compare three or more groups of data, but they must be intercompared by examination of two groups taken at a time and are preferentially compared by ANOVA. Usually this means comparison of a test group versus a control group, although two test groups may be compared as well. To determine which of the three types of *t*-tests described in this chapter should be employed, the *F*-test is usually performed first. This will tell us if the variances of the data are approximately equal, which is a requirement for the use of the parametric methods. If the *F*-test indicates homogeneous variances and the numbers of data within the groups (*N*) are equal, then Student's *t*-test is the appropriate procedure (Sokal and Rohlf, 1994). If the *F* is significant (the data are heterogeneous) and the two groups have equal numbers of data, the modified Student's *t*-test is applicable (Cochran and Cox, 1975).

The value of *t* for Student *t*-test is calculated using the formula

$$t = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{\frac{\sum D_1^2 + \sum D_2^2}{N_1 + N_2} (N_1 + N_2 - 2)}}$$

where the value of $\sum D^2 = [N \sum X^2 - (\sum X)^2] / N$

The value of *t* obtained from the earlier calculations is compared to the values in a *t*-distribution table according to the appropriate number of df. If the *F* value is not significant (i.e., variances are homogeneous), the $df = N_1 + N_2 - 2$. If the *F* was significant and $N_1 = N_2$, then the $df = N - 1$. Although this case indicates a nonrandom distribution, the modified *t*-test is still valid. If the calculated value is larger than the table value at $p = 0.05$, it may then be compared to the appropriate other table values in order of decreasing probability to determine the degree of significance between the two groups.

ASSUMPTIONS AND LIMITATIONS

1. The test assumes that the data are univariate, continuous, and normally distributed.
2. Data are collected by random sampling.
3. The test should be used when the assumptions in 1 and 2 are met and there are only two groups to be compared.

4. Do not use when the data are ranked, when the data are not approximately normally distributed, or when there are more than two groups to be compared. Do not use for paired observations.
5. This is the most commonly misused test method, except in those few cases where one is truly only comparing two groups of data and the group sizes are roughly equivalent. It is not valid for multiple comparisons (because of resulting additive errors) or where group sizes are very unequal.
6. Test is robust for moderate departures from normality and, when N_1 and N_2 are approximately equal, robust for moderate departures from homogeneity of variances.
7. The main difference between the *Z*-test and the *t*-test is that the *Z*-statistic is based on a known standard deviation, σ , while the *t*-statistic uses the sample standard deviation, s , as an estimate of σ . With the assumption of normally distributed data, the variance σ^2 is more closely estimated by the sample variance s^2 as n gets large. It can be shown that the *t*-test is equivalent to the *Z*-test for infinite degrees of freedom. In practice, a "large" sample is usually considered $n \geq 30$.

30.7.2 Cochran *t*-Test

The Cochran test should be used to compare two groups of continuous data when the variances (as indicated by the *F*-test) are heterogeneous and the numbers of data within the groups are not equal ($N_1 \neq N_2$). This is the situation, for example, when the data, though expected to be randomly distributed, were found not to be randomly distributed (Cochran and Cox, 1975, pp. 100–102).

Two *t* values are calculated for this test, the "observed" t (t_{obs}) and the "expected" t (t'). The observed *t* is obtained by

$$t_{\text{obs}} = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{W_1 + W_2}}$$

where

$$\begin{aligned} W &= \text{SEM}^2 \text{ (SEM squared)} \\ &= S^2/N \end{aligned}$$

where *S* (variance) can be calculated from

$$S = \frac{[N \sum X^2 - (\sum X)^2] / N}{N - 1}$$

The value for t' is obtained from

$$t' = \frac{t'_1 W_1 + t'_2 W_2}{W_1 + W_2}$$

where t'_1 and t'_2 are values for the two groups taken from the t -distribution table corresponding to $N-1$ df (for each group) at the 0.05 probability level (or such level as one may select).

The calculated t_{obs} is compared to the calculated t' value (or values, if t' values were prepared for more than one probability level). If t_{obs} is smaller than a t' , the groups are not considered to be significantly different at that probability level.

ASSUMPTIONS AND LIMITATIONS

1. The test assumes that the data are univariate, continuous, and normally distributed and that group sizes are unequal.
2. The test is robust for moderate departures from normality and very robust for departures from equality of variances.

30.7.3 F -Test

This is a test of the homogeneity of variances between two groups of data (Sokal and Rohlf, 1994). It is used in two separate cases. The first is when Bartlett's indicates heterogeneity of variances among three or more groups (i.e., it is used to determine which pairs of groups are heterogeneous). Second, the F -test is the initial step in comparing two groups of continuous data which we would expect to be parametric (two groups not usually being compared using ANOVA), the results indicating whether the data are from the same population and whether subsequent parametric comparisons would be valid.

The F is calculated by dividing the larger variance (S_1^2) by the smaller one (S_2^2). S^2 is calculated as

$$S^2 = \frac{[N \sum X^2 - (\sum X)^2] / N}{N - 1}$$

where N is the number of data in the group and X represents the individual values within the group. Frequently, S^2 values may be obtained from ANOVA calculations.

The calculated F value is compared to the appropriate number in an F -value table for the appropriate df ($N-1$) in the numerator (along the top of the table) and in the denominator (along the side of the table). If the calculated value is smaller, it is not significant and the variances are considered homogeneous (and Student's t -test would be appropriate for further comparison). If the calculated F value is greater, F is significant and the variances are heterogeneous (and the next test would be modified Student's t -test if $N_1 = N_2$ or the Cochran t -test if $N_1 \neq N_2$; see Figure 30.2 to review the decision tree).

ASSUMPTIONS AND LIMITATIONS

1. This test could be considered as a two-group equivalent of Bartlett's test.
2. If the test statistic is close to 1.0, the results are (of course) not significant.
3. The test assumes normality and independence of data.

30.7.4 Analysis of Variance (ANOVA)

ANOVA is used for comparison of three or more groups of continuous data when the variances are homogeneous and the data are independent and normally distributed.

A series of calculations are required for ANOVA, starting with the values within each group being added ($\sum X$) and then these sums being added ($\sum \sum X$). Each figure within the groups is squared, and these squares are then summed ($\sum X^2$) and these sums added ($\sum \sum X^2$). Next the "correction factor" (CF) can be calculated from the following formula:

$$CF = \frac{\left(\sum_{i=1}^K \sum_{j=1}^N X \right)^2}{N_1 + N_2 + \dots + N_k}$$

where N is the number of values in each group and K is the number of groups. The total sum of squares (SS) is then determined as follows:

$$SS_{\text{total}} = \sum_{i=1}^K \sum_{j=1}^N X^2 - CF$$

In turn, the SS between groups (bg) is found from

$$SS_{\text{bg}} = \frac{(\sum X_1)^2}{N_1} + \frac{(\sum X_2)^2}{N_2} + \dots + \frac{(\sum X_k)^2}{N_k} - CF$$

The SS within groups (wg) is then the difference between the last two figures, or

$$SS_{\text{wg}} = SS_{\text{total}} - SS_{\text{bg}}$$

Now, there are three types of df to determine. The first, total df, is the total number of data within all groups under analysis minus one ($N_1 + N_2 + \dots + N_k - 1$). The second figure (the df between groups) is the number of groups minus one ($K - 1$). The last figure (the df within groups or "error df") is the difference between the first two figures ($df_{\text{total}} - df_{\text{bg}}$).

The next set of calculations requires determination of the two mean squares (MS_{bg} and MS_{wg}). These are the respective SS values divided by the corresponding df figures ($MS = SS/df$).

The final calculation is that of the F ratio. For this, the MS between groups is divided by the MS within groups ($F = MS_{bg}/MS_{wg}$).

A table of the results of these calculations would appear as follows:

| | df | SS | MS | F |
|-------|----|---------|---------|------|
| bg | 3 | 0.04075 | 0.01358 | 4.94 |
| wg | 12 | 0.03305 | 0.00275 | |
| Total | 15 | 0.07380 | | |

For interpretation, the F -ratio value obtained in the ANOVA is compared to a table of F values. If $F \leq 1.0$, the results are not significant and comparison with the table values is not necessary. The df for the greater mean square (MS_{bg}) are indicated along the top of the table. Then read down the side of the table to the line corresponding to the df for the lesser mean square (MS_{wg}). The figure shown at the desired significance level (traditionally 0.05) is compared to the calculated F value. If the calculated number is smaller, there are no significant differences among the groups being compared. If the calculated value is larger, there is some difference but further (*post hoc*) testing will be required before we know which groups differ significantly.

ASSUMPTIONS AND LIMITATIONS

1. What is presented here is the workhorse of toxicology—the one-way analysis of variance. Many other forms exist for more complicated experimental designs.
2. The test is robust for moderate departures from normality if the sample sizes are large enough. Unfortunately, this is rarely the case in toxicology.
3. ANOVA is robust for moderate departures from equality of variances (as determined by Bartlett's test) if the sample sizes are approximately equal.
4. It is not appropriate to use a t -test (or a two groups at a time version of ANOVA) to identify where significant differences are within the design group. A multiple comparison *post hoc* method must be used.

30.7.5 Post Hoc Tests

There is a wide variety of *post hoc* tests available to analyze data after finding significant result in an ANOVA. Each of these tests has advantages and disadvantages, proponents, and critics. Four of the tests are commonly used in toxicology and will be presented or previewed here. These are Dunnett's t -test and Williams' t -test. Two other tests which are available in many statistical packages are Turkey's

method and the Student–Newman–Keuls method (Zar, 1974, pp. 151–161).

If ANOVA reveals no significance, it is not appropriate to proceed to perform a *post hoc* test in hope of finding differences. To do so would only be another form of multiple comparisons, increasing the type I error rate beyond the desired level.

30.7.6 Duncan's Multiple Range Test

Duncan's (1955) is used to compare groups of continuous and randomly distributed data (such as body weights and organ weights). The test normally involves three or more groups taken one pair at a time. It should only follow observation of a significant F value in the ANOVA and can serve to determine which group (or groups) differs significantly from which other group (or groups).

There are two alternative methods of calculation. The selection of the proper one is based on whether the number of data (N) is equal or unequal in the groups.

30.7.7 Groups with Equal Number of Data ($N_1 = N_2$)

Two sets of calculations must be carried out: first, the determination of the difference between the means of pairs of groups and, second, the preparation of a probability rate against which each difference in means is compared (as shown in the first of the two examples in this section).

The means (averages) are determined (or taken from the ANOVA calculation) and ranked in either decreasing or increasing order. If two means are the same, they take up two equal positions (thus, for four means we could have ranks of 1, 2, 2, and 4 rather than 1, 2, 3, and 4). The groups are then taken in pairs and the differences between the means ($\bar{X}_1 - \bar{X}_2$), expressed as positive numbers, are calculated. Usually, each pair consists of a test group and the control group; through multiple tests groups may be intracompared if so desired. The relative rank of the two groups being compared must be considered. If a test group is ranked "2" and the control group is ranked "1," then we say that there are two places between them, while if the test group were ranked "3," then there would be three places between it and the control.

To establish the probability table, the SEM must be calculated as presented earlier or as

$$\sqrt{\frac{\text{Error mean square}}{N}} = \sqrt{\frac{\text{Mean square within group}}{N}}$$

where N is the number of animals or replications per dose level. The mean square within groups (MS_{wg}) can be calculated from the information given in the ANOVA procedure (refer to the earlier section on ANOVA). The SEM is then

multiplied by a series of table values (Harter, 1960; Beyer, 1976) to set up a probability table. The table values used for the calculations are chosen according to the probability levels (note that the tables have sections for 0.05, 0.01, and 0.001 levels) and the number of means apart from the groups being compared and the number of “error” df. The “error” df is the number of df within the groups. This last figure is determined from the ANOVA calculation and can be taken from ANOVA output. For some values of df, the table values are not given and should thus be interpolated.

30.7.8 Groups with Unequal Number of Data ($N_1 \neq N_2$)

This procedure is very similar to that discussed in the preceding text. As before, the means are ranked and the differences between the means are determined ($\bar{X}_1 - \bar{X}_2$). Next, weighing values (“ a_{ij} ” values) are calculated for the pairs of groups being compared in accordance with

$$a_{ij} = \sqrt{\frac{2N_i N_j}{(N_i + N_j)}} = \sqrt{\frac{2N_1 N_2}{(N_1 + N_2)}}$$

This weighting value for each pair of groups is multiplied by ($\bar{X}_1 - \bar{X}_2$) for each value to arrive at a “ t ” value. It is the “ t ” that will later be compared to a probability table.

The probability table is set up as before except that, instead of multiplying the appropriate table values by SEM, SEM^2 is used. This is equal to $\sqrt{MS_{wg}}$.

For the desired comparison of two groups at a time, the ($\bar{X}_1 - \bar{X}_2$) value (if $N_1 = N_2$) is compared to the appropriate probability table. Each comparison must be made according to the number of places between the means. If the table value is larger at the 0.05 level, the two groups are not considered to be statistically different. If the table value is smaller, the groups are different and the comparison is repeated at lower levels of significance. Thus, the degree of significance may be determined. We might have significant differences at 0.05 but not at 0.01, in which case the probability would be represented at $0.05 > p > 0.01$.

ASSUMPTIONS AND LIMITATIONS

1. Duncan’s assures a set alpha level or type I error rate for all tests when means are separated by no more than ordered step increases. Preserving this alpha level means that the test is less sensitive than some others, such as the Student–Newman–Keuls. The test is inherently conservative and not resistant or robust.

30.7.9 Scheffe’s Multiple Comparisons

Scheffe’s is another *post hoc* comparison method for groups of continuous and randomly distributed data. It also normally involved three or more groups (Scheffe, 1959; Harris, 2013). It is widely considered a more powerful significance test than Duncan’s.

Each *post hoc* comparison is tested by comparing an obtained test value (F_{contr}) with the appropriate critical F value at the selected level of significance (the table F value multiplied by $K - 1$ for an F with $K - 1$ and $N - K$ df²). F_{contr} is computed as follows:

- a. Compute the mean for each sample (group).
- b. Denote the residual mean square by MS_{wg} .
- c. Compute the test statistic as

$$F_{\text{contr}} = \frac{(C_1 \bar{X}_1 + C_2 \bar{X}_2 + \dots + C_k \bar{X}_k)^2}{(K - 1) MS_{wg} (C_1^2 / n_1 + \dots + C_k^2 / n_k)}$$

where C_k is the comparison number such that the sum of $C_1, C_2, \dots, C_k = 0$.

ASSUMPTIONS AND LIMITATIONS

1. The Scheffe procedure is robust to moderate violations of the normality and homogeneity of variance assumptions.
2. It is not formulated on the basis of groups with equal numbers (as one of Duncan’s procedures is), and if $N_1 \neq N_2$ there is no separate weighing procedure.
3. It tests all linear contrasts among the population means (the other three methods confine themselves to pairwise comparison, except they use a Bonferroni-type correlation procedure).
4. The Scheffe procedure is powerful because of its robustness, yet it is very conservative. Type I error (the false-positive rate) is held constant at the selected test level for each comparison.

30.7.10 Dunnett’s t -Test

Dunnett’s t -test (Dunnett, 1955, 1964) has as its starting point the assumption that what is desired is a comparison of each of several means with one other mean and only one other mean, in other words, that one wishes to compare each and every treatment group with the control group, but not compare treatment groups with each other. The problem here is that, in toxicology, one is frequently interested in comparing treatment groups with other treatment groups. However, if one does want only to compare treatment groups versus a

control group, Dunnett's is a useful approach. In a study with K groups (one of them being the control), we will wish to make $K-1$ comparisons. In such a situation, we want to have a P level for the entire set of $K-1$ decisions (not for each individual decision). Dunnett's distribution is predicated on this assumption. The parameters for utilizing a Dunnett's table, such as those found in his original article, are K (as in the preceding text) and the number of df for mean square with groups (MS_{wg}). The test value is calculated as

$$t = \frac{|T_j - T_i|}{\sqrt{2MS_{wg}/n}}$$

where n is the number of observations in each of the groups. The mean square within group (MS_{wg}) is as we have defined it previously; T_j is the control group mean and T_i is the mean of, in order, each successive test group observation. Note that one uses the absolute value of the positive number resulting from subtracting T_i from T_j . This is to ensure a positive number for our final t .

ASSUMPTIONS AND LIMITATIONS

1. Dunnett's seeks to ensure that the type I error rate will be fixed at the desired level by incorporating CF into the design of the test value table.
2. Treated group sizes must be approximately equal.

30.7.11 Williams' t -Test

Williams' t -test (Williams, 1971, 1972) is popular, although its use is quite limited in toxicology. It is designed to detect the highest level (in a set of dose/exposure levels) at which there is no significant effect. It assumes that the response of interest (such as change in body weights) occurs at higher levels, but not at lower levels, and that the responses are monotonically ordered so that $X_0 \leq X_1 \leq \dots \leq X_k$. This is, however, frequently not the case. The Williams' technique handles the occurrence of such discontinuities in a response series by replacing the offending value and the value immediately preceding it with weighted average values. The test also is adversely affected by any mortality at high-dose levels. Such mortalities "impose a severe penalty, reducing the power of detecting an effect not only at level K but also at all lower doses" (Williams, 1972, p. 529). Accordingly, it is not generally applicable in toxicology studies.

30.7.12 Analysis of Covariance

ANCOVA is a method for comparing sets of data which consist of two variables (treatment and effect, with our effect variable being called the "variate"), when a third variable

(called the "covariate") exists which can be measured but not controlled and which has a definite effect on the variable of interest. In other words, it provides an indirect type of statistical control, allowing us to increase the precision of a study and to remove a potential source of bias. One common example of this is in the analysis of organ weights in toxicity studies. Our true interest here is the effect of our dose or exposure level on the specific organ weights, but most organ weights also increase (in the young, growing animals most commonly used in such studies) in proportion to increases in animal body weight. As we are not here interested in the effect of this covariate (body weight), we measure it to allow for adjustment. We must be careful before using ANCOVA, however, to ensure that the underlying nature of the correspondence between the variate and covariate is such that we can rely on it as a tool for adjustments (Anderson et al., 1980; Kotz and Johnson, 1982).

Calculation is performed in two steps. The first is a type of linear regression between the variate Y and the covariate X .

This regression, performed as described under the linear regression section, gives us the model

$$Y = a_1 + BX + e$$

which in turn allows us to define adjusted means (\bar{Y} and \bar{X}) such that $\bar{Y}_{1a} = \bar{Y}_1 - (\bar{X}_1 - \bar{X})$.

If we consider the case where K treatments are being compared such that $K=1, 2, \dots, K$, and we let X_{ik} and Y_{ik} represent the predictor and predicted values for each individual i in group k , we can let \bar{X}_k and \bar{Y}_k be the means. Then, we define the between-group (for treatment) SS and cross products as

$$T_{xx} = \sum_{k=1}^K n_k (\bar{X}_k - \bar{X})^2$$

$$T_{yy} = \sum_{k=1}^K n_k (\bar{Y}_k - \bar{Y})^2$$

$$T_{xy} = \sum_{k=1}^K n_k (\bar{X}_k - \bar{X})(\bar{Y}_k - \bar{Y})$$

In a like manner, within-group sums of squares and cross products are calculated as

$$\sum xx = \sum_{k=1}^K \sum_i (X_{ik} - \bar{X}_k)^2$$

$$\sum yy = \sum_{k=1}^K \sum_i (Y_{ik} - \bar{Y}_k)^2$$

$$\sum xy = \sum_{k=1}^K \sum_i (X_{ik} - \bar{X}_k)(Y_{ik} - \bar{Y}_k)$$

where i indicates the sum from all the individuals within each group; f =total number of subjects minus number of groups

$$S_{xx} = T_{xx} + \sum_{xx}$$

$$S_{yy} = T_{yy} + \sum_{yy}$$

$$S_{xy} = T_{xy} + \sum_{xy}$$

With these in hand, we can then calculate the residual mean squares of treatments (St^2) and error (Se^2)

$$St^2 = \frac{T_{yy} - (S_{xy}^2 / S_{xx}) + (\sum_{xy}^2 / \sum_{xx})}{lc - 1}$$

$$Se^2 = \frac{\left[\sum_{yy} - \left(\sum_{xy}^2 / \sum_{xx} \right) \right]}{f - 1}$$

These can be used to calculate an F -statistic to test the null hypothesis that all treatment effects are equal:

$$F = \frac{St^2}{Se^2}$$

The estimated regression coefficient of Y or X is

$$B = \frac{\sum_{xy}}{\sum_{xx}}$$

The estimated standard error (Se) for the adjusted difference between two groups is given by

$$Sd = Se \frac{1}{n_j} + \frac{1}{n_j} + \frac{(X_i - X_j)^2}{\sum_{xx}}$$

where n_0 and n_1 are the sample sizes of the two groups. A test of the null hypothesis that the adjusted differences between the groups are zero is provided by

$$t = \frac{Y_1 - Y_0 - B(X_1 - X_0)}{Sd}$$

The test value for the t is then looked up in the t -distribution table with $f-1$ df.

Computation is markedly simplified if all the groups are of equal size.

ASSUMPTIONS AND LIMITATIONS

1. The underlying assumptions for ANCOVA are fairly rigid and restrictive. The assumptions include the following:
 - a. The slopes of the regression lines of a Y and X are equal from group to group. This can be examined visually or formally (i.e., by a test). If this condition is not met, ANCOVA cannot be used.
 - b. The relationship between X and Y is linear.
 - c. The covariate X is measured without error. Power of the test declines as error increases.
 - d. There are no unmeasured confounding variables.
 - e. The errors inherent in each variable are independent of each other. Lack of independence effectively (but to an immeasurable degree) reduces sample size.
 - f. The variances of the errors in groups are equivalent between groups.
 - g. The measured data that form the groups are normally distributed. ANCOVA is generally robust to departures from normality.
2. Of the seven assumptions provided earlier, the most serious are the first four.

30.7.13 Modeling

The mathematical modeling of biological systems, restricted even to the field of toxicology, is an extremely large and vigorously growing area. Broadly speaking, modeling is the principal conceptual tool by which toxicology seeks to develop as a mechanistic science. In an iterative process, models are developed or proposed, tested by experiment, reformulated, and so on in a continuous cycle. Such a cycle could also be described as two related types of modeling—explanatory (where the concept is formed) and correlative (where data are organized and relationships derived). An excellent introduction to the broader field of modeling of biological systems can be found in Gold (1977) or Gad (2005).

In toxicology, modeling is of prime interest in seeking to relate a treatment variable with an effect variable and, from the resulting model, predict effects at exact points where no experiment has been done (but in the range where we have performed experiments, such as “determining” LD_{50}), to estimate how good our prediction is and, occasionally, simply to determine if a pattern of effects is related to a pattern of treatment.

For use in prediction, the techniques of linear regression, probit/logit analysis (a special case of linear regression), moving averages (an efficient approximation method), and

nonlinear regression (for doses where data cannot be made to fit a linear pattern) are presented. For evaluating the predictive value of these models, both the correlation coefficient (for parametric data) and Kendall's rank correlation (for nonparametric data) are given. And finally, the concept of trend analysis is introduced and a method presented.

When we are trying to establish a pattern between several data points (whether this pattern is in the form of a line or a curve), what we are doing is interpolating. It is possible for any given set of points to produce an infinite set of lines or curves which pass near (for lines) or through (for curves) the data points. In most cases, we cannot actually know the "real" pattern. So we apply a basic principle of science—Occam's razor. We use the simplest explanation (or, in this case, model) which fits the facts (or data). A line is, of course, the simplest pattern to deal with and describe, so fitting the best line (linear regression) is the most common form of model in toxicology.

30.7.14 Linear Regression

Foremost among the methods for interpolating within a known data relationship is regression—the fitting of a line or curve to a set of known data points on a graph and the interpolation ("estimation") of this line or curve in areas where we have no data points. The simplest of these regression models is that of linear regression (valid when increasing the value of one variable changes the value of the related variable in a linear fashion, either positively or negatively). This is the case we will explore here, using the method of least squares.

Given that we have two sets of variables, x (say, mg/kg of test material administered) and y (say, percentage of animals so dosed that die), what is required is solving for a and b in the equation $Y_i = a + bx_i$ (where the uppercase Y_i is the fitted value of y_i at x_i and we wish to minimize $(y_i - Y_i)^2$). So we solve the equations

$$b = \frac{\sum x_i y_i - n \bar{x} \bar{y}}{\sum x_i^2 - n \bar{x}^2} \quad \text{and} \quad a = \bar{y} - b \bar{x}$$

where a is the y intercept, b is the slope of the line, and n is the number of data points.

Note that in actuality, dose-response relationships are often not linear, and instead we must use either a transform (to linearize the data) or a nonlinear regression method (Gallant, 1975).

Note also that we can use the correlation test statistic (to be described in the correlation coefficient section) to determine if the regression is significant and, therefore, valid at a defined level of certainty. A more specific test for significance would be the linear regression ANOVA (Pollard, 1977). To do so we start by developing the appropriate ANOVA table.

Finally, we might wish to determine the CI for our regression line—that is, given a regression line with calculated

values for Y_i given x_i , within what limits may we be certain (with, say, a 95% probability) what the real value of Y_i is?

If we denote the residual mean square in the ANOVA by s^2 , the 95% confidence limits for a (denoted by A , the notation for the true—as opposed to the estimated—value for this parameter) are calculated as

$$t_{n-2} = \frac{a - A}{\sqrt{\left[s^2 \left(\sum x_i^2 \right) / n \sum x_i^2 - n^2 \bar{x}^2 \right]}}$$

$$\frac{9.2 - A}{\sqrt{\left[8.8(51) / \left[4(51) - (16)(10.562) \right] \right]}} = \frac{9.2 - A}{\sqrt{448 / 35.008}}$$

$$= \frac{9.2 - A}{3.58} = -4.303$$

$$9.2 - A = -15.405$$

$$A = 9.2 - 15.405$$

ASSUMPTIONS AND LIMITATIONS

1. All the regression methods are for interpolation, not extrapolation; that is, they are valid only in the range that we have data—the experimental region—not beyond.
2. The method assumes that the data are independent and normally distributed, and it is sensitive to outliers. The x -axis (or horizontal) component plays an extremely important part in developing the least square fit. All points have equal weight in determining the height of a regression line, but extreme x -axis values unduly influence the slope of the line.
3. A good fit between a line and a set of data (i.e., a strong correlation between treatment and response variables) does not imply any casual relationship.
4. It is assumed that the treatment variable can be measured without error, that each data point is independent, that variances are equivalent, and that a linear relationship does not exist between the variables.
5. The many excellent texts on regression, which is a powerful technique, include Draper and Smith (1998) and Montgomery and Smith (1983), which are not overly rigorous mathematically.

30.7.15 Probit/Log Transforms and Regression

As we noted in the preceding section, dose-response problems (among the most common interpolation problems encountered in toxicology) rarely are straightforward enough to make a valid linear regression directly from the raw data. The most common valid interpolation methods are based upon probability ("probit") and logarithmic ("log")

value scales, with percentage responses (death, tumor incidence, etc.) being expressed on the probit scale, while doses (Y_i) are expressed on the log scale. There are two strategies for such an approach. The first is based on transforming the data to these scales and then doing a weighted linear regression on the transformed data (if one does not have access to a computer or a high-powered programmable calculator, the only practical strategy is not to assign weights). The second requires the use of algorithms (approximate calculation techniques) for the probit value and regression process and is extremely burdensome to perform manually.

Our approach to the first strategy requires that we construct a table with the pairs of values of x_i and y_i listed in order of increasing values of Y_i (percentage response). Beside each of these columns, a set of blank columns should be left so that the transformed values may be listed. We then simply add the columns described in the linear regression procedure. Log and probit values may be taken from any of a number of sets of tables (such as provided in Appendix A), and the rest of the table is then developed from these transformed x'_i and y'_i values (denoted as x'_i and y'_i). A standard linear regression is then performed.

The second strategy we discussed has been broached by a number of authors (Bliss, 1935; Litchfield and Wilcoxon, 1949; Prentice, 1976; Finney, 1977). All of these methods, however, are computationally cumbersome. It is possible to approximate the necessary iterative process using the algorithms developed by Abramowitz and Stegun (1964), but even this merely reduces the complexity to a point where the procedure may be readily programmed on a small computer or programmable calculator.

ASSUMPTIONS AND LIMITATIONS

1. The probit distribution is derived from a common error function, with the midpoint (50% point) moved to a score of 5.00.
2. The underlying frequency distribution becomes asymptotic as it approaches the extremes of the range. That is, in the range of 16–84%, the corresponding probit values change gradually—the curve is relatively linear. But beyond this range, they change ever more rapidly as they approach either 0 or 100%. In fact, there are no values for either of these numbers.
3. A normally distributed population is assumed, and the results are sensitive to outliers.

30.7.16 Nonlinear Regression

More often than not in toxicology we find that our data demonstrate a relationship between two variables (such as age and body weight) which is not linear. That is, a change in one

variable (say, age) does not produce a directly proportional change in the other (e.g., body weight). But some form of relationship between the variables is apparent. If understanding such a relationship and being able to predict unknown points are of value, we have a pair of options available to us. The first, which was discussed and reviewed earlier, is to use one or more transformations to linearize our data and then to make use of linear regression. This approach, though most commonly used, has a number of drawbacks. Not all data can be suitably transformed; sometimes the transformations necessary to linearize the data require a cumbersome series of calculations; and the resulting linear regression is not always sufficient to account for the differences among sample values—there are significant deviations around the linear regression line (i.e., a line may still not give us a good fit to the data or do an adequate job of representing the relationship between the data). In such cases, we have available a second option—the fitting of data to some nonlinear function such as some form of the curve. This is, in general form, nonlinear regression and may involve fitting data to an infinite number of possible functions. But most often we are interested in fitting curves to a polynomial function of the general form

$$Y = a + bx + cx^2 + dx^2 + \dots$$

where x is the independent variable. As the number of powers of x increases, the curve becomes increasingly complex and will be able to fit a given set of data increasingly well.

Generally in toxicology, however, if we plot the log of a response (such as body weight) versus a linear scale of our dose or stimulus, we get one of four types of nonlinear curves. These are (Snedecor and Chocran, 2014):

1. Exponential growth, where
 $\log Y = A(B^x)$, such as the growth curve for the log phase of a bacterial culture
2. Exponential decay, where
 $\log Y = A(B^{-x})$, such as a radioactive decay curve
3. Asymptotic regression, where
 $\log Y = A - B(p^x)$, such as a first-order reaction curve
4. Logistic growth curve, where
 $\log Y = A/(1 + Bp^x)$, such as a population growth curve

In all these cases, A and B are constant, while p is a log transform. These curves are illustrated in Figure 30.7.

All four types of curves are fit by iterative processes—that is, best guess numbers are initially chosen for each of the constants and, after a fit is attempted, the constants are modified to improve the fit. This process is repeated until an acceptable fit has been generated. ANOVA or ANCOVA can be used to objectively evaluate the acceptability of it. Needless to say, the use of a computer generally accelerates such a curve-fitting process.

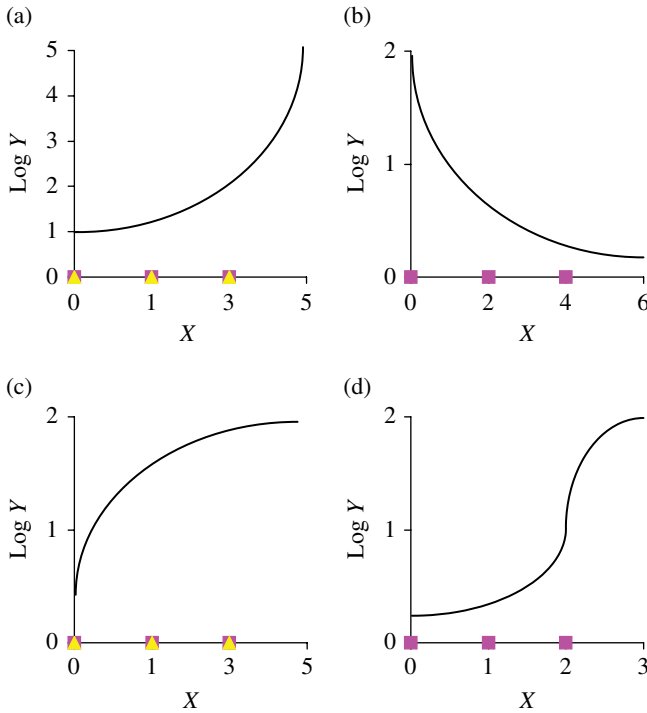


FIGURE 30.7 Common curvilinear curves: (a) exponential growth law, $\log Y = A(B^X)$; (b) exponential decay law, $\log Y = A(B^{-X})$; (c) asymptotic regression, $\log Y = A - B(\Psi^X)$; and (d) logistic growth law, $\log Y = A/(1 + B\Psi^X)$.

ASSUMPTIONS AND LIMITATIONS

1. The principle of using least squares may still be applicable in fitting the best curve, if the assumptions of normality, independence, and reasonably error-free measurement of response are valid.
2. Growth curves are best modeled using a nonlinear method.

30.7.17 Correlation Coefficient

The correlation procedure is used to determine the degree of linear correlation (direct relationship) between two groups of continuous (and normally distributed) variables; it will indicate whether there is any statistical relationship between the variables in the two groups. For example, we may wish to determine if the liver weights of dogs on a feeding study are correlated with their body weights. Thus, we will record the body and liver weights at the time of sacrifice and then calculate the correlation coefficient between these pairs of values to determine if there is some relationship.

A formula for calculating the linear correlation coefficient (r_{xy}) is as follows:

$$r_{xy} = \frac{N \sum XY - (\sum X)(\sum Y)}{\sqrt{N \sum X^2 - (\sum X)^2} \sqrt{N \sum Y^2 - (\sum Y)^2}}$$

where X is each value for one variable (such as the dog body weights in the previous example), Y is the matching value for the second variable (the liver weights), and N is the number of pairs of X and Y . Once we have obtained r_{xy} , it is possible to calculate t_r , which can be used for more precise examination of the degree of significant linear relationship between the two groups. This value is calculated as follows:

$$t_r = \frac{r_{xy} \sqrt{N-2}}{\sqrt{1-r_{xy}^2}}$$

This calculation is also equivalent to $r = \text{sample covariance} / (S_x S_y)$, as was seen earlier under ANCOVA.

The value obtained for r_{xy} can be compared to table values (Snedecor and Cochran, 2014) for the number of pairs of data involved minus two. If the r_{xy} is smaller (at the selected test probability level, such as 0.05), the correlation is not significantly different from zero (no correlation). If r_{xy} is larger than the table value, there is a positive statistical relationship between the groups. Comparisons are then made at lower levels of probability to determine the degree of relationship (note that if r_{xy} is equal to either 1.0 or -1.0, there is complete correlation between the groups). If r_{xy} is a negative number and the absolute is greater than the table value, there is an inverse relationship between the groups, that is, a change in one group is associated with a change in the opposite direction in the second group of variables.

Since the comparison of r_{xy} with the table values may be considered a somewhat weak test, it is perhaps more meaningful to compare the t_r value with values in a t -distribution table for $N-2$ df, as is done for Student's t -test. This will give a more exact determination of the degree of statistical correlation between the two groups.

Note that this method examines only possible linear relationships between sets of continuous, normally distributed data.

ASSUMPTIONS AND LIMITATIONS

1. A strong correlation does not imply that a treatment causes an effect.
2. The distances of data points from the regression line are the portions of the data not "explained" by the model. These are called residuals. Poor correlation coefficients imply high residuals, which may be due to many small contributions (variations of data from the regression line) or a few large ones. Extreme values (outliers) greatly reduce correlation.
3. X and Y are assumed to be independent.
4. Feinstein (1979) has provided a fine discussion of the difference between correlation (or association of variables) and causation.

30.7.18 Kendall's Coefficient of Rank Correlation

Kendall's rank correlation, represented by τ (tau), should be used to evaluate the degree of association between two sets of data when the nature of the data is such that the relationship may not be linear. Most commonly, this is when the data are not continuous and/or normally distributed. An example of such a case is when we are trying to determine if there is a relationship between the length of hydra and their survival time in a test medium in hours. Both of our variables here are discontinuous, yet we suspect a relationship exists. Another common use is in comparing the subjective scoring done by two different observers.

Tau is calculated at $\tau = N/n(n-1)$ where n is the sample size and N is the count of ranks, calculated as $N = 4 \sum C_i - n(n-1)$, with the computing of C_i being demonstrated in the example.

If a second variable Y_2 is exactly correlated with the first variable Y_1 , then the variates Y_2 should be in the same order as the Y_1 variates. However, if the correlation is less than exact, the order of the variates Y_2 will not correspond entirely to that of Y_1 . The quantity N measures how well the second variable corresponds to the order of the first. It has a maximum value of $n(n-1)$ and a minimum value of $-n(n-1)$.

A table of data is set up with each of the two variables being ranked separately. Tied ranks are assigned as demonstrated earlier under the Kruskal–Wallis test. From this point, disregard the original variates and deal only with the ranks. Place the ranks of one of the two variables in rank order (from lowest to highest), paired with the rank values assigned for the other variable. If one (but not the other) variable has tied ranks, order the pairs by the variables without ties (Sokal and Rohlf, 1994).

The resulting value of tau will range from -1 to $+1$, as does the familiar parametric correlation coefficient, r .

ASSUMPTION AND LIMITATION

1. A very robust estimator which does not assume normality, linearity, or minimal error of measurement.

30.7.19 Trend Analysis

Trend analysis is a collection of techniques that have been “discovered” by toxicology since the mid-1970s (Tarone, 1975). The actual methodology dates back to the mid-1950s (Cox and Stuart, 1955).

Trend analysis methods are a variation on the theme of regression testing. In the broadest sense, the methods are used to determine whether a sequence of observations taken over an ordered range of a variable (most commonly time) exhibit some form of pattern of change (either an increase or upward trend) associated with another variable of interest (in toxicology, some form or measure of dosage and exposure).

Trend corresponds to sustained and systematic variations over a long period of time. It is associated with the structural causes of the phenomenon in question, for example, population growth, technological progress, new ways of organization, or capital accumulation.

The identification of trend has always posed a serious statistical problem. The problem is not one of mathematical or analytical complexity but of conceptual complexity. This problem exists because the trend and the remaining components of a time series are latent (nonobservables) variables, and, therefore, assumptions must be made on their behavioral pattern. The trend is generally thought of as a smooth and slow movement over a long term. The concept of “long” in this connection is relative and what is identified as trend for a given series span might well be part of a long cycle once the series is considerably augmented. Often, a long cycle is treated as a trend because the length of the observed time series is shorter than one complete face of this type of cycle.

The ways in which data are collected in toxicology studies frequently serve to complicate trend analysis, as the length of time for the phenomena underlying a trend to express themselves is frequently artificially censored.

To avoid the complexity of the problem posed by a statistically vague definition, statisticians have resorted to two simple solutions: one consists of estimating trend and cyclic fluctuations together, calling this combined movement *trend-cycle*; the other consists of defining the trend in terms of the series length, denoting it as the longest nonperiodic movement.

Within the large class of models identified for trend, we can distinguish two main categories, deterministic trends and stochastic trends.

30.8 METHODS FOR THE REDUCTION OF DIMENSIONALITY

Techniques for the reduction of dimensionality are those that simplify the understanding of data, either visually or numerically, while causing only minimal reductions in the amount of information present. These techniques operate primarily by pooling or combining groups of variables into single variables, but may also entail the identification and elimination of low-information-content (or irrelevant) variables.

Descriptive statistics (calculations of means, SD, etc.) are the simplest and most familiar form of reduction of dimensionality. Here we first need to address classification, which provides the general conceptual tools for identifying and quantifying similarities and differences between groups of things which have more than a single linear scale of measurement in common (e.g., which both have been determined to have and lack a number of enzyme activities). Then we will consider two collections of methodologies which combine graphic and computational methods, multidimensional/nonmetric scaling,

and cluster analysis. Multidimensional scaling (MDS) is a set of techniques for quantitatively analyzing similarities, dissimilarities, and distances between data in a display-like manner. Nonmetric scaling is an analogous set of methods for displaying and relating data when measurements are nonquantitative (the data are described by attributes or ranks). Cluster analysis is a collection of graphic and numerical methodologies for classifying things based on the relationships between the values of the variables that they share.

The final pair of methods for reduction of dimensionality which will be tackled in this chapter are Fourier analysis and the life-table analysis. Fourier analysis seeks to identify cyclic patterns in data and can analyze the patterns or the residuals after the patterns are taken out. Life-table analysis techniques are directed to identifying and quantitating the time course of risks (such as death or the occurrence of tumors).

30.8.1 Classification

Classification is both a basic concept and a collection of techniques which are necessary prerequisites for further analysis of data when the members of a set of data are (or can be) each described by several variables. At least some degree of classification (which is broadly defined as the dividing of the members of a group into smaller groups in accordance with a set of decision rules) is necessary prior to any data collection. Whether formally or informally, an investigator has to decide which things are similar enough to be counted as the same and develop rules for governing collection procedures. Such rules can be as simple as "measure and record body weights only of live animals on study" or as complex as those demonstrated by the expanded weighting classification procedure demonstrated in the succeeding text. Such a classification also demonstrates that the selection of which variables to measure will determine the final classification of data.

Expanded Weighting Procedure

| | | |
|------|--|-------------|
| I. | Does the animal belong to the desired species? | Yes/no |
| II. | Is the animal a member of study group? | Yes/no |
| III. | Is the animal alive? | Yes/no |
| IV. | Which group does the animal belong to? | |
| | A. Control | |
| | B. Low dose | |
| | C. Intermediate dose | |
| | D. High dose | |
| V. | What sex is the animal? | Male/female |
| VI. | Is the measured weight in acceptable range? | Yes/no |

Classifications of data have two purposes (Gordon, 1981; Hartigan, 1983): data simplification (also called a descriptive function) and prediction. Simplification is necessary

because there is a limit to both the volume and complexity of data that the human mind can comprehend and deal with conceptually. Classification allows us to attach a label (or name) to each group of data, to summarize the data (i.e., assign individual elements of data to groups and to characterize the population of the group), and to define the relationships between groups (i.e., develop a taxonomy).

Prediction, meanwhile, is the use of summaries of data and knowledge of the relationships between groups to develop hypotheses as to what will happen when further data are collected (as when more animals or people are exposed to an agent under defined conditions) and as to the mechanisms which cause such relationships to develop. Indeed, classification is the prime device for the discovery of mechanisms in all of science. A classic example of this was Darwin's realization that there were reasons (the mechanisms of evolution) behind the differences and similarities in species which had caused Linnaeus to earlier develop his initial modern classification scheme (or taxonomy) for animals.

To develop a classification, one first sets bounds wide enough to encompass the entire range of data to be considered but not unnecessarily wide. This is typically done by selecting some global variables (variables every piece of data have in common) and limiting the range of each so that it just encompasses all the cases on hand. Then one selects a set of local variables (characteristics which only some of the cases have, say, the occurrence of certain tumor types, enzyme activity levels, or dietary preferences and which thus serve to differentiate between groups). Data are then collected, and a system for measuring differences and similarities is developed. Such measurements are based on some form of measurement of distance between two cases (x and y) in terms of each single variable scale. If the variable is a continuous one, then the simplest measure of distance between two pieces of data is the Euclidean distance, ($d(x, y)$), defined as

$$d(x, y) = \sqrt{(x_i - y_i)^2}$$

For categorical or discontinuous data, the simplest distance measure is the matching distance, defined as

$$d(x, y) = \text{number of times } x_i \neq y_i.$$

After we have developed a table of such distance measurements for each of the local variables, some weighting factor is assigned to each variable. A weighting factor seeks to give greater importance to those variables which are believed to have more relevance or predictive value. The weighted variables are then used to assign each piece of data to a group. The actual act of developing numerically based classifications and assigning data members to them is the realm of cluster analysis and will be discussed later in this chapter. Classification of biological data based on qualitative factors

has been well discussed (Glass, 1975; Schaper et al., 1985) and does an excellent job of introducing the entire field and mathematical concepts.

Relevant examples of the use of classification techniques range from the simple to the complex. Schaper et al. (1985) developed and used a very simple classification of response methodology to identify those airborne chemicals which alter the normal respiratory response induced by CO₂. At the other end of the spectrum, Kowalski and Bender (1972) developed a more mathematically based system to classify chemical data (a methodology they termed pattern recognition).

30.8.2 Statistical Graphics

The use of graphics in one form or another in statistics is the single most effective and robust statistical tool and, at the same time, one of the most poorly understood and improperly used (Cleveland, 1994).

Graphs are used for one of four major purposes. Each of the four is a variation on the central theme of making complex data easier to understand and use. These four major functions are exploration, analysis, communication and display of data, and graphic aids. Exploration (which may be

simply summarizing data or trying to expose relationships between variables) is determining the characteristics of data sets and deciding on one or more appropriate forms of further analysis, such as the scatterplot. Analysis is the use of graphs to formally evaluate some aspect of the data, such as whether there are outliers present or if an underlying assumption of a population distribution is fulfilled. As long ago as 1960 (Anderson, 1960), some 18 graphic methods for analyzing multivariate data relationships were developed and proposed. Table 30.7 presents a summary of major graphic techniques that are available.

Communication and display of data is the most commonly used function of statistical graphics in toxicology, whether used for internal reports, presentations at meetings, or formal publications in the literature. In communicating data, graphs should not be used to duplicate data that are presented in tables, but rather to show important trends and/or relationships in the data. Though such communication is most commonly of a quantitative compilation of actual data, it can also be used to summarize and present the results of statistical analysis. The fourth and final function of graphics is one that is largely becoming outdated as microcomputers become more widely available. Graphic aids to calculation

TABLE 30.7 Forms of Statistical Graphics (by Function)

| Data Summary | Two Variables | Three or More Variables |
|--------------------------------|---|----------------------------------|
| <i>Exploration</i> | | |
| Box and whisker plot | Autocorrelation plot | Biplot |
| Histogram ^a | Cross-correlation plot | Cluster trees ^a |
| Dot array diagram | Scatterplot ^a | Labeled scatterplot ^a |
| Frequency polygon | Sequence plot | Glyphs and metroglyphs |
| Ogive | | Fourier plots |
| Stem-and-leaf diagram | | Similarity and preference maps |
| | | MDS displays |
| | | Weathervane plot |
| Distribution Assessment | Model Evaluation and Assumption Verification | Decision Making |
| <i>Analysis</i> | | |
| Probability plot | Average versus SD | Control chart |
| Q-Q plot | Component-plus-residual plot | Custom chart |
| P-P plot | Partial residual plot | Half-normal plot |
| Hanging histogram | Residual plots | Ridge trace |
| Rootagram | | Youden plot |
| Poissonness plot | | |
| Quantitative Graphics | Summary of Statistical Analyses | Graphic Aids |
| | <i>Communication and Display of Data</i> | |
| Line chart | Means plot | Confidence limits |
| Pictogram | Sliding reference distribution | Graph paper |
| Pie chart | Notched box plot | Power curves |
| Contour plot | Factor space/response | Nomographs |
| Stereogram | Interaction plot | Sample size curves |
| Color map | Contour plot | Trilinear coordinates |
| Histogram | Predicted response plot | |
| | Confidence region plot | |

^a Reviewed in the text of this chapter.

include nomograms (the classic example in toxicology of a nomogram is that presented by Litchfield and Wilcoxon (1949) for determining median effective doses) and extrapolating and interpolating data graphically based on plotted data.

There are many forms of statistical graphics (a partial list, classified by function, is presented in Table 30.6), and a number of these (such as scatterplots and histograms) can be used for each of a number of possible functions. Most of these plots are based on a Cartesian system (i.e., they use a set of rectangular coordinates), and our review of construction and use will focus on these forms of graphs.

Construction of a rectangular graph of any form starts with the selection of the appropriate form of graph followed by the laying out of the coordinates (or axes). Even graphs which are going to encompass multivariate data (i.e., more than two variables) generally have as their starting point two major coordinates. The vertical axis, or ordinate (also called the *Y*-axis), is used to present an independent variable. Each of these axes is scaled in the units of measure which will most clearly present the trends of interest in the data. The range covered by the scale of each axis is selected to cover the entire region for which data are presented. The actual demarking of the measurement scale along an axis should allow for easy and accurate assessment of the coordinates of any data point, yet should not be cluttered.

Actual data points should be presented by symbols which present the appropriate indicators of location, and if they represent a summary of data from a normal data population, it would be appropriate to present a symbol for the mean and some indication of the variability (or error) associated with that population, commonly by using "error bars" which present the SD (or standard error) from the mean. If, however, the data are not normal or continuous, it would be more appropriate to indicate location by the median and present the range or semiquartile distance for variability estimates. The symbols which are used to present data points can also be used to present a significant amount of additional information. At the simplest level a set of clearly distinct symbols (circles, triangles, squares, etc.) are very commonly used to provide a third dimension of data (most commonly treatment group). But by clever use of symbols, all sorts of additional information can be presented. Using a method such as Chernoff's faces (Chernoff, 1973), in which faces are used as symbols of the data points (and various aspects of the faces present additional data, such as the presence or absence of eyes denoting the presence or absence of a secondary pathological condition), it is possible to present a large number of different variables on a single graph.

The three other forms of graphs that are commonly used are histograms, pie charts, and contour plots.

Histograms are graphs of simple frequency distribution. Commonly, the abscissa is the variable of interest (such as life span or litter size) and is generally shown as classes or intervals of measurements (such as age ranges of 0–10,

10–20 weeks, etc.). The ordinate, meanwhile, is the incidence or frequency of observations. The result is a set of vertical bars, each of which represents the incidence of a particular set of observations. Measures of error or variability about each incidence are reflected by some form of error bar on top of or in the frequency bars, as shown in Figure 30.8. The size of class intervals may be unequal (in effect, one can combine or pool several small class intervals), but it is proper in such cases to vary the width of the bars to indicate differences in interval size.

Pie charts are the only common form of quantitative graphic technique which is not rectangular. Rather, the figure is presented as a circle out of which several "slices" are delimited. The only major use of the pie chart is in presenting a breakdown of the components of a group. Typically the entire set of data under consideration (such as total body weight) constitutes the pie, while each slice represents a percentage of the whole (such as the percentages represented by each of several organs). The total number of slices in a pie should be small for the presentation to be effective. Variability or error can be readily presented by having a subslice of each sector shaded and labeled accordingly.

Finally, there is the contour plot, which is used to depict the relationships in a three-variable, continuous data system. That is, a contour plot visually portrays each contour as a locus of the values of two variables associated with a constant value of the third variable. An example would be a relief map that gives both latitude and longitude of constant altitude using contour lines.

The most common misuse of graphs is to either conceal or exaggerate the extent of the difference by using inappropriately scaled or ranged axis. Tufte (1983) has termed a statistic for evaluating the appropriateness of scale size, the lie factor, calculated as the ratio of the shown effect size to the range of potential change or effect. An acceptable range for the lie factor is from 0.95 to 1.05. Less means the size of an effect is being understated, more that the effect is being exaggerated.

There are a number of excellent references available for those who would like to pursue statistical graphics more. Anscombe (1973) presents an excellent short overview, while others (Schmid, 1983; Tufte, 1983; Young, 1985; Tufte 1990, 1997) provide a wealth of information.

30.8.3 Multidimensional and Nonmetric Scaling

MDS is a collection of analysis methods for data sets which have three or more variables making up each data point. MDS displays the relationships of three or more dimensional extensions of the methods of statistical graphics.

MDS presents the structure of a set of objects from data that approximate the distances between pairs of the objects. The data, called similarities, dissimilarities, distances, or proximities, must be in such a form that the degree of

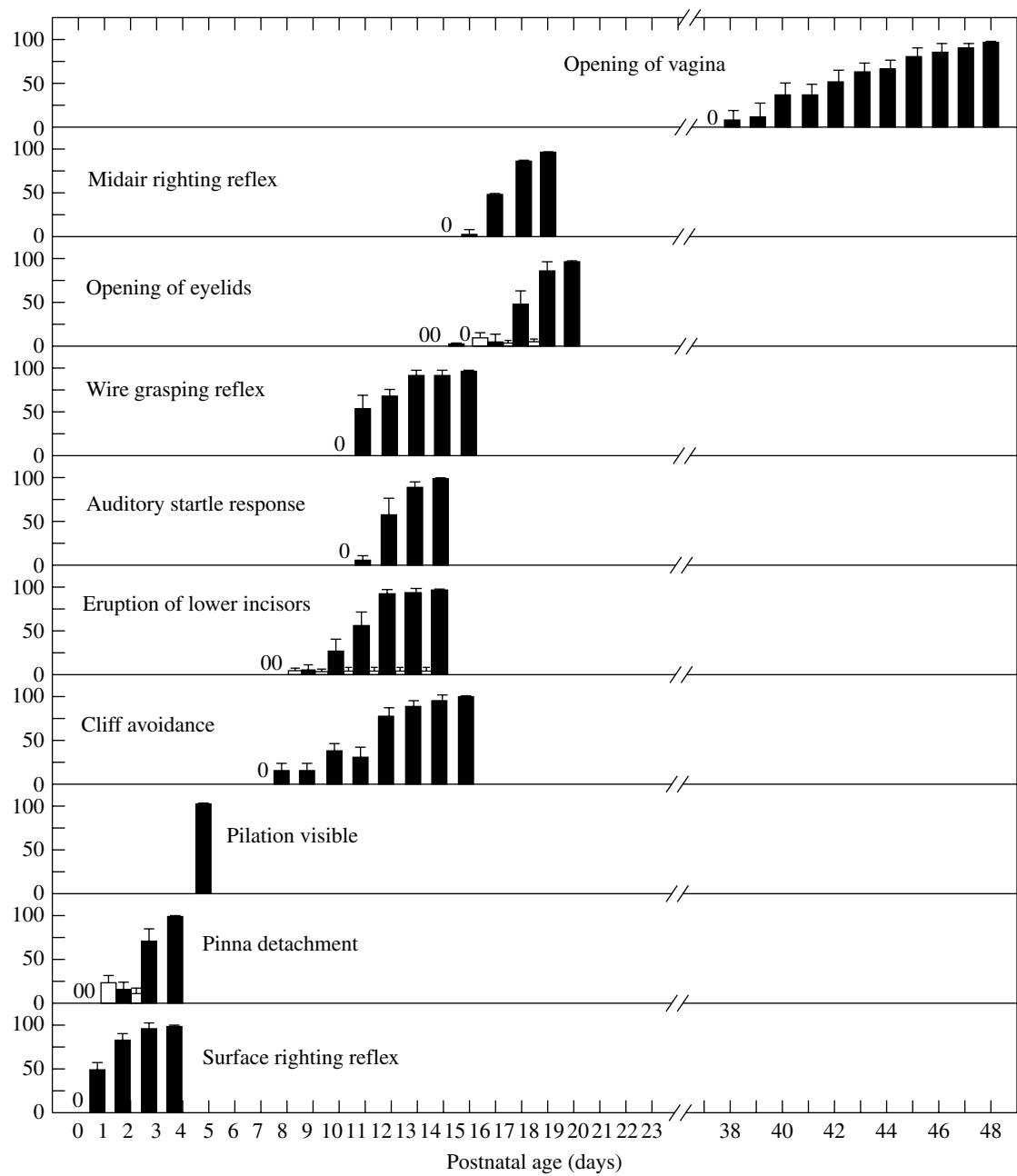


FIGURE 30.8 Acquisitions of postnatal development landmarks in rats.

similarities and differences between the pairs of the objects (each of which represents a real-life data point) can be measured and handled as a distance (remember the discussion of measures of distances under classifications). Similarity is a matter of degree—small differences between objects cause them to be similar (a high degree of similarity), while large differences cause them to be considered dissimilar (a small degree of similarity).

In addition to the traditional human conceptual or subjective judgments or similarity, data can be an “objective” similarity measure (the difference in weight between a pair

of animals) or an index calculated from multivariate data (the proportion of agreement in the results of a number of carcinogenicity studies). However, the data must always represent the degree of similarity of pairs of objects.

Each object or data point is represented by a point in a multidimensional space. These plots or projected points are arranged in this space so that the distances between pairs of points have the strongest possible relation to the degree of similarity among the pairs of objects. That is, two similar objects are represented by two points that are close together, and two dissimilar objects are represented by a pair of points

that are far apart. The space is usually a two- or three-dimensional Euclidean space, but may be non-Euclidean and may have more dimensions.

MDS is a general term which includes a number of different types of techniques. However, all seek to allow geometric analysis of multivariate data. The forms of MDS can be classified (Young, 1985) according to the nature of the similarities in the data. It can be qualitative (nonmetric) or quantitative (metric MDS). The types can also be classified by the number of variables involved and by the nature of the model used, for example, classical MDS (there is only one data matrix, and no weighting factors are used on the data), replicated MDS (more than one matrix and no weighting), and weighted MDS (more than one matrix and at least some of the data being weighted).

MDS can be used in toxicology to analyze the similarities and differences between effects produced by different agents, in an attempt to use an understanding of the mechanism underlying the actions of one agent to determine the mechanisms of the other agents. Actual algorithms and a good intermediate level presentation of MDS can be found in Davison (1983).

Nonmetric scaling is a set of graphic techniques closely related to MDS and definitely useful for the reduction of dimensionality. Its major objective is to arrange a set of objects (each object, for our purposes, consisting of a number of related observations) graphically in a few dimensions while retaining the maximum possible fidelity to the original relationships between members (i.e., values which are most different are portrayed as most distant). It is not a linear technique; it does not preserve linear relationships (i.e., A is not shown as twice as far from C as B, even though its "value difference" may be twice as much). The spacing (interpoint distances) are kept such that if the distance of the original scale between members A and B is greater than that between C and D, the distances on the model scale shall likewise be greater between A and B than between C and D. Figure 30.5, presented earlier, uses a form of this technique in adding a third dimension by using letters to present degrees of effect on the skin.

This technique functions by taking observed measures of similarity or dissimilarity between every pair of M objects and then finding a representation of the objects as points in Euclidean space that the interpoint distances in some sense "match" the observed similarities or dissimilarities by means of weighting constants.

30.8.4 Cluster Analysis

Cluster analysis is a quantitative form of classification. It serves to help develop decision rules and then use these rules to assign a heterogeneous collection of objects to a series of sets. This is almost entirely an applied methodology (as opposed to theoretical). The final result of cluster analysis is one of several forms of graphic displays and a methodology

(set of decision-classifying rules) for the assignment of new members into the classifications.

The classification procedures used are based on either density of population or distance between members. These methods can serve to generate a basis for the classification of a large number of dissimilar variables such as behavioral observations and compounds with distinct but related structures and mechanisms (Gad, 1984, 2005) or to separate tumor patterns caused by treatment from those caused by old age (Hammond et al., 1978).

There are five types of clustering techniques (Everitt et al., 2011; Romesburg, 1984):

- a. *Hierarchical Techniques* Classes are subclassified into groups, with the process being repeated at several levels to produce a tree which gives sufficient definition to groups.
- b. *Optimizing Techniques* Clusters are formed by optimization of a clustering criterion. The resulting classes are mutually exclusive; the objects are partitioned clearly into sets.
- c. *Density or Mode-Seeking Techniques* Clusters are identified and formed by locating regions in a graphic representation which contains concentrations of data points.
- d. *Clumping Techniques* A variation of density-seeking techniques in which assignment to a cluster is weighted on some variables, so that clusters may overlap in graphic projections.
- e. *Others* Methods which do not clearly fall into classes a–d.

Romesburg (1984) provides an excellent step-by-step guide to cluster analysis.

30.8.5 Fourier or Time Analysis

Fourier analysis (Bloomfield, 1976) is most frequently a univariate method used for either simplifying data (which is the basis for its inclusion in this chapter) or for modeling. It can, however, also be a multivariate technique for data analysis.

In a sense, it is like trend analysis; it looks at the relationship of sets of data from a different perspective. In the case of Fourier analysis, the approach is by resolving the time dimension variable in the data set. At the most simple level, it assumes that many events are periodic in nature, and if we can remove the variation in other variables because of this periodicity (by using Fourier transforms), we can better analyze the remaining variation from other variables. The complications to this are (i) there may be several overlying cyclic time-based periodicities and (ii) we may be interested in the time-cycle events for their own sake.

Fourier analysis allows one to identify, quantitate, and (if we wish) remove the time-based cycles in data (with their amplitudes, phases, and frequencies) by use of the Fourier transform:

$$nJ_i = x_i \exp(-iw_it)$$

where

n = length

J = the discrete Fourier transform for that case

x = actual data

i = increment in the series

w = frequency

t = time

30.8.6 Life Tables

Chronic *in vivo* toxicity studies are generally the most complex and expensive studies conducted by a toxicologist. Answers to a number of questions are sought in such a study—notably if a material results in a significant increase in mortality or in the incidence of tumors in those animals exposed to it. But we are also interested in the time course of these adverse effects (or risks). The classic approach to assessing these age-specific hazard rates is by the use of life tables (also called survivorship tables).

It may readily be seen that during any selected period of time (t_i) we have a number of risks competing to affect an animal. There are risks of (i) “natural death,” (ii) death induced by a direct or indirect action of the test compound, and (iii) death due to such occurrences of interest of tumors (Hammond et al., 1978; Salsburg, 1980). And we are indeed interested in determining if (and when) the last two of these risks become significantly different than the “natural” risks (defined as what is seen to happen in the control group). Life-table methods enable us to make such determinations as the duration of survival (or time until tumors develop) and the probability of survival (or of developing a tumor) during any period of time.

We start by deciding the interval length (t_i) we wish to examine within the study. The information we gain becomes more exact as the interval is shortened. But as interval length is decreased, the number of intervals increases and calculations become more cumbersome and less indicative of time-related trends because random fluctuations become more apparent. For a 2-year or lifetime rodent study, an interval length of a month is commonly employed. Some life-table methods, such as the Kaplan–Meier, have each new event (such as a death) define the start of a new interval.

Having established the interval length, we can tabulate our data (Cutler and Ederer, 1958). We start by establishing the

following columns in each table (a separate table being established for each group of animals—i.e., by sex and dose level):

- a. The interval of time selected (t_i)
- b. The number of animals in the group that entered that interval of the study alive (l_i)
- c. The number of animals withdrawn from study during the interval (such as those taken for an interim sacrifice or that may have been killed by a technician error) (ω_i)
- d. The number of animals that died during the interval (d_i)
- e. The number of animals at risk during the interval, $l_i = l_{i-1} - \frac{1}{2} \omega_i$, or the number on study at the start of the interval minus one half of the number withdrawn during the interval
- f. The proportion of animals that died $= D_i = d_i / l_i$
- g. The cumulative probability of an animal surviving until the end of that interval of study, $P_i = 1 - D_i$, or one minus the number of animals that died during that interval divided by the number of animals at risk
- h. The number of animals dying until that interval (M_i)
- i. Animals found to have died during the interval (m_i)
- j. The probability of dying during the interval of the study $c_i = 1 - (M_i + m_i / l_i)$ or the total number of animals dead until that interval plus the animals discovered to have died during that interval divided by the number of animals at risk through the end of that interval
- k. The cumulative proportion surviving, p_i , equivalent to the cumulative product of the interval probabilities of survival (i.e., $p_i = p_1 \cdot p_2 \cdot p_3 \cdots p_x$)
- l. The cumulative probability of dying, C_i , equal to the cumulative product of the interval probabilities to that point (i.e., $C_i = c_1 \cdot c_2 \cdot c_3 \cdots c_x$)

With such tables established for each group in a study, we may now proceed to test the hypotheses that each of the treated groups has a significantly shorter duration of survival or that each of the treated groups died more quickly (note that plots of total animals dead and total animals surviving will give one an appreciation of the data, but can lead to no statistical conclusions).

There are a multiplicity of methods for testing significance in life tables, with (as is often the case) the power of the tests increasing as does the difficulty of computation (Cox, 1972; Tarone, 1975; Haseman, 1977; Salsburg, 1980).

We begin our method of statistical comparison of survival at any point in the study by determining the standard error of the K interval survival rate as (Garrett, 1947)

$$S_K = P_K \sqrt{\sum_{i=1}^K \left(\frac{D_i}{l'_x - d_x} \right)}$$

We may also determine the effective sample size (l_1) in accordance with

$$l_1 = \frac{P(1-P)}{S^2}$$

We may now compute the standard error of difference for any two groups (1 and 2) as

$$S_D = \sqrt{S_1^2 + S_2^2}$$

The difference in survival probabilities for the two groups is then calculated as

$$P_D = P_1 - P_2$$

We can then calculate a test statistic as

$$t' = \frac{P_D}{S_D}$$

This is then compared to the z distribution table. If $t' > z$ at the desired probability level, it is significant at that level. With increasing recognition of the effects of time (both as age and length of exposure to unmeasured background risks), life-table analysis has become a mainstay in chronic toxicology. An example is the reassessment of the ED₀₁ study (SOT ED₀₁ Task Force, 1981) which radically changed interpretation of the results and understanding of underlying methods when adjustment for time on study was made.

The increased importance and interest in the analysis of survival data have not been restricted to toxicology, but rather has encompassed all the life sciences. Those with further interest should consult Lee (1980) or Elandt-Johnson and Johnson (1980), both general in their approach to the subject.

30.9 META-ANALYSIS

Meta-analysis (meaning “analysis among”) is being used increasingly in biomedical research to try to obtain a qualitative or quantitative synthesis of the research literature on a particular issue. The technique is usually applied to the synthesis of several separate but comparable studies.

30.9.1 Selection of the Studies to Be Analyzed

The issue of study selection is perhaps the most troublesome issue for those doing meta-analysis. Several questions need to be addressed:

1. Should studies be limited to those which are published? It is well known that negative studies that

report little or no benefit from following a particular course of action are less likely to be published than are positive studies. Therefore, the published literature may be biased toward studies with positive results, and a synthesis of these studies would give a biased estimate of the impact of pursuing some courses of action. Unpublished studies, however, may be of lower quality than the published studies, and poor research methods often produce an underestimate of impact. Moreover, the unpublished studies may be difficult to discover.

2. Should studies be limited to those which appear in peer-reviewed publications? Peer review is considered the primary method for quality control in scientific publishing. Some investigators recommend that only those studies which are published in peer-reviewed publications be considered in meta-analysis. Although this may seem an attractive option, it might produce an even more highly biased selection of studies.
3. Should studies be limited those which meet additional quality control criteria? If investigators impose an additional set of criteria before including a study in meta-analysis, this may further improve the average quality of the studies used, but it introduces still greater concerns about selection bias. Moreover, different investigators might use different criteria for a “valid” study and therefore select a different group of studies for meta-analysis.
4. Should studies be limited to randomized controlled studies? This is a variant of the earlier question concerning quality control. At one time, rigid quality standards were more likely to be met by randomized controlled studies than by observational studies. Increasingly, however, observational methods have been used to evaluate certain kinds of effects, particularly those that are uncommon. A larger issue may well be that of combining data from studies performed in different laboratories and, even more so, using different strains of a single animal species.
5. Should studies be limited to those using identical methods? For practical purposes, this would mean using only separately published studies from the same lab in a limited time frame, for which the methods were the same for all and the similarity of methods was monitored. This criterion is very difficult to achieve.

30.9.2 Pooled (Quantitative) Analysis

Usually, the main purpose of meta-analysis is quantitative. The goal is to develop better overall estimates of the degree of benefit achieved by specific exposure and dosing techniques, based on the combining (pooling) of estimates

found in the existing studies of the interventions. This type of meta-analysis is sometimes called a pooled analysis (Gerbarg and Horwitz, 1988) because the analysts pool the observations of many studies and then calculate parameters such as risk ratios or odds ratios from the pooled data.

Because of the many decisions regarding inclusion or exclusion of studies, different meta-analyses might reach very different conclusions on the same topic. Even after the studies are chosen, there are many other methodological issues in choosing how to combine means and variances (e.g., what weighting methods should be used). Pooled analysis should report both relative risks and risk reductions as well as absolute risks and risk reductions (Sinclair and Bracken, 1994).

30.9.3 Methodological (Qualitative) Analysis

Sometimes the question to be answered is not how much toxicity is induced by the use of a particular exposure but whether there is any biologically significant toxicity at all. In this case, a qualitative meta-analysis may be done, in which the quality of the research is scored according to a list of objective criteria. The meta-analyst then examines the methodologically superior studies to determine whether or not the question of toxicity is answered consistently by them. This qualitative approach has been called methodological analysis (Gerbarg and Horwitz 1988) or quality scores analysis (Greenland 1994). In some cases, the methodologically strongest studies agree with one another and disagree with the weaker studies, which may or may not be consistent with one another.

30.10 BAYESIAN INFERENCE

It is useful to know the sensitivity and specificity of a test. Once a researcher decides to use a certain test, two important questions require answers: If the test results are positive, what is the probability that the researcher has the condition of interest? If the test results are negative, what is the probability that the patient does not have the disease? Bayes' theorem provides a way to answer these questions.

Bayes' theorem, which was first described centuries ago by the English clergyman after whom it is named, is one of the most imposing statistical formulas in the biomedical sciences (Lindley, 1971). Put in symbols more meaningful for researchers such as pathologists, the formula is as follows:

$$P(D|T+) = \frac{p(T+|D+)p(D+)}{[p(T+|D+)p(D+)] + [p(T+|D-)p(D-)]}$$

Where p denotes probability; D+ means that the animal has the effect in question; D- means that the animal does not

have the effect; T+ means that a certain diagnostic test for the effect is positive; T- means that the test is negative; and the vertical line (|) means "conditional upon" what immediately follows.

Most researchers, even those who can deal with sensitivity, specificity, and predictive values, throw in the towel when it comes to Bayes' theorem. This is odd, because a close look at the earlier equation reveals that Bayes' theorem is merely the formula for the positive predictive value (Box and Tiao, 1973).

The numerator of Bayes' theorem merely describes **cell a** (the true-positive results). The probability of being in cell a is equal to the prevalence times the sensitivity, where $p(D+)$ is the prevalence (the probability of being in the effected column) and where $p(T+|D+)$ is the sensitivity (the probability of being in the top row, *given the fact of being in the effected column*). The denominator of Bayes' theorem consists of two terms, the first of which once again describes **cell a** (the true-positive results) and the second of which describes **cell b** (the false-positive error rate, or $p(T+|D-)$), is multiplied by the prevalence of noneffected animals, or $p(D-)$. The true-positive results (a) divided by the true-positive plus false-positive results ($a+b$) give $a/(a+b)$, which is the positive predictive value.

In genetics, an even simpler-appearing formula for Bayes' theorem is sometimes used. The numerator is the same, but the denominator is mere $p(T+)$. This makes sense because the denominator in $a/(a+b)$ is equal to all of those who have positive test results, whether they are true-positive or false-positive results.

30.10.1 Bayes' Theorem and Evaluation of Safety Assessment Studies

In a population with a low prevalence of a particular toxicity, most of the positive results in a screening program for that lesion or effect would be falsely positive. Although this does not automatically invalidate a study or assessment program, it raises some concerns about cost-effectiveness, and these can be explored using Bayes' theorem (Racine et al., 1986).

A program employing an immunochemical stain-based test to screen tissues for a specific effect will be discussed as an example. This test uses small amounts of antibody tissues for a specific effect, and the presence of an immunologically bound stain is considered a positive result. If the sensitivity and specificity of the test and the prevalence of biochemical effect are known, Bayes' theorem can be used to predict what proportion of the tissues with positive test results will have true-positive results (actually be effected).

Example 30.3 shows how the calculations are made. If the test has a sensitivity of 96% and if the true prevalence is 1%, only 13.9% of tissues with a positive test result are predicted actually to be effected.

EXAMPLE 30.3**USE OF BAYES' THEOREM OR A 2×2 TABLE TO DETERMINE THE POSITIVE PREDICTIVE VALUE OF A HYPOTHETICAL TUBERCULIN SCREENING PROGRAM***Part 1* Beginning data

| | |
|---------------------------------------|------------|
| Sensitivity of immunological stain | 96% = 0.96 |
| False-negative error rate of the test | 4% = 0.04 |
| Specificity of the test | 94% = 0.94 |
| False-positive error rate of the test | 6% = 0.06 |
| Prevalence of effect in the tissues | 1% = 0.01 |

Part 2 Use of Bayes' theorem

$$\begin{aligned}
 p(D^+ | T^+) &= \frac{p(T^+ | D^+)p(D^+)}{[p(T^+ | D^+)p(D^+)] + [p(T^+ | D^-)p(D^-)]} \\
 &= \frac{(\text{sensitivity})(\text{prevalence})}{[(\text{sensitivity})(\text{prevalence})] + (\text{false} - \text{positive error rate})(1 - \text{prevalence})} \\
 &= \frac{(0.96)(0.01)}{[(0.96)(0.01)] + [(0.06)(0.99)]} = \frac{0.0096}{0.0096 + 0.0594} = \frac{0.0096}{0.0690} \\
 &= 0.139 = \mathbf{13.9\%}
 \end{aligned}$$

Part 3 Use of a 2×2 table, with numbers based on the assumption that 10 000 tissues are in the study

| | True Disease Status | | | | | |
|----------|---------------------|------------|-------------|------------|--------|------------|
| | Diseased | | Nondiseased | | Total | |
| | Number | Percentage | Number | Percentage | Number | Percentage |
| | <i>Test Result</i> | | | | | |
| Positive | 96 | 96 | 594 | 6 | 690 | 7 |
| Negative | 4 | 4 | 9306 | 94 | 9310 | 93 |
| Total | 100 | 100 | 9900 | 100 | 10000 | 100 |

Positive predictive value = 96/690 = 0.139 = **13.9%**.

Pathologists and toxicologists can quickly develop a table that lists different levels of test sensitivity, test specificity, and effect prevalence and shows how these levels affect the proportion of positive results that are likely to be true-positive results. Although this calculation is fairly straightforward and is extremely useful, it has seldom been used in the early stages of planning for large studies or safety assessment programs.

30.10.2 Bayes' Theorem and Individual Animal Evaluation

Suppose a pathologist is uncertain about an animal's cause of death and obtains a positive test result for a certain

pathology. Even if the pathologist knows the sensitivity and specificity of the test, that does not solve the problem, because to calculate the positive predictive value, it is necessary to know the prevalence of the particular tissue/effect that the test is designed to detect. The prevalence is thought of as the expected prevalence in the population from which the animal comes. The actual prevalence is usually not known, but often a reasonable estimate can be made.

Say, for example, a pathologist evaluates a male primate that was observed to have easy fatigability and has signs of kidney stones but has no other symptoms or signs of parathyroid disease on physical examination. The pathologist considers the probability of hyperparathyroidism and decides that it is now perhaps 2% (reflecting that in 100 such

primates, probably only two of them would have the disease). This probability is called the prior probability, reflecting the fact that it is estimated prior to the performance of laboratory tests and is based on the estimated prevalence of a particular pathology among primates with similar signs and symptoms. Although the pathologist believes that the probability of hyperparathyroidism is low, he or she considers the results of the serum calcium test to "rule" out the diagnosis. Somewhat to his or her surprise, the results of the test were positive, with an elevated level of 12.2 mg dL⁻¹. He or she could order more special tests or stains for parathyroid disease, but some test results might come back positive and some negative.

Under the circumstances, Bayes' theorem could be used to make a second estimate of probability, which is called the posterior probability, reflecting the fact that it is made after the test results are known. Calculation of the posterior probability is based on the sensitivity and specificity of the test that was performed, which in this case was the serum calcium test, and on the prior probability, which in this case was 2%. If the serum calcium test had a 90% sensitivity and a 95% specificity, that means it had a false-positive error rate of 5% (specificity plus the false-positive error rate equals 100%). When this information is used in the Bayes' equation, as shown in Example 30.4, the result is a posterior

probability of 27%. This means that the patient is now in a group of primates with a significant possibility of parathyroid disease. In Example 30.4, note that the result is the same (i.e., 27%) when a 2×2 table is used. This is true because, as discussed in the preceding text, the probability based on the Bayes' theorem is identical to the positive predictive value.

EXAMPLE 30.4

USE OF BAYES' THEOREM OR A 2×2 TABLE TO DETERMINE THE POSTERIOR PROBABILITY AND THE POSITIVE PREDICTIVE VALUE

Part 1 Beginning data

Sensitivity of first test 90%=0.90
Specificity of first test 95%=0.95
Prior probability of disease 2%=0.02

Part 2 Use of Bayes' theorem

$$\begin{aligned} p(D^+ | T^+) &= \frac{p(T^+ | D^+)p(D^+)}{[p(T^+ | D^+)p(D^+)] + [p(T^+ | D^-)p(D^-)]} \\ &= \frac{(0.90)(0.02)}{[(0.90)(0.02) + (0.05)(0.98)]} \\ &= \frac{0.018}{0.018 + 0.049} = \frac{0.018}{0.067} = 0.269 = \mathbf{27\%} \end{aligned}$$

Part 3 Use of a 2×2 table

| | True Disease Status | | | | | |
|-------------|---------------------|------------|-------------|------------|--------|------------|
| | Diseased | | Nondiseased | | Total | |
| | Number | Percentage | Number | Percentage | Number | Percentage |
| Test Result | | | | | | |
| Positive | 18 | 90 | 49 | 5 | 67 | 6.7 |
| Negative | 2 | 10 | 931 | 95 | 933 | 93.3 |
| Total | 20 | 100 | 980 | 100 | 1000 | 100.0 |

Positive predictive value = $18/67 = 0.269 = \mathbf{27\%}$.

In light of the 27% posterior probability, the pathologist decides to order a parathyroid hormone radioimmunoassay, even though this test is expensive. If the radioimmunoassay had a sensitivity of 95% and a specificity of 98% and the results turned out to be positive, the Bayes' theorem could again be used to calculate the probability of parathyroid disease. This time, however, the posterior probability for the first test (27%) would be used as the prior probability for the second test. The result of the calculation, as shown in Example 30.5, is a new probability of 94%. Thus, the primate in all probabilities did have hyperparathyroidism.

EXAMPLE 30.5

USE OF BAYES' THEOREM OR A 2×2 TABLE TO DETERMINE THE SECOND POSTERIOR PROBABILITY AND THE SECOND POSITIVE PREDICTIVE VALUE

Part 1 Beginning data

Sensitivity of the first test 95%=0.95
Specificity of the first test 98%=0.98
Prior probability of disease 27%=0.27

Part 2 Use of Bayes' theorem

$$\begin{aligned} p(D^+ | T^+) &= \frac{p(T^+ | D^+)p(D^+)}{[p(T^+ | D^+)p(D^+)] + [p(T^+ | D^-)p(D^-)]} \\ &= \frac{(0.95)(0.27)}{[(0.95)(0.27) + (0.02)(0.73)]} \\ &= \frac{0.257}{0.257 + 0.0146} = \frac{0.257}{0.272} = 0.9449^a = \mathbf{94\%} \end{aligned}$$

Part 3 Use of a 2×2 table

| | True Disease Status | | | | | |
|-------------|---------------------|------------|-------------|------------|--------|------------|
| | Diseased | | Nondiseased | | Total | |
| | Number | Percentage | Number | Percentage | Number | Percentage |
| Test Result | | | | | | |
| Positive | 256 | 95 | 15 | 2 | 271 | 27.1 |
| Negative | 13 | 5 | 716 | 98 | 729 | 72.9 |
| Total | 269 | 100 | 731 | 100 | 1000 | 100.0 |

Positive predictive value = $256/271 = 0.9446^a = \mathbf{94\%}$.

^aThe slight difference in the results for the two approaches is due to rounding errors. It is not important biologically.

Why did the posterior probability increase so much the second time? One reason was that the prior probability was considerably higher in the second calculation than in the first (27% vs. 2%), based on the fact that the first test yielded positive results. Another reason was that the specificity of the second test was quite high (98%), which markedly reduced the false-positive error rate and therefore increased the positive predictive value.

30.11 DATA ANALYSIS APPLICATIONS IN SAFETY ASSESSMENT STUDIES

Having reviewed basic principles and provided a set of methods for statistical handling of data, the remainder of this book will address the practical aspects and difficulties encountered in working safety assessment.

There are now common practices in the analysis of safety data, though they are not necessarily the best. These are discussed in the remainder of this chapter, which seeks to review statistical methods on a use-by-use basis and to provide a foundation for the selection of alternatives in specific situations. Some of the newer available methodologies (meta-analysis and Bayesian approaches) should be kept in mind, however.

30.11.1 Body and Organ Weights

Among the sets of data commonly collected in studies where animals are dosed with (or exposed to) a chemical are body weight and the weights of selected organs. In fact, body weight is frequently the most sensitive indication of an adverse effect. How to best analyze this and in what form to analyze the organ weight data (as absolute weights, weight changes, or percentages of body weight) have been the subject of a number of articles (Jackson, 1962; Weil, 1962; Weil and Gad, 1980).

Both absolute body weights and rates of body weight change (calculated as changes from a baseline measurement value which is traditionally the animal's weight immediately prior to the first dosing with or exposure to test material) are almost universally best analyzed by ANOVA, followed, if called for, by a *post hoc* test. Even if the groups were randomized properly at the beginning of a study (no group being significantly different in mean body weight from any other group and all animals in all groups within two SD of the overall mean body weight), there is an advantage to performing the computationally slightly more cumbersome (compared to absolute body weights) changes in body weight analysis. The advantage is an increase in sensitivity, because the adjustment of starting points (the setting of initial weights as a "zero" value) acts to reduce the amount of initial variability. In this case, Bartlett's test is performed first to ensure homogeneity of variance and the appropriate sequence of analysis follows.

With smaller sample sizes, the normality of the data becomes increasingly uncertain and nonparametric methods such as Kruskal–Wallis may be more appropriate (Zar, 1974).

The analysis of relative (to body weight) organ weights is a valuable tool for identifying possible target organs (Bickis, 1990; Lee and Lovell, 2009). How to perform this analysis is still a matter of some disagreement, however. Weil (1962) presented evidence that organ weight data expressed as percentages of body weight should be analyzed separately for each sex. Furthermore, since the conclusions from organ weight data of males differed so often from those of females, data from animals of each sex should be used in this measurement. Others (Boyd and Knight, 1963; Grubbs, 1969; Boyd, 1972; Weil, 1973) have discussed in detail other factors which influence organ weights and must be taken into account.

The two competing approaches to analyzing relative organ weights call for either:

1. Calculating organ weights as a percentage of total body weight (at the time of necropsy) and analyzing the results by ANOVA
2. Analyzing results by ANCOVA, with body weights as the covariates as discussed previously by the author (Weil and Gad, 1980)

A number of considerations should be kept in mind when these questions are addressed. First, one must keep a firm grasp on the difference between biological significance and statistical significance. In this particular case, we are especially interested in examining organ weights when an organ weight change is not proportional to changes in whole body weights. Second, we are now required to detect smaller and small changes while still retaining a similar sensitivity (i.e., the $p < 0.05$ level).

There are several devices to attain the desired increase in power. One is to use larger and larger sample sizes (number of animals) and the other is to utilize the most powerful test we can. However, the use of even currently employed numbers of animals is being vigorously questioned and the power of statistical tests must, therefore, now assume an increased importance in our considerations.

The biological rationale behind analyzing both absolute body weight and the organ weight to body weight ratio (this latter as opposed to a covariance analysis of organ weights) is that in the majority of cases, except for the brain, the organs of interest in the body change weight (except in extreme cases of obesity or starvation) in proportion to total body weight. We are particularly interested in detecting cases where this is not so. Analysis of actual data from several hundred studies (unpublished data) has shown no significant difference in rates of weight change of target organs (other than the brain) compared to total body weight for healthy animals in those species commonly used for repeated-dose studies (rats, mice, rabbits, and dogs). Furthermore, it should be noted that ANCOVA is of questionable validity in analyzing body weight and related organ weight changes, since a primary assumption is the independence of treatment—that the relationship of the two variables is the same for all treatments (Ridgemen, 1975). Plainly, in toxicology this is not true.

In cases where the differences between the error mean squares are much greater, the ratio of F ratios will diverge in precision from the result of the efficiency of covariance adjustment. These cases are where either sample sizes are much larger or where the differences between means themselves are much larger. This latter case is one which does not occur in the designs under discussion in any manner that would leave ANCOVA as a valid approach, because group

means start out being very similar and cannot diverge markedly unless there is a treatment effect. As we have discussed earlier, a treatment effect invalidates a prime underpinning assumption of ANCOVA.

30.11.2 Clinical Chemistry

A number of clinical chemistry parameters are commonly determined on the blood and urine collected from animals in chronic, subchronic, and, occasionally, acute toxicity studies. In the past (and still, in some places), the accepted practice has been to evaluate these data using univariate parametric methods (primarily *t*-tests and/or ANOVA). However, this can be shown to be not the best approach on a number of grounds.

First, such biochemical parameters are rarely independent of each other. Neither is our interest often focused on just one of the parameters. Rather, there are batteries of the parameters associated with toxic actions at particular target organs. For example, increases in creatinine phosphokinase (CPK), γ -hydroxybutyrate dehydrogenase (γ -HBDH), and lactate dehydrogenase (LDH), occurring together, are strongly indicative of myocardial damage. In such cases, we are not just interested in a significant increase in one of these, but in all three. Detailed coverage of the interpretation of such clinical laboratory tests can be found in other references (Martin et al., 1975; Harris, 1978; Loeb and Quimby, 1999; Gad and Rousseaux, 2013; Gad, 2015) or elsewhere in this text.

Similarly, the serum electrolytes (sodium, potassium, and calcium) interact with each other; a decrease in one is frequently tied, for instance, to an increase in one of the others. Furthermore, the nature of the data (in the case of some parameters), either because of the biological nature of the parameter or the way in which it is measured, is frequently either not normally distributed (particularly because of being markedly skewed) or not continuous in nature. This can be seen in some of the reference data for experimental animals in Mitruka and Rawnsley (1977) or Weil (1982) in, for example, creatinine, sodium, potassium, chloride, calcium, and blood.

30.11.3 Hematology

Much of what we said about clinical chemistry parameters is also true for the hematologic measurements made in toxicology studies. Which test to perform should be evaluated by use of a decision tree until one becomes confident as to the most appropriate methods. Keep in mind that sets of values and (in some cases) population distribution vary not only between species but also between the commonly used strains of species and that “control” or “standard” values will “drift” over the course of only a few years.

Again, the majority of these parameters are interrelated and highly dependent on the method used to determine them. RBC count, platelet counts, and mean corpuscular volume

(MCV) may be determined using a device such as a Coulter counter to take direct measurements, and the resulting data are usually stable for parametric methods. The hematocrit, however, may actually be a value calculated from the RBC and MCV values and, if so, is dependent on them. If the hematocrit is measured directly, instead of being calculated from the RBC and MCV, it may be compared by parametric methods.

Hemoglobin is directly measured and is an independent and continuous variable. However, and probably because at any one time a number of forms and conformations (oxyhemoglobin, deoxyhemoglobin, methemoglobin, etc.) of hemoglobin are actually present, the distribution seen is not typically a normal one, but rather may be a multimodal one. Here a nonparametric technique such as the Wilcoxon or multiple rank-sum is called for.

Consideration of the WBC and differential counts leads to another problem. The total WBC is, typically, a normal population amenable to parametric analysis, but differential counts are normally determined by counting, manually, one or more sets of 100 cells each. The resulting relative percentages of neutrophils are then reported as either percentages or are multiplied by the total WBC count with the resulting “count” being reported as the “absolute” differential WBC. Such data, particularly in the case of eosinophils (where the distribution does not approach normality), should usually be analyzed by nonparametric methods. It is widely believed that “relative” (%) differential data should not be reported because they are likely to be misleading.

Lastly, it should always be kept in mind that it is rare for a change in any single hematologic parameter to be meaningful. Rather, because these parameters are so interrelated, patterns of changes in parameters should be expected if a real effect is present, and analysis and interpretation of results should focus on such patterns of changes. Classification analysis techniques often provide the basis for a useful approach to such problems.

30.11.4 Histopathological Lesion Incidence

The last 30 years have seen increasing emphasis placed on histopathological examination of tissues collected from animals in subchronic and chronic toxicity studies. While it is not true that only those lesions which occur at a statistically significantly increased rate in treated/exposed animals are of concern (for there are the cases where a lesion may be of such a rare type that the occurrence of only one or a few such in treated animals “raises a flag”), it is true that, in most cases, a statistical evaluation is the prime method to determine if what we see in treated animals is significantly worse than what has been seen in control animals. And although cancer is not our only concern, this category of lesions is that of greatest interest (Gad and Rousseaux, 2013).

Typically, comparison of incidences of any one type of lesion between controls and treated animals is made using the

multiple 2×2 chi-square test or Fisher's exact test with a modification of the number of animals as the denominators. Too often, experimenters exclude from consideration all those animals (in both groups) which died prior to the first animals being found with a lesion at that site. The special case of carcinogenicity bioassays will be discussed in detail in the next chapter.

An option which should be kept in mind is that, frequently, a pathologist can not only identify a lesion as present but also grade those present as to severity. This represents a significant increase in the information content of the data which should not be given up by performing an analysis based only on the perceived quantal nature (present/absent) of the data. Quantal data, analyzed by chi-square or Fisher's exact tests, are a subset (the 2×2 case) of categorical or contingency table data. In this case it also becomes ranked (or "ordinal") data—the categories are naturally ordered (e.g., no effect < mild lesion < moderate lesion < severe lesion). This gives a $2 \times R$ table if there are only one treatment and one control group or an $N \times R$ ("multiway") table if there are three or more groups of animals.

The traditional method of analyzing multiple, cross-classified data has been to collapse the $N \times R$ contingency table over all but two of the variables and to follow this with the computation of some measure of association between these variables. For an N -dimensional table this results in $N(N-1)/2$ separate analyses. The result is crude, "giving away" information, and even (by inappropriate pooling of data) yielding a faulty understanding of the meaning of data. Though computationally more laborious, a multiway ($N \times R$ table) analysis should be utilized.

30.11.5 Carcinogenesis

In the experimental evaluation of substances for carcinogenesis based on experimental results in a nonhuman species at some relatively high-dose or exposure level, an attempt is made to predict the occurrence and level of tumorigenesis in humans at much lower levels. An entire chapter could be devoted to examining the assumptions involved in this undertaking and review

of the aspects of design and interpretation of animal carcinogenicity studies. Such is beyond the scope of this effort. The reader is referred to Gad (2005) for such an examination.

The single most important statistical consideration in the design of carcinogenicity bioassays in the past was based on the point of view that what was being observed and evaluated was a simple quantal response (cancer occurred or it didn't) and that a sufficient number of animals are needed to be used to have reasonable expectations of detecting such an effect. Though the single fact of whether or not the simple incidence of neoplastic tumors is increased due to an agent of concern is of interest, a much more complex model must now be considered. The time-to-tumor, patterns of tumor incidence, effects on survival rate, and age at first tumor all must now be included in a model.

The rationale behind this assumption is that though humans may be exposed at very low levels, detecting the resulting small increase (over background) in the incidence of tumors would require the use of an impractically large number of test animals per group. This point was illustrated by Table 30.1, where, for instance, while only 46 animals (per group) are needed to show a 10% increase over a zero background (i.e., a rarely occurring tumor type), 770 000 animals (per group) would be needed to detect a tenth of a percent increase above a five percent background. As we increase dose, however, the incidence of tumors (the response) will also increase until it reaches the point where a modest increase (say, 10%) over a reasonably small background level (say, 1%) could be detected using an acceptably small-sized group of test animals (in Table 30.8 we see that 51 animals would be needed for this example case). There are, however, at least two real limitations to the highest dose level. First, the test rodent population must have a sufficient survival rate after receiving a lifetime (or 2 years) of regular doses to allow for meaningful statistical analysis. Second, we really want the metabolism and mechanism of action of the chemical at the highest level tested to be the same as at the low levels where human exposure would

TABLE 30.8 Average Number of Animals Needed to Detect a Significant Increase in the Incidence of an Event (Tumors, Anomalies, etc.) Over the Background Incidence (Control) at Several Expected Incidence Levels Using Fisher's Exact Probability Test ($p = 0.05$)

| Background Incidence (%) | Expected Increase in Incidence (%) | | | | | |
|--------------------------|------------------------------------|-----------|--------|-------|-----|-----|
| | 0.01 | 0.1 | 1 | 3 | 5 | 10 |
| 0 | 46 000 000 ^a | 460 000 | 4 600 | 511 | 164 | 46 |
| 0.01 | 46 000 000 | 460 000 | 4 600 | 511 | 164 | 46 |
| 0.1 | 47 000 000 | 470 000 | 4 700 | 520 | 168 | 47 |
| 1 | 51 000 000 | 510 000 | 5 100 | 570 | 204 | 51 |
| 5 | 77 000 000 | 770 000 | 7 700 | 856 | 304 | 77 |
| 10 | 100 000 000 | 1 000 000 | 10 000 | 1 100 | 400 | 100 |
| 20 | 148 000 000 | 1 480 000 | 14 800 | 1 644 | 592 | 148 |
| 25 | 160 000 000 | 1 600 000 | 16 000 | 1 840 | 664 | 166 |

^a Number of animals needed in each group—controls as well as treated.

occur. Unfortunately, toxicologists usually must select the high-dose level based only on the information provided by a subchronic or range-finding study (usually 90 days in length), but selection of either too low or too high a dose will make the study invalid for detection of carcinogenicity and may seriously impair the use of the results for risk assessment.

There are several solutions to this problem. One of these has been the rather simplistic approach of the NTP bioassay program, which is to conduct a 3-month range-finding study with sufficient dose levels to establish a level which significantly (10%) decreases the rate of body weight gain. This dose is defined as the maximum tolerated dose (MTD) and is selected as the highest dose. Two other levels, generally one-half MTD and one-quarter MTD, are selected for testing as the intermediate- and low-dose levels. In many earlier NCI studies, only one other level was used.

The dose range-finding study is necessary in most cases, but the suppression of body weight gain is a scientifically questionable benchmark when dealing with establishment of safety factors. Physiological, pharmacologic, or metabolic markers generally serve as better indicators of systemic response than body weight. A series of well-defined acute and subchronic studies designed to determine the "chronicity factor" and to study onset of pathology can be more predictive for dose setting than body weight suppression.

Also, the NTP's MTD may well be at a level where the metabolic mechanisms for handling a compound at real-life exposure levels have been saturated or overwhelmed, bringing into play entirely artifactual metabolic and physiological mechanisms (Gehring and Blau, 1977). The regulatory response to questioning the appropriateness of the MTD as a high-dose level (Haseman, 1985) has been to acknowledge that occasionally an excessively high dose is selected, but to counter by saying that using lower doses would seriously decrease the sensitivity of detection.

REFERENCES

- Abramowitz M, Stegun IA. (1964) *Handbook of Mathematical Functions*. National Bureau of Standards, Washington, DC, pp. 925–964.
- Anderson E. (1960) A semigraphical method for the analysis of complex problems. *Technometrics*. 2:387–391.
- Anderson TW. (1971) *The Statistical Analysis of Time Series*. John Wiley & Sons, Inc., New York.
- Anderson S, Auquier A, Hauck WW, Oakes D, Vandaele W, Weisburg HL. (1980) *Statistical Methods for Comparative Studies*. John Wiley & Sons, Inc., New York.
- Anscombe FJ. (1973) Graphics in statistical analysis. *Am. Stat.* 27:17–21.
- Antonello JM, Clark RL, Heyse JF. (1993) Application of Tahey trend test procedures to assess developmental and reproductive toxicity. I. Measurement data. *Fundam. Appl. Toxicol.* 21:52–58.
- Armitage P. (1955) Tests for linear trends in proportions and frequencies. *Biometrics*. 11:375–386.
- Beyer WH. (1976) *Handbook of Tables for Probability and Statistics*. CRC Press, Boca Raton, FL.
- Bickis MG. (1990) Experimental design. In: *Handbook of In Vivo Toxicity Testing* (Arnold DL, Grice HC and Krewski DR, Eds.). Academic Press, San Diego, CA, pp. 113–166.
- Bliss CI. (1935) The calculation of the dosage-mortality curve. *Ann. Appl. Biol.* 22:134–167.
- Bloomfield P. (1976) *Fourier Analysis of Time Series: An Introduction*. John Wiley & Sons, Inc., New York.
- Box GEP, Tiao GC. (1973) *Bayesian Inference in Statistical Analysis*. Addison-Wesley, Reading, MA.
- Boyd EM. (1972) *Predictive Toxicometrics*. Williams & Wilkins, Baltimore, MD.
- Boyd EM, Knight LM. (1963) Postmortem shifts in the weight and water levels of body organs. *Toxicol. Appl. Pharmacol.* 5:119–128.
- Breslow N. (1984) Comparison of survival curves. In: *Cancer Clinical Trials: Methods and Practice* (Buse MF, Staguet MJ and Sylvester RF, Eds.). Oxford University Press, Oxford, pp. 381–406.
- Chambers JM, Cleveland WS, Kleiner B, Tukey PA. (1983) *Graphical Methods for Data Analysis*. Wadsworth, Belmont, CA.
- Chernoff H. (1973) The use of faces to represent points in K-dimensional space graphically. *J. Am. Stat. Assoc.* 68:361–368.
- Cleveland WS. (1985) *The Elements of Graphing Data*. Wadsworth Advanced Books, Monterey, CA.
- Cleveland WS. (1994) *The Elements of Graphing Data*, 2nd Ed. Wadsworth Advanced Books, Monterey, CA.
- Cleveland WS, McGill R. (1984) Graphical perception: theory, experimentation, and application to the development of graphical methods. *J. Am. Stat. Assoc.* 79:531–554.
- Cochran WF. (1954) Some models for strengthening the common χ^2 tests. *Biometrics*. 10:417–451.
- Cochran WG, Cox GM. (1975) *Experimental Designs*, 2nd Ed. John Wiley & Sons, Inc., New York.
- Conover JW, Inman RL. (1981) Rank transformation as a bridge between parametric and nonparametric statistics. *Am. Stat.* 35:124–129.
- Cox DR. (1972) Regression models and life-tables. *J. R. Stat. Soc.* 34B:187–220.
- Cox DR, Stuart A. (1955) Some quick tests for trend in location and dispersion. *Biometrics*. 42:80–95.
- Crowley J, Breslow N. (1984) Statistical analysis of survival data. *Annu. Rev. Public Health.* 5:385–411.
- Cutler SJ, Ederer F. (1958) Maximum utilization of the life table method in analyzing survival. *J. Chronic Dis.* 8:699–712.
- Davison ML. (1983) *Multidimensional Scaling*. John Wiley & Sons, Inc., New York.
- Diamond WJ. (1981) *Practical Experimental Designs*. Lifetime Learning Publications, Belmont, CA.
- Diem K, Lentner C. (1975) *Documenta Geigy Scientific Tables*. Geigy, New York, pp. 158–159.
- Draper NR, Smith H. (1998) *Applied Regression Analysis*, 3rd Ed. John Wiley & Sons, Inc., New York.
- Duncan DB. (1955) Multiple range and multiple F tests. *Biometrics*. 11:1–42.

- Dunnett CW. (1955) A multiple comparison procedure for comparing several treatments with a control. *J. Am. Stat. Assoc.* 50:1096–1121.
- Dunnett CW. (1964) New tables for multiple comparison with a control. *Biometrics.* 16:671–685.
- Dykstra RL, Robertson T. (1983) On testing monotone tendencies. *J. Am. Stat. Assoc.* 78:342–350.
- Elandt-Johnson RC, Johnson NL. (1980) *Survival Models and Data Analysis*. John Wiley & Sons, Inc., New York.
- Engelman L, Hartigan JA. (1969) Percentage points of a test for clusters. *J. Am. Stat. Assoc.* 64:1647–1648.
- Everitt BS, Hand DJ. (1981) *Finite Mixture Distributions*. Chapman & Hall, New York.
- Everitt B, Landou S, Leese M, Stahl D. (2011) *Cluster Analysis*, 5th Ed. John Wiley & Sons, Inc., New York.
- Federal Register. (1985) No. 50, Vol. 50. Washington, DC.
- Federer WT. (1955) *Experimental Design*. Macmillan, New York.
- Feinstein AR. (1979) Scientific standards vs. statistical associations and biological logic in the analysis of causation. *Clin. Pharmacol. Ther.* 25:481–492.
- Finney DK. (1977) *Probit Analysis*, 3rd Ed. Cambridge University Press, Cambridge, UK.
- Finney DJ, Latscha R, Bennet BM, Hsu P. (1963) *Tables for Testing Significance in a 2×2 Contingency Table*. Cambridge University Press, Cambridge, UK.
- Gad SC. (1984) Statistical analysis of behavioral toxicology data and studies. *Arch. Toxicol. Suppl.* 5:256–266.
- Gad SC. (2005) *Statistics and Experimental Design for Toxicologists*, 4th Ed. CRC Press, Boca Raton, FL.
- Gad SC. (2015) *Animal Models in Toxicology*, 3rd Ed. Informa, Boca Raton, FL.
- Gad SC, Rousseaux CG. (2013) Use and misuse of statistics in toxicologic pathology. In: *Handbook of Toxicologic Pathology* (Hascheck WM and Rousseaux CG, Eds.), 3rd Ed. Academic Press, San Diego, CA, pp. 327–418.
- Gad SC, Taulbee SM. (1996) *Handbook of Data Recording, Maintenance and Management for the Biomedical Sciences*. CRC Press, Boca Raton, FL.
- Gad SC, Reilly C, Siino KM, Gavigan FA. (1985) Thirteen cationic ionophores: neurobehavioral and membrane effects. *Drug Chem. Toxicol.* 8(6):451–468.
- Gallant AR. (1975) Nonlinear regression. *Am. Stat.* 29:73–81.
- Garrett HE. (1947) *Statistics in Psychology and Education*. Longmans, Green, New York, pp. 215–218.
- Gehring PJ, Blau GE. (1977) Mechanisms of carcinogenicity: dose response. *J. Environ. Pathol. Toxicol.* 1:163–179.
- Gerbarg ZB, Horwitz RI. (1988) Resolving conflicting clinical trials: guidelines for meta-analysis. *J. Clin. Epidemiol.* 41:503–509.
- Ghent AW. (1972) A method for exact testing of 2×2 , 2×3 , 3×3 and other contingency tables, employing binomial coefficients. *Am. Midl. Nat.* 88:15–27.
- Glass L. (1975) Classification of biological networks by their qualitative dynamics. *J. Theor. Biol.* 54:85–107.
- Gold HJ. (1977) *Mathematical Modeling of Biological System—An Introductory Guidebook*. John Wiley & Sons, Inc., New York.
- Gordon AD. (1981) *Classification*. Chapman & Hall, New York.
- Greenland S. (1994) Invited commentary: a critical look at some popular meta-analytic methods. *Am. J. Epidemiol.* 140:290–296.
- Grubbs FE. (1969) Procedure for detecting outlying observations in samples. *Technometrics.* 11:1–21.
- Hammond EC, Garfinkel L, Lew EA. (1978) Longevity, selective mortality, and competitive risks in relation to chemical carcinogenesis. *Environ. Res.* 16:153–173.
- Harris EK. (1978) Review of statistical methods of analysis of series of biochemical test results. *Ann. Biol. Clin.* 36:194–197.
- Harris RJ. (2013) *A Primer of Multivariate Statistics*, 3rd Ed. Academic Press, New York.
- Harter AL. (1960) Critical values for Duncan's new multiple range test. *Biometrics.* 16:671–685.
- Hartigan JA. (1983) Classification. In: *Encyclopedia of Statistical Sciences* (Katz S and Johnson NL, Eds.), Vol. 2. John Wiley & Sons, Inc., New York.
- Haseman JK. (1977) Response to use of statistics when examining life time studies in rodents to detect carcinogenicity. *J. Toxicol. Environ. Health.* 3:633–636.
- Haseman JK. (1985) Issues in carcinogenicity testing: dose selection. *Fundam. Appl. Toxicol.* 5:66–78.
- Hicks CR. (1982) *Fundamental Concepts in the Design of Experiments*. Holt, Rinehart, and Winston, New York.
- Hoaglin DC, Mosteller F. (2000) *Understanding Robust and Explanatory Data Analysis*, 2nd Ed. John Wiley & Sons, Inc., New York.
- Hollander M, Wolfe DA. (2013) *Nonparametric Statistical Methods*, 3rd Ed. John Wiley & Sons, Inc., New York, pp. 137–142.
- Jackson B. (1962) Statistical analysis of body weight data. *Toxicol. Appl. Pharmacol.* 4:432–443.
- Kotz S, Johnson NL. (1982) *Encyclopedia of Statistical Sciences*, Vol. 1. John Wiley & Sons, Inc., New York, pp. 61–69.
- Kowalski BR, Bender CF. (1972) Pattern recognition, a powerful approach to interpreting chemical data. *J. Am. Chem. Soc.* 94:5632–5639.
- Kraemer HC, Thiernann G. (1987) *How Many Subjects? Statistical Power Analysis in Research*, Vol. 27. Sage Publications, Newbury Park, CA.
- Lee ET. (1980) *Statistical Methods for Survival Data Analysis*. Lifetime Learning Publications, Belmont, CA.
- Lee PN, Lovell D. (2009) Statistics for toxicology. In: *General and Applied Toxicology* (Ballantyne B, Marrs T and Syversen T, Eds.), 3rd Ed, Vol. 2. Grove's Dictionaries, Inc., New York, pp. 675–690.
- Lindley SV. (1971) *Bayesian Statistics: A Review*. SIAM, Philadelphia, PA.
- Litchfield JT, Wilcoxon F. (1949) A simplified method of evaluating dose effect experiments. *J. Pharmacol. Exp. Ther.* 96:99–113.
- Loeb WF, Quimby FW. (1999) *The Clinical Chemistry of Laboratory Animals*, 2nd Ed. Taylor & Francis, Philadelphia, PA.
- Marriott FHC. (1991) *The Dictionary of Statistical Terms*. Longman Scientific & Technical, Essex, UK.

- Martin HF, Gudzinowicz BJ, Fanger H. (1975) *Normal Values in Clinical Chemistry*. Marcel Dekker, New York.
- Mendell NR, Finch SJ, Thode HC Jr. (1993) Where is the likelihood ratio test powerful for detecting two component normal mixtures? *Biometrics*. 49:907–915.
- Mitruka BM, Rawnsley HM. (1977) *Clinical Biochemical and Hematological Reference Values in Normal Animals*. Masson, New York.
- Montgomery DC, Smith EA. (1983) *Introduction to Linear Regression Analysis*. John Wiley & Sons, Inc., New York.
- Myers JL. (1972) *Fundamentals of Experimental Designs*. Allyn and Bacon, Boston, MA.
- Peto R, Pike MC. (1973) Conservatism of approximation $\Sigma(0-E)^2/E$ in the log rank test for survival data on tumour incidence data. *Biometrics*. 29:579–584.
- Peto R, Pike MC, Armitage P, Breslow NE, Cox DR, Howard SV, Kantel N, McPherson K, Peto J, Smith PG. (1977) Design and analysis of randomized clinical trials requiring prolonged observations of each patient, II analyses and examples. *Br. J. Cancer*. 35:1–39.
- Peto R, Pike M, Day N, Gray R, Lee P, Parish S, Peto J, Richards S, Wahrendorf J. (1980) *Guidelines for Simple, Sensitive Significance Tests for Carcinogenic Effects in Long-Term Animal Experiments, IARC Monographs on the Evaluation of the Carcinogenic Risk of Chemicals to Humans, Supplement 2, Long-Term and Short-Term Screening Assays for Carcinogens: A Critical Appraisal*. International Agency for Research in Cancer, Lyon, pp. 311–346.
- Pollard JH. (1977) *Numerical and Statistical Techniques*. Cambridge University Press, New York.
- Portier C, Hoel D. (1984) Type I error of trend tests in proportions and the design of cancer screens. *Comm. Stat. Theory Meth.* A13:1–14.
- Prentice RL. (1976) A generalization of the probit and logit methods for dose response curves. *Biometrics*. 32:761–768.
- Racine A, Grieve AP, Fluhler H. (1986) Bayesian methods in practice: experiences in the pharmaceutical industry. *Appl. Stat.* 35:93–150.
- Ridgemen WJ. (1975) *Experimentation in Biology*. John Wiley & Sons, Inc., New York, pp. 214–215.
- Roe FJ. (1989) Relevance for man of the effects of lactose, polzols and other carbohydrates on calcium metabolism. *Human Toxicol.* 8(2):87–98.
- Romesburg HC. (1984) *Cluster Analysis for Researchers*, Vol. 43. Lifetime Learning Publications, Belmont, CA, pp. 45–58.
- Salsburg D. (1980) The effects of life-time feeding studies on patterns of senile lesions in mice and rats. *Drug Chem. Toxicol.* 3:1–33.
- Schaper M, Thompson RD, Alarie Y. (1985) A method to classify airborne chemicals which alter the normal ventilatory response induced by CO₂. *Toxicol. Appl. Pharmacol.* 79:332–341.
- Scheffe H. (1959) *The Analysis of Variance*. John Wiley & Sons, Inc., New York.
- Schmid CF. (1983) *Statistical Graphics*. John Wiley & Sons, Inc., New York.
- Siegel S. (1956) *Nonparametric Statistics for the Behavioral Sciences*. McGraw-Hill, New York.
- Sinclair JC, Bracken MB. (1994) Clinically useful measures of effect in binary analyses of randomized trials. *J. Clin. Epidemiol.* 47:881–889.
- Snedecor GW, Cochran WG. (2014) *Statistical Methods*, 8th Ed. Iowa State University Press, Ames, IA.
- Sokal RR, Rohlf FJ. (1994) *Biometry*, 3rd Ed. W.H. Freeman, San Francisco, CA.
- Sokal RR, Rohlf FJ. (2011) *Biometry*, 4th Ed. W.H. Freeman, San Francisco, CA.
- SOT ED₀₁ Task Force. (1981) Reexamination of the ED₀₁ study-adjusting for time on study. *Fundam. Appl. Toxicol.* 1:8–123.
- Tarone RE. (1975) Tests for trend in life table analysis. *Biometrika*. 62:679–682.
- Tufte ER. (1983) *The Visual Display of Quantitative Information*. Graphics Press, Cheshire, CT.
- Tufte ER. (1990) *Envisioning Information*. Graphics Press, Cheshire, CT.
- Tufte ER. (1997) *Visual Explanations*. Graphics Press, Cheshire, CT.
- Tukey JW. (1977) *Exploratory Data Analysis*. Addison-Wesley Publishing Co., Reading, MA.
- Tukey JW, Ciminera JL, Heyes JF. (1985) Testing the statistical certainty of a response to increasing doses of a drug. *Biometrics*. 41:295–301.
- Velleman PF, Hoaglin DC. (1981) *Applications, Basics and Computing of Exploratory Data Analysis*. Duxbury Press, Boston, MA.
- Weil CS. (1962) Applications of methods of statistical analysis to efficient repeated-dose toxicological tests. I. General considerations and problems involved. Sex differences in rat liver and kidney weights. *Toxicol. Appl. Pharmacol.* 4:561–571.
- Weil CS. (1973) *Experimental design and interpretation of data from prolonged toxicity studies*. Proceedings of the Fifth International Congress on Pharmacology, Beacon Press, San Francisco, Vol. 2, pp. 4–12.
- Weil CS. (1982) Statistical analysis and normality of selected hematologic and clinical chemistry measurements used in toxicologic studies. *Arch. Toxicol. Suppl.* 5:237–253.
- Weil CS, Gad SC. (1980) Applications of methods of statistical analysis to efficient repeated-dose toxicologic tests. 2. Methods for analysis of body, liver and kidney weight data. *Toxicol. Appl. Pharmacol.* 52:214–226.
- Williams DA. (1971) A test for differences between treatment means when several dose levels are compared with a zero dose control. *Biomet.* 27:103–117.
- Williams DA. (1972) The comparison of several dose levels with a zero dose control. *Biomet.* 28:519–531.
- Young FW. (1985) Multidimensional scaling. In: *Encyclopedia of Statistical Sciences* (Katz S and Johnson NL, Eds.), Vol. 5. John Wiley & Sons, Inc., New York, pp. 649–659.
- Zar JH. (1974) *Biostatistical Analysis*. Prentice-Hall, Englewood Cliffs, NJ, p. 50.

COMBINATION PRODUCTS: DRUGS AND DEVICES

31.1 COMBINATION PRODUCTS

The last three decades have seen a vast increase in the number of new therapeutic products which are not purely drug, device, or biologic but rather a combination of two or more of these. Classical examples are implanted drug delivery systems (whose primary function is drug delivery) and drug-impregnated devices (in which drug delivery is an adjunct to the device function) (see Table 31.1 for a more extensive list of examples). Congress first acknowledged the need for specific regulation of such combination products in the 1990 Safe Medical Device Act, and subsequent regulations have clarified regulatory expectations (CDRH, 1994; Chapekar, 1996; FDA, 2006; Siegal, 2008; Crawford, 2015; Gad and Gad-McDonald, 2015).

31.1.1 Historical Background

The history of this category includes an ever-evolving (and expanding) variety of product types, dating at least from the perfection of the hypodermic needle (1855). There are many modern examples of implanted therapeutic delivery systems, such as the insulin pump (1980), pain drug pumps, and artificial pancreas implants (device/cell combination). One fundamental driving force for delivery systems has been the growth of new pharmaceutical products, especially since the dramatic expansion of drug research after 1945. That research has led to the synthesis and testing of millions of compounds for pharmacological and antimicrobial properties. Indeed, today much of that development is performed in automated computer-controlled systems, leading to an even greater acceleration of the process. The continued emergence of a stream of novel and more complex combination products has

blurred any distinguishing lines of regulatory authority and has complicated product designation and regulation. The issue of products combining a device and a drug, such as an asthma-metered dose inhaler, has received considerable scrutiny as more attention has been brought to bear of device leachables and extractables. But products combining a device and a biologic, such as organ replacement or assist devices, had received less attention.

Even less than drug and device combinations, device and biologic products—which include, among other things, cellular and tissue implants, infused or encapsulated cells, artificial and replacement organs, heart valves and pumps, and cardiac, neural, and neuromuscular stimulation devices—do not fit neatly into existing regulatory paradigms. For example, as part of the question of regulation, FDA must take into account the possibility of tissue contamination and other hazards involved in using animal-derived tissues.

What has come to pass is a settled regulatory process. The written guidelines are fixed and the day-to-day process seemingly settled (Merrill, 1994; March, 1998; Segal, 1999; Gopalaswamy and Gopalaswamy, 2008; Siegal, 2008). It starts with determination of a primary mode of action (PMOA) which governs which center of FDA will have the primary regulatory oversight. The FDA Office of Combination Products (OCP) reviews requests for such determinations (RFDs—requests for designations) and advises the parties concerned as to which centers will be the primary for product review. More recently, FDA (2006) has promulgated a guidance on the nonclinical safety evaluation of drug/drug combinations. Such combinations are now frequent, most commonly seeking approval by the 505(b)(2) route. If the combination is of two approved and marketed products, the primary safety concern is for international

TABLE 31.1 Examples of Existing Device/Drug Combination Products

| | | |
|---|---|--|
| Cardiac output catheter | Heparin | As device in the United Kingdom |
| Extracorporeal sets | Heparin | As device in the United Kingdom |
| Viscose/rayon dressings | Povidone iodine | As drug in the United Kingdom |
| Cardiovascular oxygenator | Heparin | As drug in UK defoamer reservoir |
| Paste bandages | Clioquinol, coal tar, calamine | As drug in the United Kingdom (if they have ichthammol ancillary action) |
| Medicated tulle dressings | Chlorhexidine | As drug in the United Kingdom |
| Antimicrobial drape | Iodophor | As device in the United Kingdom |
| Antiseptic wipes | Chlorhexidine, cetrimide, alcohol | As drug in the United Kingdom |
| Cardiovascular guidewires | Heparin | As device in Spain |
| Guidewires | Heparin | As device in Spain, Switzerland, and the United Kingdom |
| Antibiotic bone cement | Antibiotic (for example, Gentamicin sulfate), colistin sulfomethate, sodium, erythromycin | As drug (but soon to be regulated as device) |
| Extracorporeal cardiectomy reservoirs and filters | Heparin | As devices in Spain, Benelux, and Italy |
| Extracorporeal venous reservoirs and filters | Heparin | As devices in Spain, Benelux, and Italy |
| Bacteriostatic urological catheters | Silver | As devices in three Benelux countries |
| Antiseptic island dressing | Chlorhexidine digluconate | As device in Italy |
| Spermicidal condoms | Nonoxynol-9 | As device in Germany |
| Pacemaker lead with a porous tip (seulte) | Dexamethasone | As device |
| Pacemaker lead with protector mannitol capsule (sweet tip) | Mannitol | As device |
| Biomedicus centrifugal pump | Heparin | Not applicable |
| Peripheral vascular cannulae | Heparin | Not applicable |
| Surgical gauzes or nonwoven fabrics impregnated with iodophor | Iodophor | Not applicable |
| Surgical gauzes or nonwoven fabrics impregnated with—alginates and Clioquinol | Clioquinol (NaCa alginates, clauden powder) | As device in Germany |
| Vascular prosthesis | Collagen, albumen | As devices in the United Kingdom |

(pharmacokinetics and pharmacodynamics). As such, the primary nonclinical safety studies that must be conducted are systemic toxicity studies of the fixed combination itself with toxicokinetic sampling components. Genetic toxicity and safety pharmacology studies of the combination are usually not required.

31.1.2 Future Trends

Table 31.2 presents anticipated developments in the device combination product category up until 2020 which lead to new clinical products. Three types of developments are generally expected. First, additional products designed for implanted delivery of insulin and other drugs. These include implanted pumps, possibly intelligent devices with improved biosensors to monitor concentrations in body fluids and make dynamic adjustments in delivery rates. Also, there is the likely development of new polymeric timed-release devices which could improve the delivery of long-acting pharmaceuticals at optimized locations and rates.

Second, new developments in drug-impregnated devices are expected. Examples included new types of cardiac

implants with antithrombogenic and anti-infective drugs, as well as orthopedic implants with bacteriostatic coatings.

Finally, under development are new developments in drug delivery systems to simplify reliable use by unsophisticated patients in home settings, including the increasing elderly population. Examples included nasal and inhalation products.

Device regulation designation is by PMOA, which is generally straightforward but can become less clear as precedents accumulate and technology becomes more complex. Table 31.3 provides some examples of classifications.

Although both extracorporeal and peritoneal dialysis systems are regulated as devices, dialysate concentrate for use with the former is a device, but prepackaged dialysate for use with the latter is a drug. Sometimes consistency was elusive even when there was no combination but just a single product. For example, *in vitro* diagnostics for detecting antibodies to HIV are regulated as biologics when they are used for screening the blood supply but as medical devices when used for diagnostic or other screening purposes. When the FDA decides quickly and unequivocally on the regulatory status of a product, whether it was deemed a single product

TABLE 31.2 Likelihood New Combination Device Drug Technologies

| | |
|---|--|
| Biosensors | Biosensors (for glucose, implantation, and systemic infection markers), genetic diagnostics, laser diagnosis and treatment, minimally invasive devices |
| Blood vessel prosthetics | Genetic therapy, tissue-engineered vessels, nerves, and devices |
| Bone prosthetics/growth | Artificial organs, tissue-engineered devices, scaffolds with bone-growth-stimulating agents |
| Cardiac stimulation | Intelligent devices, microminiaturized devices |
| Cartilage prosthetics | Tissue-engineered device |
| Computer-aided clinical labs | Computer-aided diagnosis, networks of devices |
| Drug-impregnated devices | Device/drug/biological products |
| Endoscopy | Minimally invasive devices, telemedicine, virtual reality diagnostics |
| Genetics—cancer | Genetic diagnostics, genetic therapy |
| Hearing aids | Intelligent devices, microminiaturized devices, nonimplanted sensory aids |
| Heart pumps | Artificial organs |
| Heart valves | Artificial organs, tissue-engineered devices, device/drug/biological products |
| Home diagnostics | Home/self monitoring and diagnosis |
| Image contrast agents | Medical imaging |
| Imaging: functional, content | Medical imaging, minimally invasive devices, networks of devices |
| Implanted drug | Biosensors, device/drug/biological products, delivery systems, home/self-therapy, intelligent devices, robotic devices |
| Integrated patient medical info systems | Computer-aided diagnosis, networks of devices, telemedicine |
| Kidney prosthetics | Artificial organs, home/self-therapy, tissue-engineered devices |
| Laser surgery | Laser diagnosis and treatment |
| Liver prosthetics | Artificial organs, tissue-engineered devices |
| Min. invasive cardiology | Minimally invasive devices vascular surgery |
| Min. invasive neurosurgery | Minimally invasive devices |
| MRI | Greater resolution imaging |
| Nanotechnology | Microminiaturized devices |
| Nerve regeneration | Tissue-engineered devices |
| Neural stimulation | Artificial organs, electrical stimulation, intelligent devices |
| Neuromuscular stimulation | Electrical stimulation, home/self-therapy |
| Ocular prosthetics | Artificial organs, electrical stimulation, intelligent devices |
| Pancreas prosthetics | Artificial organs, tissue-engineered devices |
| Patient smart cards | Computer-aided diagnosis, networks of devices, telemedicine |
| PET imaging | Combined PET and CAT imaging |
| Robotic surgery | Microminiaturized devices, robotic devices |
| Skin prosthetics | Tissue-engineered devices |
| Telemedicine—home use | Home/self monitoring and diagnosis, telemedicine |
| Telemedicine—radiology | Telemedicine |

or was in combination with another product, there was relatively little opportunity for objection to the agency's decisions about how to regulate combination products and products whose status was uncertain. In the case of blood devices, the EU has affirmed this process (Anon, 2000).

In the Safe Medical Devices Act of 1990 (SMDA), Congress took these issues in hand and amended the Federal Food, Drug, and Cosmetic Act (FDCA) to make it easier for the FDA to regulate combination products in a rational fashion. The new provisions altered the substantive provisions of the FDCA only in minor respects. The main thrust of the new law was managerial, directing the FDA to make decisions about which center would have "primary jurisdiction" over a combination product, based on the agency's understanding of the PMOA of the product.

For these products, center jurisdiction turns on the PMOA. If the PMOA is that of a drug, then the Center for Drug Evaluation and Research (CDER) has primary jurisdiction; if it is that of a device, jurisdiction is with the Center for Devices and Radiological Health (CDRH); if that of a biological product, the Center for Biologics Evaluation and Research (CBER) has this jurisdiction. As the statute prescribed, the regulations go on to state that the center with primary jurisdiction may consult with other agency components.

Although neither the statute nor the regulations explain what "primary jurisdiction" means, it seems clear that the FDA intends it to mean that the center that has primary jurisdiction will review the combination product and ordinarily give it just one approval, that is, an NDA, PMA, or biologic license

TABLE 31.3 Classification Examples of Products

Combination products that have been classified as drugs:

- Prefilled syringes
- Patches for transdermal drug delivery
- Implants whose primary purpose is to release a drug
- Wound dressings whose primary purpose is to deliver a drug
- Dental products impregnated with a drug whose primary purpose is to deliver a drug
- Red blood cell processing solutions
- Contrast media
- Peritoneal dialysis solutions
- Alcohol swabs

Combination products that have been classified as devices:

- Drug-coated devices such as catheters, shunt sensors, or pacemaker leads
- Drug-impregnated devices
- Wound dressings and surgical barriers containing an antimicrobial agent
- Wound dressings whose primary purpose is to act as a barrier to pathogens
- Blood bags containing anticoagulant or preservation solutions
- Bone cement containing antibiotic and novel bone void fillers, for example, collagen matrix with bone morphogenic protein
- Injectable collagen
- Sodium hyaluronate nasal solution
- Urea breath test (accessory to device)
- Device for ex vivo photodynamic cell processing

Combinations of drugs and devices to which this policy does not apply and which must comply with both the food and drug regulations and the medical devices regulations:

- Kits (e.g., epidural tray containing drugs and devices, first aid kit containing drugs and devices)

Products for which neither set of regulations applies:

- Organ preservation solutions
- Minimally manipulated tissue

Source: Adapted from Gopaldaswamy and Gopaldaswamy (2008).

application (BLA) as appropriate. Section 3.4(b) makes it clear, however, that the FDA's designation of one agency component as having primary jurisdiction does not preclude, in appropriate cases, the requirement for separate application, for example, a 510(k) and a BLA. When separate applications are required, both can be reviewed by the lead center, but "exceptional" cases may involve a second application to be reviewed by a different center. To facilitate this, the agency published new delegations giving officials in each of the three centers the authority to clear devices and to approve devices, drugs, biologics, or any combination of two or more of them (FDA, 1991).

Contemporaneous with publication of the new regulations, the FDA made public three new intercenter agreements between CDRH and CBER, CDRH and CDER, and CDER and CBER. They describe the allocations of responsibility for numerous categories of specific products, both

combination and noncombination. According to the regulations, these intercenter agreements are not binding; they are intended to "provide useful guidance to the public" and, as a practical matter, to FDA staff as well.

The intercenter agreements are a treasure trove of information. In addition to explicit guidance about which center has the lead with respect to particular products and whether one center or two will work on particular issues, they contain information and hints about whether the FDA believes it can regulate certain products at all and, if so, how (Adams et al., 1997; Pilot and Waldemann, 1998).

The regulations and intercenter agreements, however, do not answer every question, and the regulations recognize a role for the sponsor in cases of uncertainty. When the identity of the center with primary jurisdiction is unclear or in dispute, or a sponsor believes its combination product is not covered by the intercenter agreements, a sponsor can request a designation from the FDA's product jurisdiction officer. A sponsor "should" file a request for designation with the product jurisdiction officer before submitting its application for marketing approval or an investigational notice. In practice, though, disputes or lack of clarity may not become evident until well into the review process, and it seems likely that the FDA would, if necessary, entertain requests for designation submitted at a later time.

Section 3.7(c) of the regulations lists the information to be included in the request, all of which must fit on 15 pages or less, including the identity of the sponsor, detailed information on the product, where the developmental work stands, the product's known modes of action and its PMOA, and, importantly, the sponsor's recommendation for which center should have primary jurisdiction and the reasons for the recommendation.

The FDA promises to check the request for designation for completeness within 5 working days of receipt and to issue a letter of designation within 60 days of receipt of a complete request. If the FDA does not meet the 60-day time limit, then the sponsor's recommendation for the appropriate lead center is honored.

The agency's letter of designation can be changed only with the sponsor's written consent or, if the sponsor does not consent, "to protect the public health or for other compelling reasons." A sponsor must be given prior notice of any proposed nonconsensual change and must be given an opportunity to object in writing and at a "timely" meeting with the product jurisdiction officer and appropriate center officials.

The CDRH is designated the center for major policy development and for the promulgation and interpretation of procedural regulations for medical devices under the Act. The CDRH regulates all medical devices inclusive of radiation-related device that are not assigned categorically or specifically to CDER. In addition, CDRH will independently

administer the following activities (references to “sections” are the provisions of the Act):

- A. Small business assistance programs under Section 10 of the amendments (see PL 94-295). Both CDER and CDRH will identify any unique problems relating to medical device regulation for small business.
- B. Registration and listing under Section 510 including some CDER-administered device applications. The CDER will receive printouts and other assistance, as requested.
- C. Color additives under Section 706, with review by CDER, as appropriate.
- D. Good Manufacturing Practices (GMPs) Advisory Committee. Under Section 520(f)(3), CDER will regularly receive notices of all meetings, with participation by CDER, as appropriate.
- E. Medical Device Reporting. The manufacturers, distributors, importers, and users of all devices, including those regulated by CDER, shall report to CDRH under Section 519 of the Act as required. The CDRH will provide monthly reports and special reports as needed to CDER for investigation and follow-up of those medical devices regulated by CDER.

Table 31.4 presents the primary product responsibilities of CDER and CBER.

TABLE 31.4 Product Class Review Responsibilities

| |
|---|
| Center for Drug Evaluation and Review |
| Natural products purified from plant or mineral sources |
| Products produced from solid tissue sources (excluding procoagulants, venoms, blood products, etc.) |
| Antibiotics, regardless of method of manufacture |
| Certain substances produced by fermentation |
| Disaccharidase inhibitors |
| HMG-CoA inhibitors |
| Synthetic chemicals |
| Traditional chemical synthesis |
| Synthesized mononuclear or polynuclear products including antisense chemicals |
| Hormone products |
| Center for Biologics Evaluation and Review |
| Vaccines, regardless of manufacturing method |
| <i>In vivo</i> diagnostic allergenic products |
| Human blood products |
| Protein, peptide, and/or carbohydrate products produced by cell culture (other than antibiotics and hormones) |
| Immunoglobulin products |
| Products containing intact cells or microorganisms |
| Proteins secreted into fluids by transgenic animals |
| Animal venoms |
| Synthetic allergens |
| Blood banking and infusion adjuncts |

31.1.2.1 Device Programs That CDER and CDRH Each Will Administer Both CDER and CDRH will administer and, as appropriate, enforce the following activities for medical devices assigned to their respective centers (references to “sections” are the provisions of the Act):

- A. Surveillance and compliance actions involving general controls violations, such as misbranded or adulterated devices under Sections 301, 501, and 502
- B. Warning letters, seizures, injunctions, and prosecutions under Sections 302, 303, and 304
- C. Civil penalties under Section 303(f) and administrative restraint under Section 304(g)
- D. Nonregulatory activities, such as educational programs directed at users, participation in voluntary standards organizations, etc
- E. Promulgation of performance standards and applications of special controls under Section 514
- F. Premarket notification, investigational device exemptions (IDE) including humanitarian exemptions; premarket approval; product development protocols; classification; device tracking; petitions for reclassification; postmarket surveillance under Sections 510(k), 513, 515, 519, 520(g) and (m), and 522; and the advisory committees necessary to support these activities
- G. Banned devices under Section 516
- H. FDA-requested and firm-initiated recalls whether under Section 518 or another authority and other Section 518 remedies such as recall orders
- I. Exemptions, variances, and applications of CGMP regulations under Section 520(f)
- J. Government-wide quality assurance program
- K. Requests for export approval under Sections 801(e) and 802

31.1.2.2 Coordination The centers will coordinate their activities in order to assure that manufacturers do not have to independently secure authorization to market their product from both centers unless this requirement is specified in Section VII.

31.1.2.3 Submissions Submissions should be made to the appropriate center, as specified herein, at the addresses provided in the following:

Address update:

- Food and Drug Administration
- Center for Drug Evaluation and Research (CDER)
- Central Document Room (CDR)

- 5901-B Ammendale Road
- Beltsville, MD 20705-1266

or

- Food and Drug Administration
- Center for Devices and Radiological Health
- Document Mail Center (HFZ-401)
- 9200 Corporate Blvd.
- Rockville, MD 20850

For submissions involving medical devices and/or drugs that are not clearly addressed in this agreement, sponsors are referred to the product jurisdiction regulations (21 CFR Part 3). These regulations have been promulgated to facilitate the determination of regulatory jurisdiction but do not exclude the possibility for a collaborative review between the centers.

31.1.2.4 Center Jurisdiction The following subsections provide details concerning status, market approval authority, special label/regulatory considerations, investigational options, and intercenter consultations for the categories of products specified. Section VII provides the general criteria that CDRH and CDER will apply in reaching decisions as to which center will regulate a product:

- A. 1. (a) Device with primary purpose of delivering or aiding in the delivery of a drug that is distributed without a drug (i.e., unfilled).

Examples:

- Devices that calculate drug dosages
- Drug delivery pump and/or catheter infusion pump for implantation of iontophoreses device
- Medical or surgical kit (e.g., tray) with reference in instructions for use with specific drug (e.g., local anesthetic)
- Nebulizer
- Small particle aerosol generator (SPAG) for administering drug to ventilated patient
- Splitter block for mixing nitrous oxide and oxygen
- Syringe, jet injector, storage, and dispensing equipment

Status: Device and drug, as separate entities.

Market approval authority: CDRH and CDER, respectively, unless the intended use of the two products, through labeling, creates a combination product.

Special label/regulatory considerations: The following specific procedures will apply depending on the status of the drug delivery device and drugs that will be delivered with the device:

- (i) It may be determined during the design or conduct of clinical trials for a new drug that it is not possible to develop adequate performance specifications data on those characteristics of the device that are required for the safe and effective use of the drug. If this is the case, then drug labeling cannot be written to contain information that makes it possible for the user to substitute a generic, marketed device for the device used during developments to use with the marketed drug. In these situations, CDER will be the lead center for regulation of the device under the device authorities.
- (ii) For a device intended for use with a category of drugs that are on the market, CDRH will be the lead center for regulation for the device under the device authorities. The effects of the device use on drug stability must be addressed in the device submission when relevant. An additional showing of clinical effectiveness of the drug when delivered by the specific device will generally not be required. The device and drug labeling must be mutually conforming with respect to indication, general mode of delivery (e.g., topical, I.V.), and drug dosage/schedule equivalents.
- (iii) For a drug delivery device and drug that are developed for marketing to be used together as a system, a lead center will be designated to be the contact point with the manufacturer(s). If a drug has been developed and marketed and the development and studying of device technology predominate, the PMOA will be deemed to be that of the device, and CDRH would have the lead. If a device has been developed and marketed and the development and studying of drug predominate, then, correspondingly, CDER would have the lead. If neither the drug nor the device is on the market, the lead center will be determined on a case-by-case basis.

Investigation options: IDE or IND, as appropriate.

Intercenter Consultation: CDER, when lead center, will consult with CDRH if CDER determines that a specific device is required as part of the NDA process. CDRH as lead center will consult with CDER. If the device is intended for use with a marketed drug and the device creates a significant change in the

intended use, mode of delivery (e.g., topical, I.V.), or dose/schedule of the drug.

- (b) The device with primary purpose of delivering or aiding in the delivery of a drug and distributed containing a drug (i.e., “prefilled delivery system”).

Examples

Nebulizer

Oxygen tank for therapy and OTC emergency use

Prefilled syringe

Transdermal patch

Status: Combination product

Market approval authority: CDER using drug authorities and device authorities, as necessary

Special label/regulatory considerations: None

Investigation options: IND

Intercenter consultations: Optional

2. Device incorporating a drug component with the combination product having the primary intended purpose of fulfilling a device function.

Examples:

Bone cement containing antimicrobial agent

Cardiac pacemaker lead with steroid-coated tip

Condom, diaphragm, or cervical cap with contraceptive or antimicrobial agent (including virucidal) agent

Dental device with fluoride

Dental wood wedge with hemostatic agent

Percutaneous cuff (e.g., for a catheter or orthopedic pin) coated/impregnated with antimicrobial agent

Skin closure or bandage with antimicrobial agent

Surgical or barrier drape with antimicrobial agent

Tissue graft with antimicrobial or other drug agent

Urinary and vascular catheter coated/impregnated with antimicrobial agent

Wound dressing with antimicrobial agent

Status: Combination product.

Market approval authority: CDRH using device authorities.

Special label/regulatory considerations: These products have a drug component that is present to augment the safety and/or efficacy of the device.

Investigation options: IDE.

Intercenter Consultation: Required if a drug or the chemical form of the drug has not been legally

marketed in the United States as a human drug for the intended effect.

3. Drug incorporating a device component with the combination product having the primary intended purpose of fulfilling a drug function.

Examples:

Skin-prep pads with antimicrobial agent

Surgical scrub brush with antimicrobial agent

Status: Combination product.

Market approval authority: CDER using drug authorities and, as necessary, device authorities.

Special label/regulatory considerations: Marketing of such a device requires a submission of an NDA with safety and efficacy data on the drug component, or it meets monograph specifications as generally recognized as safe (GRAS) and generally recognized as effective (GRAE). Drug requirements, for example, CGMPs, registration and listing, and experience reporting, apply to products.

Investigation options: IND.

Intercenter consultation: Optional.

4. (a) Device used in the production of a drug either to deliver directly to a patient or for the use in the producing medical facility (excluding use in a registered drug manufacturing facility).

Examples:

Oxygen concentrators (home or hospital)

Oxygen generator (chemical)

Ozone generator

Status: Device.

Market approval authority: CDER, applying both drug and device authorities.

Special label/regulatory consideration: May also require an NDA if the drug produced is a new drug. Device requirements, for example, CGMPs, registration and listing, and experience reporting, will apply to products.

Investigation options: IDA or NDA, as appropriate.

Intercenter consultation: Optional.

- (b) Drug/device combination product intended to process a drug into a finished package form.

Examples:

Device that uses drug concentrates to prepare large volume parenterals

Oxygen concentrator (hospital) output used to fill oxygen tanks for use within that medical facility

Status: Combination product.

Market approval authority: CDER, applying both drug and device authorities.

Special label/regulatory considerations: Respective drug and device requirements, for example, CGMPs, registration and listing, and experience reporting, will apply.

Investigation options: IDE or NDA, as appropriate.

Intercenter consultation: Optional but will be routinely obtained.

- B. 1. Device used concomitantly with a drug to directly activate or to augment drug effectiveness.

Examples:

Biliary lithotripter used in conjunction with dissolution agent

Cancer hyperthermia used in conjunction with chemotherapy

Current generator used in conjunction with an implanted silver electrode (drug) that produces silver ions for an antimicrobial purpose

Materials for blocking blood flow temporarily to restrict chemotherapy drug to the intended site of action

UV and/or laser activation of oxisoralen for psoriasis or cutaneous T-cell lymphoma

Status: Device and drug, as separate entities.

Market approval authority: CDRH and CDER, respectively.

Special label/regulatory considerations: The device and drug labeling must be mutually conforming with respect to indications, general mode of delivery (e.g., topical, I.V.), and drug dosage/schedule equivalence. A lead center will be designated to be the contact point with the manufacturer. If a drug has been developed and approved for another use and the development and studying of device technology predominate, then CDRH would have lead. If a device has been developed and marketed for another use and the development and studying of drug action predominate, then CDER would have lead. If neither the drug nor the device is on the market, the lead center will be determined on a case-by-case basis. If the labeling of the drug and device creates a combination product, as defined in the combination product regulations, then the designation of the lead center for both applications will be based upon a determination of the product's PMOA.

Investigation options: IDE or IND, as appropriate.

Intercenter consultations: Required.

2. Device kits labeled for use with drugs that include both device(s) and drug(s) as separate entities in one package with the overall primary intended purpose of the kit fulfilling a device function.

Examples:

Medical or surgical kit (e.g., tray) with drug component

Status: Combination product.

Market approval authority: CDRH, using device authorities, is responsible for the kit if the manufacturer is repackaging a market drug. Responsibility for overall packaging resides with CDRH. CDER will be consulted as necessary on the use of drug authorities for the repackaged drug component.

Special label/regulatory consideration: Device requirements, for example, CGMPs, registration and listing, and experience reporting, apply to kits. Device manufacturers must assure that manufacturing steps do not adversely affect drug components of the kit. If the manufacturing steps do affect the marketed drug (e.g., the kit is sterilized by irradiation), then ANDA or NDA would also be required with CDRH as lead center.

Investigation options: IDA or IND, as appropriate.

Intercenter consultation: Optional if ANDA or NDA is not required.

- C. Liquids, gases, or solids intended for use as devices (e.g., implanted or components, parts, or accessories to devices).

Examples:

Dye for tissues used in conjunction with laser surgery to enhance absorption of laser light in target tissue

Gas mixtures for pulmonary function testing devices

Gases used to provide "physical effects"

Hemodialysis fluids

Hemostatic devices and dressings

Injectable silicon, collagen, and Teflon

Liquids functioning through physical action applied to the body to cool or freeze tissues for therapeutic purposes

Liquids intended to inflate, flush, or moisten (lubricate) indwelling device (in or on the body)

Lubricants and lubricating jellies

Ophthalmic solutions for contact lenses

Organ/tissue transport and/or perfusion fluid with antimicrobial or other drug agent, that is, preservation solutions

Powders for lubricating surgical gloves

Sodium hyaluronate or hyaluronic acid for use as a surgical aid

Solution for use with dental "chemical drill"

Spray on dressings not containing a drug component

Status: Device.

Market approval authority: CDRH.

Special label/regulatory considerations: None.

Investigation options: IDE, best preceded by a presubmeeting.

Intercenter consultation: Required if the device has direct contact with the body and the drug or the chemical form of the drug has not been legally marketed as a human drug.

D. Products regulated as drugs.

Examples:

- Irrigation solutions
- Purified water or saline in prefilled nebulizers for use in inhalation therapy
- Skin protectants (intended for use on intact skin)
- Sun screens
- Topical/internal analgesic–antipyretic

Status: Drug.

Market approval authority: CDER.

Special label/regulatory considerations: None.

Investigation options: IND.

Intercenter consultations: Optional.

E. Ad hoc jurisdictional decisions.

| Examples | Status | Center |
|---|--------|--------|
| Motility marker constructed of radiopaque plastic | Device | CDRH |
| Brachytherapy capsules, needles, etc. that are radioactive and may be removed from the body after radiation therapy has been administered | Device | CDRH |
| Skin markers | Device | CDRH |

Status: Device or drug.

Market approval authority: CDRH or CDER as indicated.

Special label/regulatory considerations: None.

Investigation options: IDE or IND, as appropriate.

Intercenter consultation: Required to assure agreement on drug/device status.

31.1.2.5 General Criteria Affecting Drug/Device Determination The following represent the general criteria that will apply in making device/drug determinations.

A. Device criteria:

1. A liquid, powder, or other similar formulation intended only to serve as a component, part, or accessory to a device with a PMOA that is physical in nature will be regulated as a device by CDRH.
2. A product that has the physical attributes described in 201(h) (e.g., instrument, apparatus) of the Act and does not achieve its primary intended purpose through chemical action within or on the body, or

by being metabolized, will be regulated as a device by CDRH.

3. The phrase “within or on the body” as used in 201(h) of the Act does not include extracorporeal systems or the solutions used in conjunction with such equipment. Such equipment and solutions will be regulated as devices by CDRH.
4. An implant, including an injectable material, placed in the body for primarily a structural purpose even though such an implant may be absorbed or metabolized by the body after it has achieved its primary purpose will be regulated as a device by CDRH.
5. A device containing a drug substance as a component with the primary purpose of the combination being to fulfill a device function is a combination product and will be regulated as a device by CDRH.
6. A device (e.g., machine or equipment) marketed to the user, pharmacy, or licensed practitioner that produces a drug will be regulated as a device or combination product by CDER. This does not include equipment marketed to a registered drug manufacturer.
7. A device whose labeling or promotional materials make reference to a specific drug or generic class of drugs unless it is prefilled with a drug ordinarily remains a device regulated by CDRH. It may, however, also be subject to the combination products regulation.

B. Drug criteria

1. A liquid, powder, tablet, or other similar formulation that achieves its primary intended purpose through chemical action within or on the body, or by being metabolized, unless it meets one of the specified device criteria, will be regulated as a drug by CDER.
2. A device that serves as a container for a drug or a device that is a drug delivery system attached to the drug container where the drug is present in the container is a combination product that will be regulated as a drug by CDER.
3. A device containing a drug substance as a component with the primary purpose of the combination product being to fulfill a drug purpose is a combination product and will be regulated as a drug by CDER.
4. A drug whose labeling or promotional materials make reference to a specific device or generic class of devices ordinarily remains a drug regulated by CDER. It may, however, also be subject to the combination products regulation.
5. An imaging agent.

For the device component of combination product, required biocompatibility testing is specified in ISO 10993-1—in particular, the June 2016 FDA version (FDA, 2016).

If an approved device is being used to deliver a drug (such as the case for syringes and metered-dose inhalers), then generally only a leachables study is required to identify, quantitate, and assess the risk of any chemical entities migrating from the device into the drug (Gad and Gad-McDonald, 2015).

REFERENCES

- Adams DG, Cooper RM, Kahan JS. (1997) *Fundamentals of Law and Regulation: An In-Depth Look at Therapeutic Products*. FDLI, Washington, DC.
- Anon. (2000) EU Nations agree that directive will cover human blood devices. *Eur. Drug Device Rep.* 10(14):1–21.
- CDRH. (1994) *Premarket Notifications: 510 (K) Regulatory Requirements for Medical Devices*. HHS FDA 1990: 4158. Center for Devices and Radiological Health, Rockville, MD.
- Chapekar MS. (1996) Regulatory concerns in the development of biologic–biomaterials combination. *J. Biomed. Mater. Res.* 33:199–203.
- Crawford M. (2015) Compliance with Medical Device Regulations of Combination Products: CDRH Expectations. *AAPS News Magazine*, December 11, 2015:28–30.
- FDA. (1991) *Assignment of Agency Component for Review of Premarket Applications*, Finale Rule, 56 Fed. Reg. 58,754 (November 21, 1991).
- FDA. (2006) *Nonclinical Safety Evaluation of Drug or Biologic Combinations*. Center for Drug Evaluation and Research (CDER), Rockville, MD.
- FDA. (2016) *Use of International Standard ISO-10993, “Biological Evaluation of Medical Devices Part 1: Evaluation and Testing”*. CDRH, June 16, 2016.
- Gad SC, Gad-McDonald SE. (2015) *Biomaterials, Medical Devices and Combination Products: Biocompatibility Testing and Safety Assessment*. CRC Press, Boca Raton, FL.
- Gopalaswamy S, Gopalaswamy V. (2008) *Combination Products: Regulatory Challenges and Successful Product Development*. CRC Press, Boca Raton, FL.
- March E. (1998) Combination products: who’s regulating what and how—Part I. *Pharm. Eng.*, July/August, 18(4): 34–38.
- Merrill R. (1994) Regulation of drugs and devices: an evolution. *Health Aff.* 13:47–69.
- Pilot L, Waldemann D. (1998) Food and Drug Modernization Act of 1997: medical device provisions. *Food Drug Law J.* 53:267–295.
- Segal SA. (1999) Device and biologic combination products—understanding the evolving regulation. *Med. Device Diagn. Ind.*, January 1999, 180–184.
- Siegel EB. (2008) *Development and Approval of Combination Products: a Regulatory Perspective*. John Wiley & Sons, Inc, Hoboken, NJ.

QUALIFICATION OF IMPURITIES, DEGRADANTS, RESIDUAL SOLVENTS, METALS, AND LEACHABLES IN PHARMACEUTICALS

Impurities (from either materials intended to be part of the process of making a drug or formulating it), degradants (unintentionally formed by unintended reactions after the drug substance or product is produced and prone to increase in quantity over time due to instability of the product under conditions of storage), metals and elements, and residual solvents (purposely added to the synthesis product to facilitate synthesis, formulation, or dosage form production—always liquids of some degree of volatility) can become part of a drug product or substance in multiple ways. Additionally, drugs which are in containers or delivered/distributed (such as a metered-dose inhaler (MDI)) may also have substances from such a container “leach” into the drug product itself (with the safety of such leachables requiring evaluation). But the extent of their presence is now strictly governed by a series of ICH, FDA, and EMA guidelines (see CFR, 2005 and FDA, 2005). These ICH guidelines call for these materials to be present at levels no greater than in product specifications and (because they serve no functional purpose in the drug) are both to be kept to a practical minimum and must be qualified for safety at the highest specification levels under the assumption of maximum potential patient use of the drug (and, therefore, exposure to the unintended substances). Contaminants, coming from the vessels and machinery used to manufacture a drug but not intended to be present, and leachables/extractants are materials which transfer to a drug product from packaging or delivery system. Undesired chemical substances which end up in pharmaceutical products may have a range of sources, but all of these must be evaluated for potential risks to patients. Depending on the source of the undesired substance, it is categorized and regulated somewhat differently. Categories of such substances include impurities, degradants, residual solvents, contaminants, leachables, and extractables. Each of these will be considered in this chapter.

Biologics and biotechnology products have both similar and unique process impurity issues. The process needs to ensure that there are no residual cellular components in the biologic product. Biotechnology processing needs to avoid causing structural deformities to the protein. In all cases, the process must be scrutinized closely. Checking for impurities at various steps throughout the manufacturing phase may help to pinpoint where the impurities are produced.

32.1 IMPURITIES

The ICH *Guidance for Industry, Q3A Impurities in New Drug Substances*, published in February 2003 and finalized in October of 2006 (ICH, 2006a), is intended to “provide guidance for registration applications on the content and qualification of impurities in new drug substances produced by chemical syntheses and not previously registered in a region or member state.” A new drug substance is not the final marketed product but the active ingredient used in the marketed product. Impurities in new drug substances are addressed from both a chemistry and safety perspective (Ball et al., 2007; FDA, 2010).

The guidance is not intended to apply to new drug substances during the clinical research stage of development (though such drugs in development must have consideration of meeting, these requirements at the time of marketing approval and FDA frequently takes the view that such limitations must be met by a drug product entering clinical development) and addresses safety concerns associated with such substances during development. Nor does it cover natural product or biological process-produced drugs or extraneous contaminants that should not occur in new drug substances and are more appropriately addressed as good

TABLE 32.1 Thresholds for Action on Impurities in a Drug Product

| Maximum Daily Dose (g day ⁻¹) ^a | Reporting Threshold (%) ^{b,c} | Identification Threshold ^c | Qualification Threshold ^c |
|--|--|---|---|
| ≤2 | 0.05 | 0.10% or 1.0 mg day ⁻¹ intake (whichever is lower) | 0.15% or 1.0 mg day ⁻¹ intake (whichever is lower) |
| >2 | 0.03 | 0.05% | 0.05% |

^aThe amount of drug substance administered per day.^bHigher reporting thresholds should be scientifically justified.^cLower thresholds can be appropriate if the impurity is unusually toxic.

manufacturing practice (GMP) issues. The guidance further describes the circumstances in which impurities need to be reported, identified, and qualified.

The rationale for the reporting and control, identification, and qualification of impurities is discussed in the guidance. Organic impurities need to be summarized based on the actual and potential impurities most likely to arise during the synthesis, purification, and storage of a new drug substance. This discussion can be limited to those impurities that might reasonably be expected based on knowledge of the chemical reactions and conditions involved.

Studies conducted to characterize the structure of impurities present in a new drug substance at a level greater than the identification threshold (Table 32.1) should be described, and any impurity from any batch or degradation product from stability studies should be identified. If identification of an impurity or degradant is not feasible, a summary of the laboratory studies demonstrating the unsuccessful effort should be included in the application. If an impurity is pharmacologically or toxicologically active, identification of the compound should be conducted even if the impurity level is below the identification threshold.

The guidance also states that “qualification is the process of acquiring and evaluating data that establishes the biological safety of an individual impurity or a given impurity profile at the level(s) specified. The applicant should provide a rationale for establishing impurity acceptance criteria that includes safety considerations. The level of any impurity that is present in a new drug substance that has been adequately tested in safety and/or clinical studies would be considered qualified. Impurities that are also significant metabolites present in animal and/or human studies are generally considered qualified. A level of a qualified impurity higher than that present in a new drug substance can also be justified based on an analysis of the actual amount of impurity administered in previous relevant safety studies. If data are unavailable to qualify the proposed acceptance criterion of an impurity, safety studies to obtain such data can be appropriate when the usual qualification thresholds are exceeded.”

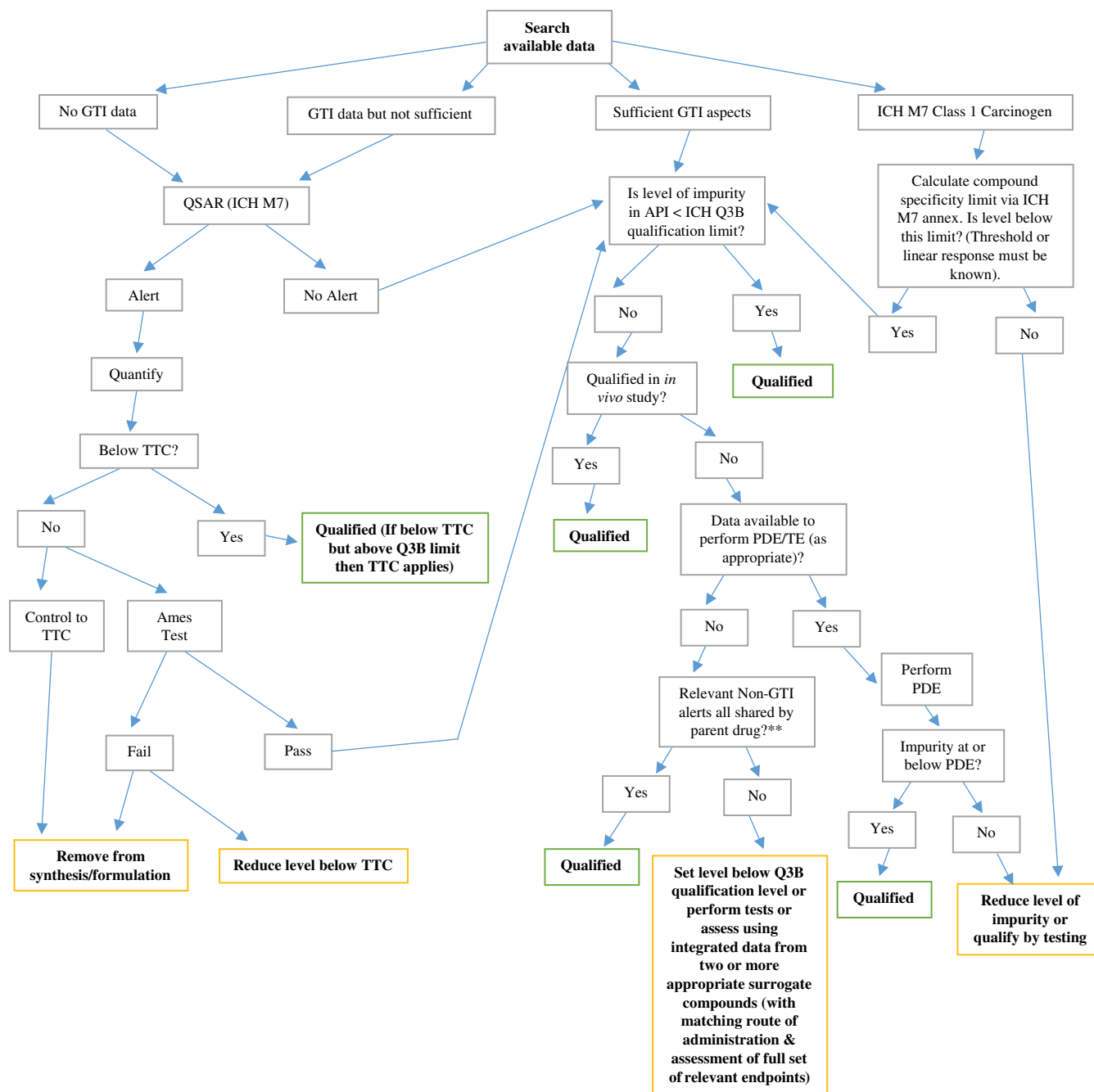
Q3B(R) describes considerations for the qualification of impurities when thresholds are exceeded. If the level of impurity cannot be decreased below the threshold, or if adequate data is not available in the scientific literature to justify safety,

then additional safety testing should be considered. The studies considered appropriate to qualify an impurity will depend on a number of factors, including the patient population, daily dose, and route and duration of administration. Toxicology studies are discussed briefly later in this chapter and in more detail in other chapters in this volume. Such studies can be conducted on the new drug substance containing the impurities to be controlled, although studies using isolated impurities can sometimes be appropriate.

ICH Q3A states that “safety assessment studies to qualify an impurity should compare the new drug substance containing a representative amount of the new impurity with previously qualified material. Safety assessment studies using a sample of the isolated impurity can also be considered.” The latter is especially important to consider for genetic toxicology studies and the importance of testing the isolated impurity is discussed in more detail at the end of this chapter.

Therefore, according to the guidance, if the maximum daily dose of the drug is less than 2 g day⁻¹ and the impurity intake is more than 0.15% or 1.0 mg day⁻¹, the qualification threshold has been reached, meaning safety studies will need to be performed. Lower thresholds can be appropriate if the impurity is unusually toxic. In addition, the impurity will need to be reported and identified. These studies include general and genetic toxicology studies, and possibly other specific toxicology end points, as appropriate. Discussion of specific toxicity testing with the relevant FDA division is recommended.

If considered desirable, a minimum screen (e.g., genotoxic potential) should be conducted. A study to detect point mutations and one to detect chromosomal aberrations, both *in vitro*, are considered an appropriate minimum screen. However, for genotoxic impurities (GTIs), the concern has evolved to being focused on mutagenic impurities. Indeed, ICH M7 (2014b) provides several approaches to establishing that impurities are either not GTIs or that if they are, they occur at acceptable levels (most commonly, below the appropriate threshold of toxicological concern (TTC)). If an impurity (in the broadest sense, including leachables, solvents, and all other subsets) is identified (“alerted”) to be mutagenic by literature or a QSAR determination, and is above the relevant TTC limit, there are a number of further steps (of increasing cost) that can be taken (Lee, 2015) (Figure 32.1). Among these may be *in vitro* or *in vivo* testing



*On rare occasion Q3A

Impurity does **not have to share all the alerts with parent

FIGURE 32.1 Hierarchical approach to qualification (Q3B)*.

to determine if there is a relevant risk to patients (see Table 32.2) (Teasdale, 2010).

Qualification studies for impurities are essentially bridging studies. If general toxicity studies are desirable, one or more studies should be designed to allow comparison of unqualified to qualified material. The study duration should be based on available relevant information and performed in the species most likely to maximize the potential to detect

the toxicity of a degradation product. On a case-by-case basis, single-dose studies can be appropriate, especially for single-dose drugs. In general, a minimum duration of 14 days and a maximum duration of 90 days would be considered appropriate.

The genetic toxicology studies can include a minimum screen (a study to detect point mutations and one to detect chromosome aberrations, both *in vitro*). The general toxicology

TABLE 32.2 Tests to Investigate the *In Vivo* Relevance of *In Vitro* Mutagens (Positive Bacterial Mutagenicity)

| <i>In Vitro</i> Test | Mechanistic Data to Justify Choice of Test as Fit for Purpose |
|--|--|
| Transgenic mutation assays | For any bacterial mutagenicity positive. Justify selection of assay tissue/organ |
| Pig-a assay (blood) | For directly acting mutagens (bacterial mutagenicity positive without S9) ^a |
| Micronucleus test (blood or bone marrow) | For directly acting mutagens (bacterial mutagenicity positive without S9) and compounds known to be clastogenic ^a |
| Rat liver UDS test | In particular for bacterial mutagenicity positive with S9 only Responsible liver metabolite known To be generated in test species used To induce bulky adducts |
| Comet assay | Justification needed (chemical class specific mode of action to form alkaline labile sites or single-strand breaks as preceding DNA damage that can potentially lead to mutations) |
| Others | Justify selection of assay tissue/organ With convincing justification |

Source: Adapted from Lee (2015).

Note: UDS, unscheduled DNA synthesis.

^aFor indirectly acting mutagens (requiring metabolic activation), justification is needed for sufficient exposure to metabolites.

studies should include one or more studies designed to allow comparison of unqualified to qualified material. The study duration should be based on available relevant information and performed in the species most likely to maximize the potential to detect the toxicity of an impurity. On a case-by-case basis, single-dose studies can be appropriate, especially for single-dose drugs. In general, a minimum duration of 14 days and a maximum duration of 90 days would be considered appropriate.

Inorganic impurities are normally detected and quantified using pharmacopeial or other appropriate procedures. The need for inclusion or exclusion of inorganic impurities in a new drug substance specification should be discussed. Acceptance criteria should be based on pharmacopeia standards or known safety data. The control of residues of the solvents used in the manufacturing process for a new drug substance should be discussed and presented according to ICH Q3C.

A registration application should include documented evidence that the analytical procedures are validated and suitable for the detection and quantification of impurities (ICH Q3A(R2), 2006a; ICH Q3B(R2), 2006b; ICH Q6A, 1999b; ICH Q6B, 1999a). Organic impurity levels can be measured by a variety of techniques, including those that compare an analytical response for an impurity to that of an appropriate reference standard or to the response of the new drug substance itself. Differences in the analytical procedures used during development and those proposed for the commercial product should be discussed in the registration application.

TABLE 32.3 Threshold for Degradation Products in New Drug Products

| Maximum Daily Dose ^a | Threshold ^{b,c} |
|----------------------------------|--|
| <i>Reporting thresholds</i> | |
| ≤1 g | 0.1% |
| >1 g | 0.05% |
| <i>Identification thresholds</i> | |
| <1 mg | 1.0% or 5 µg TDI, whichever is lower |
| 1–10 mg | 0.5% or 20 µg TDI, whichever is lower |
| >10 mg to 2 g | 0.2% or 2 mg TDI, whichever is lower |
| >2 g | 0.10% |
| <i>Qualification thresholds</i> | |
| <10 mg | 1.0% or 50 µg TDI, whichever is lower |
| 10–100 mg | 0.5% or 200 µg TDI, whichever is lower |
| >100 mg to 2 g | 0.2% or 3 mg TDI, whichever is lower |
| >2 g | 0.15% |

^aThe amount of drug substance administered per day.

^bThresholds for degradation products are expressed either as a percentage of the drug substance or as total daily intake (TDI) of the degradation product. Lower thresholds can be appropriate if the degradation product is unusually toxic.

^cHigher thresholds should be scientifically justified.

Analytical results should be provided in an application for all batches of a new drug substance used for clinical, safety, and stability testing, as well as for batches representative of the proposed commercial process. The application should also contain a table that links the specific new drug substance batch to each safety study and each clinical study in which the new drug substance has been used. Any impurity at a level greater than the reporting threshold (Table 32.1) and total impurities observed in these batches of the new drug substance should be reported with the analytical procedures indicated. Table 32.3 is an illustration of reporting impurity results for identification and qualification in an application.

The guidance also states that when analytical procedures change, results provided in the application should be linked to the procedure used, with appropriate validation information provided, including representative chromatograms of representative batches. The applicant should ensure that complete impurity profiles (e.g., chromatograms) of individual batches are available, if requested.

The ICH Q3A guidance also states that the specification for a new drug substance should include a list of impurities. Individual impurities with specific acceptance criteria included in the specification for a new drug substance are referred to as specified impurities. Specified impurities can be identified or unidentified. A rationale for the inclusion or exclusion of impurities in a specification should be presented.

“Acceptance criteria should be set no higher than the level that can be justified by safety data and should be consistent with the level achievable by the manufacturing process and the analytical capability. Where there is no safety concern, impurity acceptance criteria should be based on data generated on batches of a new drug substance manufactured by the

TABLE 32.4 Illustration of Reporting Degradation Product Results for Identification and Qualification in an Application

| Raw Result (%) | Reported Result (%) | Total Daily Intake (TPI) of the Degradation Product | Action | |
|---------------------------------------|------------------------|--|--------------------------|-------------------------|
| | | | Identification Threshold | Qualification Threshold |
| 50 mg maximum daily dose ^a | | | | |
| 0.04 | Not reported | 20 | None | None |
| 0.2143 | 0.2 | 100 | None | None |
| 0.349 | 0.3" | 150 | Yes | None" |
| 0.550 | 0.6" | 300 | Yes | Yes" |
| 1.9 g maximum daily dose ^b | | | | |
| 0.049 | Not reported | 1 | None | None |
| 0.079 | 0.08 | 2 | None | None |
| 0.183 | 0.18" | 3 | Yes | None ^{c,d} |
| 0.192 | 0.19" | 4 | Yes | Yes ^c |

^aReporting threshold 0.1%; TDI rounded result in-micrograms; identification threshold 0.2%; qualification threshold, TDI equivalent to 0.4%.

^bReporting threshold 0.05%; TDI rounded result in milligrams; identification threshold 2 mg TDI (equivalent to 0.11%); qualification threshold 3 mg TDI (equivalent to 0.16%).

^cAfter identification, if the response factor is determined to differ significantly from the original assumptions, it can be appropriate to remeasure the actual amount of the degradation product present and reevaluate against the qualification threshold.

^dAlthough the reported result of 0.18% exceeds the calculated threshold value of 0.16%, in this case the action is acceptable since the TDI (when rounded) does not exceed 3 mg. Chromatograms with peaks labeled (or equivalent data if other analytical procedures are used) from representative batches, including chromatograms from analytical procedure validation studies and from long-term and accelerated stability studies, should be provided. The applicant should ensure that complete degradation product profiles (e.g., chromatograms) of individual batches are available, if requested.

proposed commercial process, allowing sufficient latitude to deal with normal manufacturing and analytical variation and the stability characteristics of the new drug substance. Although normal manufacturing variations are expected, significant variation in batch-to-batch impurity levels can indicate that the manufacturing process of the new drug substance is not adequately controlled and validated (ICH Q6A)."

ICH Q3B(R2) *Impurities in New Drug Products* was published in June 2006 and is intended to provide guidance for registration applications on the content and qualification of impurities in new drug products produced from chemically synthesized new drug substances not previously registered in a region or member state. A new drug product is a finished dosage form, for example, a tablet, capsule, or solution, that contains a drug substance, generally, but not necessarily, in association with one or more other ingredients. The Q3B(R2) guidance (ICH, 2006b) complements the ICH guidance Q3A *Impurities in New Drug Substances*, which should be consulted for basic principles along with ICH Q3C(R5) *Impurities: Residual Solvents* (ICH, 2011), and ICH Q3D *Guideline for Elemental Impurities* (ICH, 2014a) when appropriate.

Q3A addresses only those impurities in new drug products classified as degradation products of the drug substance or reaction products of the drug substance with an excipient and/or immediate container closure system (collectively referred to as *degradation products*). Generally, impurities present in a new drug substance need not be monitored or specified in new drug product unless they are also degradation products. This guidance does not address impurities arising from excipients present in a new drug product or extracted or leached from the container closure system. This guidance also does not apply to new drug products used

during the clinical research stages of development. It also does not cover the same types of products as in 3QA(R): biological/biotechnological, peptides, oligonucleotides, radiopharmaceuticals, fermentation products and associated semisynthetic products, herbal products, and crude products of animal or plant origin. Also excluded from this guidance are extraneous contaminants that should not occur in new drug products and are more appropriately addressed as GMP issues, polymorphic forms, and enantiomeric impurities.

Qualification of an impurity for a new drug product has similar concerns as Q3A. The main differences are the reporting, identification, and qualification thresholds (Table 32.3). The thresholds are basically higher than they were in Q3A; however, there are more categories for dosages. If the qualification thresholds given in Table 32.4 are exceeded and data are unavailable to qualify the proposed acceptance criterion of a degradation product, additional studies to obtain such data may be appropriate.

US FDA (CDER) *Guidance for Industry, NDAs: Impurities in Drug Substances* (FDA, 2000) was published in February 2000. The guidance refers applicants to ICH Q3A *Impurities in New Drug Substances* when seeking guidance on identification, qualification, and reporting of impurities in drug substances that are not considered new drug substances. Q3A was developed by the ICH to provide guidance on the information that should be provided in a new drug application (NDA) in support of impurities in new drug substances that are produced by chemical syntheses. The FDA believes that such guidance on identification, qualification, and reporting of impurities should also be considered when evaluating impurities in drug substances produced by chemical syntheses that are not considered new drug substances. ICH

Q3A defines a new drug substance (also referred to as a new molecular entity or new chemical entity) as a designated therapeutic moiety that has not been previously registered in a region or member state. The definition also states that a new drug substance may be a complex, a simple ester, or a salt of a previously approved drug substance.

32.2 RESIDUAL SOLVENTS

ICH Q3C is intended to provide guidance for recommending acceptable amounts for residual solvents in pharmaceuticals for the safety of the patient. The guidance recommends use of less toxic solvents and describes levels considered to be toxicologically acceptable for some residual solvents. A complete list of the solvents included in this guidance is provided in a companion document entitled *ICH Q3C-Tables and List* which can be found at the ICH or FDA website, but may not be complete due to ongoing qualifications effort. The list is not exhaustive, and other solvents may be used and later added to the list.

Residual solvents in pharmaceuticals are defined here as organic volatile chemicals that are used or produced in the manufacture of drug substances or excipients or in the preparation of drug products. The solvents are not completely removed by practical manufacturing techniques. Appropriate selection of the solvent for the synthesis of drug substance may enhance the yield or determine characteristics such as crystal form, purity, and solubility. Therefore, the solvent may sometimes be a critical parameter in the synthetic process. This guidance does not address solvents deliberately used as excipients nor does it address solvates. However, the content of solvents in such products should be evaluated and justified.

As there are no therapeutic benefits from residual solvents, all residual solvents should be removed to the extent possible to meet product specifications, GMPs, or other quality-based requirements. Drug products should contain no higher levels of residual solvents than can be supported by safety data. Some solvents that are known to cause unacceptable toxicities (carcinogens), such as benzene and carbon tetrachloride (Class 1, see table 1 in ICH, 2014a), should be avoided in the production of drug substances, excipients, or drug products unless their use can be strongly justified in a risk–benefit assessment. Some solvents associated with less severe toxicity (nongenotoxic animal carcinogens or possible causative agents of other irreversible toxicity such as neurotoxicity or teratogenicity), such as acetonitrile and chlorobenzene (Class 2), should be limited in order to protect patients from potential adverse effects. Ideally, less toxic solvents, such as acetic acid and acetone (Class 3), should be used where practical.

This guidance does not apply to potential new drug substances, excipients, or drug products used during the clinical research stages of development nor does it apply to previously existing marketed drug products.

The guidance applies to all dosage forms and routes of administration. Higher levels of residual solvents may be acceptable in certain cases such as short-term (30 days or

less) or topical application. Justification for these levels should be made on a case-by-case basis and discussed with the appropriate FDA division.

The limits of residual solvents may include a value for the permitted daily exposure (PDE), which is the maximum acceptable intake per day of residual solvent in pharmaceutical products. These limits vary depending on the class.

For solvents where quantities are limited to set values in pharmaceutical products because of their inherent toxicity, Class 2 (table 2 of the guidance list) should be consulted. PDEs are given to the nearest 0.1 mg day⁻¹, and concentrations are given to the nearest 10 ppm.

For solvents with low toxic potential, solvents in Class 3 (table 3) may be regarded as less toxic and of lower risk to human health. Class 3 includes no solvent known as a human health hazard at levels normally accepted in pharmaceuticals. However, there are no long-term toxicity or carcinogenicity studies for many of the solvents in Class 3. Available data indicate that they are less toxic in acute or short-term studies and negative in genotoxicity studies. It is considered that amounts of these residual solvents of 50 mg day⁻¹ or less (corresponding to 5000 ppm or 0.5% under Option 1) would be acceptable without justification. Higher amounts may also be acceptable provided they are realistic in relation to manufacturing capability and GMPs.

For solvents for which no adequate toxicological data were found, the solvents listed (table 4 in ICH, 2014a) may also be of interest to manufacturers of excipients, drug substances, or drug products. However, no adequate toxicological data on which to base a PDE were found. Manufacturers should supply justification for residual levels of these solvents in pharmaceutical products.

The Gaylor–Kodell method of risk assessment (Gaylor and Kodell, 1980) is appropriate for Class 1 carcinogenic solvents. Only in cases where reliable carcinogenicity data are available should extrapolation by the use of mathematical models be applied to setting exposure limits. Exposure limits for Class 1 solvents could be determined with the use of a large safety factor (i.e., 10 000–100 000) with respect to the NOEL. Detection and quantitation of these solvents should be by state-of-the-art analytical techniques.

Acceptable exposure levels in this guidance for Class 2 solvents were established by calculation of PDE values according to the procedures for setting exposure limits in pharmaceuticals (*Pharmacopeial Forum*, November–December 1989) and the method adopted by IPCS for Assessing Human Health Risk of Chemicals (EHC 170, WHO, 1994). These methods are similar to those used by the US EPA (IRIS) and the US FDA (Red Book) and others. The method is outlined here to give a better understanding of the origin of the PDE values. It is not necessary to perform these calculations in order to use the PDE values tabulated in Section 4 of this document:

$$\text{PDE} = \frac{\text{NOEL} \times \text{weight adjustment}}{F1 \times F2 \times F3 \times F4 \times F5} \quad (32.1)$$

PDE is derived from the NOEL or the LOEL in the most relevant animal study as follows:

The PDE is derived preferably from a NOEL. If no NOEL is obtained, the LOEL may be used. Modifying factors proposed here, for relating the data to humans, are the same kind of *uncertainty factors* used in EHC (EHC 170, WHO, Geneva, 1994) and *modifying factors* or *safety factors* in *Pharmacoepial Forum*. The assumption of 100% systemic exposure is used in all calculations regardless of route of administration.

The modifying factors are as follows:

F1 = A factor to account for extrapolation between species

F1 = 5 for extrapolation from rats to humans

F1 = 12 for extrapolation from mice to humans

F1 = 2 for extrapolation from dogs to humans

F1 = 2.5 for extrapolation from rabbits to humans

F1 = 3 for extrapolation from monkeys to humans

F1 = 10 for extrapolation from other animals to humans

F1 takes into account the comparative surface area:body weight ratios for the species concerned and for man. Surface area (S) is calculated as

$$S = kM^{0.67} \quad (32.2)$$

in which M = body mass and the constant k have been taken to be 10. The body weights used in the equation are 50 kg for an adult and 10 kg for a child.

F2 = A factor of 10 to account for variability between individuals

A factor of 10 is generally given for all organic solvents, and 10 is used consistently in this guidance:

F3 = A variable factor to account for toxicity studies of short-term exposure

F3 = 1 for studies that last at least one half-lifetime (1 year for rodents or rabbits and 7 years for cats, dogs, and monkeys)

F3 = 1 for reproductive studies in which the whole period of organogenesis is covered

F3 = 2 for a 6-month study in rodents or a 3.5-year study in nonrodents

F3 = 5 for a 3-month study in rodents or a 2-year study in nonrodents

F3 = 10 for studies of a shorter duration.

In all cases, the higher factor has been used for study durations between the time points (e.g., a factor of 2 for a 9-month rodent study).

F4 = A factor that may be applied in cases of severe toxicity (e.g., nongenotoxic carcinogenicity, neurotoxicity, or

teratogenicity). In studies of reproductive toxicity, the following factors are used:

F4 = 1 for fetal toxicity associated with maternal toxicity

F4 = 5 for fetal toxicity without maternal toxicity

F4 = 5 for a teratogenic effect with maternal toxicity

F4 = 10 for a teratogenic effect without maternal toxicity

F5 = A variable factor that may be applied if the NOEL was not established

When only a LOEL is available, a factor of up to 10 could be used depending on the severity of the toxicity.

The weight adjustment assumes an arbitrary adult human body weight for either sex of 50 kg. This relatively low weight provides an additional safety factor against the standard weights of 60 or 70 kg that are often used in this type of calculation. It is recognized that some adult patients weigh less than 50 kg; these patients are considered to be accommodated by the built-in safety factors used to determine a PDE. If the solvent was present in a formulation specifically intended for pediatric use, an adjustment for a lower body weight would be appropriate.

As an example of the application of this equation, consider a toxicity study of acetonitrile in mice that is summarized in *Pharmeuropa*, vol. 9, no. 1, Supplement, April 1997, p. S24. The NOEL is calculated to be $50.7 \text{ mg kg}^{-1} \text{ day}^{-1}$. The PDE for acetonitrile in this study is calculated as follows:

$$\text{PDE} = \frac{50.7 \text{ mg kg}^{-1} \text{ day}^{-1} \times 50 \text{ kg}}{12 \times 10 \times 5 \times 1 \times 1} = 4.22 \text{ mg day}^{-1}$$

In this example:

F1 = 12 to account for the extrapolation from mice to humans

F2 = 10 to account for differences between individual humans

F3 = 5 because the duration of the study was only 13 weeks

F4 = 1 because no severe toxicity was encountered

F5 = 1 because the NOEL was determined

32.3 EXTRACTABLES AND LEACHABLES

Leachables are chemical entities, either organic or inorganic, that migrate from pharmaceutical container closure (or delivery) system components into a drug product formulation (Osterberg, 2005a, b). Since patients can be exposed to leachables during normal use of a drug product, leachables are of potential safety concern. Extractables are compounds that are forced out of container closure system materials and

components under laboratory experimental conditions. All extractables from a given pharmaceutical container closure system and its components are, therefore, potential leachables in a drug product incorporating the same container closure system components. Regulatory concern regarding leachables and extractables is directly related to the potential for contamination and/or interaction of the drug product formulation with the container closure system, with the greatest concern focused on orally inhaled and nasal drug products (OINDP), which include MDIs, dry powder inhalers (DPIs), inhalation solutions, suspensions and sprays, and nasal sprays (Colton, 2007; Norwood, 2007; Norwood et al., 2007; Bestwick and Colton, 2009). Controlled extraction studies are an extremely important part of the pharmaceutical development process for OINDP and should be performed on critical components as identified by the manufacturer and regulatory authority. As stated in the PQRI L&E recommendations: "A controlled extraction study is a laboratory investigation into the qualitative and quantitative nature of extractables profiles of critical components of an OINDP container closure system. The purpose of a controlled extraction study is to systemically and rationally identify and quantify potential leachables, that is, extractables, to the extent practicable, and within certain defined analytical threshold parameters."

Controlled extraction studies result in extractables profiles of OINDP components (FDA, 1998). Extractables profiles contain information which allows the identification, to the extent possible, and quantitation of individual extractables from a given component and therefore an early indication of potential leachables of concern. Controlled extraction studies generally establish a basis for the development and validation of routine quality control methods for drug product leachables and, finally, allow for the correlation of extractables and leachables profiles. Although information on component composition from suppliers is very useful, helping to inform component selection and guide controlled extraction studies, such knowledge does not provide a complete extractables profile and therefore does not alleviate the requirement for controlled extraction studies no matter how "complete" the information might appear to be.

It is, therefore, critical that controlled extraction studies be performed properly and thoroughly. Specific expectations for "proper and thorough" controlled extraction studies will ultimately depend on the nature of the OINDP being developed. However, the PQRI L&E Working Group was able to establish some general best practice recommendations for OINDP controlled extraction studies based on the data that the group generated by conducting its own controlled extraction studies on the elastomer and polymer test articles. These recommendations are:

- Controlled extraction studies should employ vigorous extraction with multiple solvents of varying polarity.
- Controlled extraction studies should incorporate multiple extraction techniques.

- Controlled extraction studies should include careful sample preparation based on knowledge of analytical techniques used.
- Controlled extraction studies should employ multiple analytical techniques.
- Controlled extraction studies should include a defined and systematic process for identification of individual extractables.
- Controlled extraction study "definitive" extraction techniques and methods should be optimized.
- During the controlled extraction studies, sponsors should revisit supplier information describing component information.
- Controlled extraction studies should be guided by Analytical Evaluation Thresholds (AET) that are based on an accepted safety concern threshold.
- Qualitative and quantitative extractables profiles should be discussed with and reviewed by toxicologists so that any potential safety concerns regarding individual extractables, that is, potential leachables, are identified early in the development process.
- Polynuclear aromatics (PNAs), *N*-nitrosamines, and 2-mercaptobenzothiazole (MBT) are "special case" compounds, requiring evaluation by specific analytical techniques and technology defined thresholds.

The characterization and control of leachables and extractables represent possibly the most significant challenge facing a pharmaceutical development team responsible for the development, registration, and manufacture of an OINDP. Indeed, detecting, identifying, and quantifying organic leachables are formidable tasks. In contrast to drug substance or excipient-related impurities, organic leachables can represent a diversity of chemical structures and compound classes and are potentially present at widely varying concentrations in any particular OINDP. Additionally, the information available to a pharmaceutical development team on container closure system component composition and processing, which is provided by the component supplier, is often incomplete. In some cases, the supplier may provide no information. Thus, when an extractables study is first undertaken, the development team may only have a limited idea of what to look for and what extraction techniques and analytical methods to use for identification and assessment of potential leachables.

32.4 RESIDUAL METALS AND ELEMENTS

In early 2008, the EMEA promulgated a standard for metals as impurities in pharmaceuticals (EMA, 2008). They organized metals of concern into categories, as presented in Table 32.5. USP followed with their own version of a guidance (USP, 2008), and, subsequently, an ICH guidance

TABLE 32.5 Class Exposure and Concentration Limits for Individual Metal Catalysts and Metal Reagents

| Classification | Oral Exposure | | Parenteral Exposure | | Inhalation Exposure ^a |
|--------------------------------------|--------------------------------|---------------------|--------------------------------|---------------------|----------------------------------|
| | PDE ($\mu\text{g day}^{-1}$) | Concentration (ppm) | PDE ($\mu\text{g day}^{-1}$) | Concentration (ppm) | PDE (ng day^{-1}) |
| Class 1A: Pt, Pd | 100 | 10 | 10 | 1 | Pt: 70 ^a |
| Class 1B: Ir, Rh, Ru, Os | 100 ^b | 10 ^b | 10 ^b | 1 ^b | |
| Class 1C: Mo, Ni, Cr, V | 250 | 25 | 25 | 2.5 | Ni: 100 |
| Metals of significant safety concern | | | | | Cr (VI): 10 |
| Class 2: Cu, Mn | 2 500 | 250 | 250 | 25 | |
| Metals with low safety concern | | | | | |
| Class 3: Fe, Zn | 13 000 | 1300 | 1300 | 130 | |
| Metals with minimal safety concern | | | | | |

^aPt as hexachloroplatinic acid.^bSubclass limit: the total amount of listed metals should not exceed the indicated limit.

(Q3D, 2013) was also promulgated, referring to metals and elements by which was meant boron.

If synthetic processes of pharmaceutical substances are known or suspected to lead to the presence of metal residues due to the use of a specific metal catalyst or metal reagent, a concentration limit and validated test for residues of each specific metal should be set. All concentration limits should be realistic in relation to analytical precision, manufacturing capability, and reasonable variation in the manufacturing process. Since the use of metal catalysts or metal reagents during synthesis is restricted to validated and controlled chemical reactions, limitation of their residues in pharmaceutical substances itself will normally be sufficient. A limit for a metal residue in the pharmaceutical substance may however be replaced by a limit for that metal residue in the final medicinal product, as described in the following text.

For pharmaceutical products administered via the oral, parenteral, or inhalation route of administration, two options are available when setting a concentration limit for a metal residue:

Option 1: For each metal, the concentration limit in parts per million (ppm) as stated in Table 1 can be used. The concentration limits in Table 1 have been calculated using Equation 32.3 in the following by assuming a daily dose of 10 g of the drug product:

$$\text{Concentration ppm} = \frac{\text{PDE } \mu\text{g day}^{-1}}{\text{daily dose g day}^{-1}} \quad (32.3)$$

If all pharmaceutical substances in a drug product meet the Option 1 concentration limit for all metals potentially present, then all these substances may be used in any proportion in the drug product as long as the daily dose of the drug product does not exceed 10 g day⁻¹. When the daily dose of the drug product is greater than 10 g day⁻¹, Option 2 should be applied:

Option 2a: The PDE in terms of microgram per day as stated in Table 1 can be used together with the actual daily dose of a pharmaceutical substance in the drug

product to calculate the concentration of residual metal allowed in that pharmaceutical substance.

Option 2b: Alternatively, it is not considered necessary for each pharmaceutical substance to comply with the limits given in Option 1 or the calculated limits using Option 2a.

The PDE in terms of microgram per day as stated in Table 1 can also be used with the known maximum daily dose of the drug product to determine the concentration of a metal residue originating from any of the pharmaceutical substances in the drug product (not the substance). This approach is considered acceptable provided that it has been demonstrated that the metal residue has been reduced to the practical minimum in every substance. This approach implies that the maximum levels of a metal in certain substances may be higher than the Option 1 or Option 2a limit, but that this should then be compensated by lower maximum levels in the other substances.

For pharmaceutical products applied via other routes of administration, the concentration limits should be set in consideration of the route of administration.

Without proper justification, parenteral limits/PDEs should be used for pharmaceutical substances that are administered by other routes of administration, including inhalation. Oral limits/PDEs may be applied if the absorption by other routes of administration is not likely to exceed the absorption following oral administration. For example, for cutaneous administration, oral concentration limits/PDEs are considered acceptable.

Platinum salts have been shown to be allergenic, with hexachloroplatinic acid being clearly the most allergenic (Malo, 2005). Consequently, a specific limit for inhalation exposure for this molecule has been set at 70 ng day⁻¹ (see monograph). Chromium VI and Nickel, when inhaled, have been associated with carcinogenicity. Therefore, specific limits for inhalation exposure have been set for Chromium VI at 10 ng day⁻¹ and for Nickel at 100 ng day⁻¹ (see respective monographs).

For pharmaceutical products used for short-term and for life-saving indications, as the PDEs and concentration limits mentioned in the TTC guideline are based on chronic use,

higher PDEs and concentration limits may be acceptable in cases of short-term use (30 days or less). For instance, this may be applicable to contrasting agents, antidotes, or products for diagnostic use. This may however only be applied if neither an Option 1 nor an Option 2 limit is feasible.

Specific risk–benefit considerations, such as for compounds used for lifesaving indications, may also warrant the use of higher limits. Justifications should be made on a case-by-case basis.

REFERENCES

- Ball D, Blanchard J, Jacobson-Kram D, McClellan R, McGovern T, Norwood DL, Vogel M, Wolff R, Nagao LM. (2007) Development of safety qualification thresholds and their use in drug product evaluation. *Toxicol. Sci.* 97(2):226–236.
- Bestwick D, Colton R. (2009) Extractables and leachables from single use disposables. *BioProc. Int.* 7:88–94.
- Code of Federal Regulations; Title 21, Volume 5; (April 1, 2005). 21 CFR 314.3(b); Title 21 Food and Drugs, Chapter I—Food and Drug Administration Department of Health and Human Services, Subchapter D—Drugs for Human Use. Part 314, *Applications for FDA Approval to Market a New Drug*, Subpart A—General Provisions, Sec. 314.3 Definitions. Revised as of April 1, 2005. Available at: www.fda.gov.
- Colton R. (2007) Recommendations for extractables and leachables: part 1: introduction, regulatory issues and risk assessment. *BioProc. Int.* 12:36–44.
- EMA. (2008) *Guideline on the Specification Limits for Residues of Metal Catalysts or Metal Reagents*, February 21, 2008. European Medicines Agency, Geneva, Switzerland.
- FDA. (1998) *Guidance for Industry, Metered Dose Inhaler (MDI) and Dry Powder Inhaler (DPI), Drug Products Chemistry, Manufacturing, and Controls Documentation*. Center for Drug Evaluation and Research (CDER), Rockville, MD. Available at: http://www.fda.gov/ohrms/dockets/ac/00/backgrd/3634b1c_sectiond.pdf (accessed May 11, 2016).
- FDA. (2000) *Guidance for Industry, NDAs: Impurities in Drug Substances*. Center for Drug Evaluation and Research (CDER), Rockville, MD. Available at: <http://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/ucm070577.pdf> (accessed May 11, 2016).
- FDA. (2010) *Guidance for Industry, ANDAs: Impurities in Drug Products*. Center for Drug Evaluation and Research (CDER), Rockville, MD. Available at: <http://www.fda.gov/downloads/drugs/guidancecomplianceinformation/guidances/ucm072861.pdf> (accessed May 11, 2016).
- FDA website. (2005) *GMP packaging requirement*. Available at: <http://www.fda.gov/cder/guidance/cGMPs/packaging.htm> (accessed March 31, 2016).
- Gaylor DW, Kodell RL. (1980) Linear interpolation algorithm for low dose assessment of toxic substances. *J. Environ. Pathol. Toxicol.* 4:305.
- ICH. (1999a) *Guidance for Industry, Q6B Specifications: Test Procedures and Acceptance Criteria for Bio-technological/Biological Products*. Available at: <http://www.ich.org/products/guidelines/quality/article/quality-guidelines.html> (accessed March 10, 1999).
- ICH. (1999b) *Guidance for industry, Q6A Specifications: Test Procedures and Acceptance Criteria for New Drug Substances and New Drug Products: Chemical Substances* (October 6, 1999). Available at: <http://www.ich.org/products/guidelines/quality/article/quality-guidelines.html> (accessed October 6, 1999).
- ICH. (2006a) *Guidance for Industry, Q3A(R2) Impurities in New Drug Substances*. Available at: <http://www.ich.org/products/guidelines/quality/article/quality-guidelines.html> (accessed October 25, 2006).
- ICH. (2006b) *Guidance for Industry, Q3B(R2) Impurities in New Drug Products*. Available at: <http://www.ich.org/products/guidelines/quality/article/quality-guidelines.html> (accessed June 6, 2006).
- ICH. (2011) *Guidance for Industry, Q3C(R5) Impurities: Guideline for Residual Solvents*. Available at: <http://www.ich.org/products/guidelines/quality/article/quality-guidelines.html> (accessed February 4, 2011).
- ICH. (2014a) *Guidance for Industry, Q3D Guideline for Elemental Impurities* (December 16, 2014). Available at: <http://www.ich.org/products/guidelines/quality/article/quality-guidelines.html> (accessed July 26, 2013).
- ICH. (2014b) *Guidance for Industry, M7 Assessment and Control of DNA Reactive (Mutagenic) Impurities in Pharmaceutical to Limit Potential Carcinogenic Risk*. Available at: <http://www.ich.org/products/guidelines/multidisciplinary/article/multidisciplinary-guidelines.html> (accessed June 23, 2014).
- Lee H. (Ed.) (2015) *Pharmaceutical Industry Practices on Genotoxic Impurities*. Chromatographic Science Series (Book 107). CRC Press, Boca Raton, FL.
- Malo J-L. (2005) Occupational rhinitis and asthma due to metal salts. *Allergy*. 60:138–139.
- Norwood DL. (2007) Understanding the challenges of extractables and leachables for the pharmaceutical industry—safety and regulatory environment for pharmaceuticals. *Am. Pharm. Rev.* 10(2):32–39.
- Norwood DL, Paskiet D, Roberto M, Feinberg T, Schroeder A, Poochikian G, Wang Q, Deng TJ, DeGrazio F, Munos MK, Nagao LM. (2007) Best practices for extractables and leachables in orally inhaled and nasal drug products: an overview of the PQRI recommendations. *Pharm. Res.* 25:727–738.
- Osterberg RE. (2005a) Extractables and leachables in drug products. *Am. Coll. Toxicol. Newsl.* 25(1):2–3.
- Osterberg RE. (2005b) Potential toxicity of extractables and leachables in drug products. *Am. Pharm. Rev.* 8(2):64–67.
- Teasdale A. (Ed.) (2010) *Genotoxic Impurities*. John Wiley & Sons, Inc, Hoboken, NJ.
- USP. (2008) General chapter on inorganic impurities: heavy metals. *Pharm. Forum.* 34(5):1345–1347.
- World Health Organization (WHO). (1994) Environmental Health Criteria (EHC) 170: assessing human health risks of chemicals: derivation of guidance values for health-based exposure limits. World Health Organization, Geneva, Switzerland. Available at: <http://www.inchem.org/documents/ehc/ehc/ehc170.htm> (accessed May 11, 2016).

TISSUE, CELL, AND GENE THERAPY

This chapter seeks to overview three closely related therapeutic products that are either living biologic components or genetic material directly derived from such components (delivered in or by another living component—a virus). All three share many of the same potential mechanisms of toxicity. These are sometimes called advanced therapies:

1. Infection if improperly prepared or contaminated with pathogens
2. Inflammatory responses if contaminated with endotoxins
3. “Off-target hits”—the therapeutic entity going somewhere it is not intended to be and remaining viable

Gene therapy medicinal products (GTMPs) have additional potential adverse effects as a result of modifying genetic function—potentially including cancer.

There are numerous regulatory guidances on the production of clinical material for these products. This is as it should be, as tight manufacturing controls and ensuring stability of cell and genetic stock involved are the key to safety, and what are to be used in any nonclinical safety evaluations are intended clinical materials. But there are limited guidances available for the nonclinical safety evaluation—EMA (2008) gene therapy product (GTPs) and Japan (Pharmaceutical and Medical Device Agency—PMDA) (2000–2008) and FDA (2013) for cellular and GTPs.

All of the cell-based “advanced therapy medicinal products” (ATMPs) have their safety assessment predicated on a risk-based approach, based on careful consideration of the product itself (both as manufactured and as administered to the intended patients). These products share specific issues such as cell gene viability, longevity, distribution, growth, replication, differentiation, and migration (of somatic

cellular/tissue products) or transmission through the germline and/or reactivation of delivery vectors (for GTPs).

Specific considerations for safety assessment include what is the duration of observation after administration (with the assumption that product administrations to patients will usually be one-time events), what is the appropriate animal model to use for preclinical evaluation, and how are we to extrapolate from animals and predelivery/preexposure profiles to the relevant human situations.

Add to these the issues associated with understanding distribution, incorporation, and potential elimination of administered agents (for traditional PK/TK methods and models neither apply nor are effective here), and the problem of both evaluating safety and developing the optimal clinical product is daunting.

So any development program must proceed through a series of first exploratory and then confirmatory clinical (and potentially preclinical) studies (or study cases).

There are “tagant” compounds that can be easily incorporated in minute amounts into all the advanced therapy products to allow tracking of their distribution in the body by noninvasive means, such as MRI. Because of the very small amounts tagged into or on such a product, such tagants require a strong “signal generator,” such as manganese.

Traditional toxicologic study designs and animal models are not dependably predictive for efficacy or safety in humans, particularly in regard to immunogenicity and tumorigenicity. Differences between biodistributions of introduced cells in animals and humans must be understood, and extrapolation of relative potency between species (remember TGN412, the small molecule discussed in Chapter 11) and the issues seen in the long term with administered monoclonal antibodies are just beginning to be understood.

All nonclinical safety studies should be conducted in accordance with GLP. Genetic toxicity and safety pharmacology are generally not appropriate or required for these products.

33.1 SAFETY ASSESSMENT OF CELL THERAPY (CT) PRODUCTS

The overarching consideration for an adequate preclinical program for a cell therapy product includes (Sharper et al., 2012; FDA, 2013; Galli and Serabian, 2015; Guenther et al., 2016):

1. Establishment of biological plausibility (of therapeutic efficacy).
2. Identification of biologically active administration levels.
3. Selection of potential starting dose level, dose-escalation schedule, and dosing regimen for clinical trials.
4. Establishment of feasibility and reasonable safety of the investigational product's proposed clinical route of administration (ROA).
5. Support for determination of patient eligibility criteria.
6. Identification of physiologic parameters that can guide clinical monitoring.
7. Identification of potential public health risks (e.g., to the general public, caregivers, family members, close contacts (for example, coworkers), and intimate partners).
8. Traditional and well-established transplant tissues (such as bones, corneas, skin, heart valves, and vessels) are not considered to be products subject to these evaluations.

33.1.1 Recommendations for General Preclinical Program Design

When possible, the same investigational CT product that will be administered to the patient population should be used in the definitive preclinical studies. Each lot of an investigational CT product used in the preclinical *in vitro* and *in vivo* studies should be characterized according to appropriate criteria, consistent with the stage of product development. Similarities and differences between product lots intended for preclinical use and lots intended for clinical use should be highlighted and discussed in the IND submission. However, in certain cases, due to the species-specific nature of the clinical product (e.g., some vector-expressed human transgenes: human-derived cellular therapy (CT) products), testing the CT product intended for clinical administration in animals may not be informative, and therefore testing of an

analogous product may be a suitable alternative. In these situations, the design of the preclinical testing program is considered on a product-by-product basis. Considerations regarding investigational product incompatibility issues as outlined in the operative FDA guidance are essential.

33.1.2 Model Species Selection

The selected model for assessment of bioactivity and safety should have demonstrated a biological response to the investigational CT product similar to that expected in humans in order to generate data to guide clinical trial design. Some factors that should be considered when determining the relevant species include (i) comparability of physiology and anatomy to that of humans; (ii) permissiveness/susceptibility to infection by, and replication of, viral vectors or microbial vectors for gene therapy (GT); (iii) immune tolerance to a human CT product; and (iv) feasibility of using the planned clinical delivery system/procedure.

Assessment of these factors necessitates consideration of the specific product and clinical indication. "Nonstandard" test species, such as genetically modified rodents (i.e., transgenics or knockouts) or large animals (e.g., sheep, pigs, goats, and horses) may be acceptable when adequate justification is provided. Although safety and effectiveness of the investigational CT product *in vitro* and *in vivo* can possibly be evaluated in one animal species (and usually is), other contributory factors (e.g., source of the CGT product, ROA) may result in the need for testing in more than one species. Prior to initiation of the definitive preclinical studies, it is recommended that conduct of *in vitro* studies (e.g., functional assays, immunophenotyping, morphologic evaluation) and *in vivo* pilot studies be conducted to establish the biological relevance of a specific animal species to the investigational product(s).

Detailed assessment of the relevancy of each animal species used in support of each potential clinical trial should be conducted prior to initiation of said trial. A summary of this assessment should be submitted as part of the preclinical section of the IND.

33.1.3 Selection of Animal Models of Disease/Injury

Preclinical studies performed in animal models of disease/injury may provide insight regarding the relationships of dose to activity and toxicity and are recommended where feasible. Animal models of disease/injury used in basic research or discovery science phases of product development are also potentially useful for generating data to support clinical trials for CT products. Due to features of CT products (e.g., potentially prolonged duration of intended product effect, product persistence *in vivo*, complex mode of action (MOA) involving interaction between the CT product and the disease environment, and invasive ROA), animal models of disease/injury may be preferable to healthy animals

to assess the activity and safety of these products. Therefore, preclinical studies in disease/injury models are encouraged to better define the risk/benefit ratio associated with investigational CT products. In addition, use of disease/injury models provides the opportunity for possible identification of activity-risk biomarkers that may be applicable for monitoring in clinical trials.

33.1.4 Information Describing Limitations of Potential Animal Model(s)

Potential limitations of these preclinical animal models may exist and should be identified and considered. Examples of these limitations include:

- (i) Inherent variability of the model (potentially age or gender dependent)
- (ii) Limited historical/baseline data for the model
- (iii) Technical limitations with the physiological and anatomical constraints of the model
- (iv) Animal care issues
- (v) Limited fidelity in modeling human pathophysiology of the disease/injury of interest

Every model has inherent strengths and weaknesses and these must be well understood. No single model will predict with complete accuracy the efficacy and safety outcome of the investigational CT product in the patient population. The activity and safety profile of the CT product may be influenced by the timing of administration relative to the onset of disease; thus the disease state at the initiation of product administration should be characterized and documented in the IND submission.

33.1.5 Information Supporting the Choice of Animal Model(s)

- (i) The similarities and differences between the pathophysiology of the disease/injury animal model and the pathophysiology of the disease/injury of humans (Gad, 2015)
- (ii) The effect of the disease/injury status of the animal on the cellular viability and the pharmacology/toxicology of the investigational CT product (i.e., altered sensitivity of the animal model to the specific product under study)
- (iii) Detrimental effects of the administered product on existing disease/injury status (i.e., exacerbation of an existing disease/injury condition or induction of a new disease/toxicity)

It is recommended that consideration be given to using a tiered approach for determining selection of an appropriate

animal model. Performance of pilot studies involving the intended investigational CT product may assist in evaluating the suitability of a particular animal species/model for use in the definitive preclinical studies. Moreover, multiple animal models may be necessary to adequately identify functional aspects and potential toxicities of a single product under study. In these situations, the preclinical testing paradigm may include the use of (i) large and small animal models, (ii) multiple small animal models, or (iii) only large animal models.

The number and type of studies performed should be guided by the biological attributes of the investigational product. One should refer to current CBER guidances (FDA, 1998, 2013) that include information and recommendations regarding tiered testing approaches for the products.

33.1.6 Proof-of-Concept (POC) Studies

A primary objective of proof-of-concept (POC) studies is to establish the feasibility and rationale for use of an investigational CT product in the targeted patient population. POC studies help inform the benefit side of the risk/benefit assessment of the CT product. Such data may be essential in the assessment of novel products with substantial inherent risks. In addition, data from POC studies can contribute significantly to animal species selection.

POC studies should investigate the following:

- a. The pharmacologically effective dose range (i.e., minimally effective dose and optimal biological dose)
- b. Optimization of the ROA and confirmation that the product reaches the target anatomic site/tissue/cell
- c. Optimization of the timing of product administration (regimen of administration) relative to onset of disease/injury
- d. Optimization of the dosing schedule (regimen)
- e. Characterization of the putative MOA or hypothesized biological activities of the investigational product

Collectively, this information serves to establish the rationale for, and feasibility of, the proposed clinical trial. Features of study design, such as the inclusion of appropriate concurrent controls, randomization, or blinding methods, may increase the strength of the resulting study data and thus should be considered.

Preclinical *in vitro* assays intended to assess aspects of the biological activity of an investigational product (e.g., growth factor secretion, immunological response profile, expression of a neurotransmitter) can provide supporting POC information.

Data derived from *in vitro* and *in vivo* preclinical POC testing are useful in guiding the design of both preclinical

toxicology studies and the early-phase clinical trials while contributing to defining reasonable risk for the investigational product in the intended patient population.

33.1.7 Toxicology Studies

Preclinical assessment of the safety of an investigational product contributes to the definition of an acceptable risk/benefit ratio for a proposed clinical trial. The safety assessment should be sufficiently comprehensive to permit identification, characterization, and quantification of potential local and systemic toxicities, their onset (i.e., acute or delayed), the possibility for resolution of any toxicities, and the effect of product dose level on toxicity findings and provides guidance on how to monitor for this in nonclinical and clinical studies.

33.1.7.1 Primary Considerations for Toxicology Study Design Each of the following should be considered in the design of the toxicology study:

- (i) The proposed clinical indication
- (ii) The amount and quality of published preclinical or clinical safety information for the specific CT product under investigation or for a similar product (i.e., known toxicities or adverse effects)
- (iii) The amount and quality of existing pharmacology (*in vitro/in vivo*) or POC data for the specific CT product under investigation or for a similar product
- (iv) Previous preclinical/clinical experience with the proposed clinical delivery device/delivery procedure or with any related device/procedure
- (v) The biological responsiveness of the animal species to the investigational CT product
- (vi) The putative MOA of the CT product
- (vii) The intrinsic properties of the CT product
- (viii) The pathophysiology of the animal disease/injury model, if one is used

Animal species in which the product is biologically active should be used in the nonclinical safety studies. Although healthy animals represent the standard model test system employed to conduct traditional toxicology studies, study designs using animal models of disease/injury are frequently modified to incorporate important safety parameters that allow for assessment of the potential toxicology of an investigational product (i.e., hybrid pharmacology/toxicology study design). Such data can supplement or possibly be used in lieu of toxicology studies in healthy animals.

33.1.7.2 Secondary Considerations for Toxicology Study Design The overall design of the toxicology studies should

mimic the proposed clinical trial design as closely as possible. Preclinical toxicology study designs should include the following, as applicable:

- (i) Adequate numbers of animals per gender randomized to each group. If the number of animals that can be dosed in a single day is limited due to the complexity of the dosing procedure or the timing of product administration relative to disease status of the model, then appropriate blocked randomization methods or other factors should be considered in an attempt to reduce study bias as much as possible. The number of animals required for each group will vary depending on the safety concerns for the investigational product, the species, the model, and the delivery system.
- (ii) Appropriate control groups. Examples include animals that are left untreated, receive sham surgery, or are administered formulation vehicle only, adjuvant alone, null vector, or scaffold alone. Justification should be provided for the specific control group(s) selected.
- (iii) Multiple dose levels of the investigational product should bracket the proposed clinical dose range, if feasible. Results obtained from POC studies should guide selection of the target dose levels for both preclinical safety assessment and for clinical development. The highest dose level used in preclinical models may be restricted due to animal size, tissue volume/size, ROA, or product manufacturing capacity. Justification, with supporting data, should be provided for the specific dose levels selected.
- (iv) A dosing schedule that reflects the intended clinical dosing regimen, to the extent possible.
- (v) An ROA that mimics the intended clinical route as closely as possible. The delivery device intended for use in the clinical studies should be used to administer the investigational product in the definitive toxicology studies; justification should be provided if the intended clinical delivery device is not used.
- (vi) Multiple sacrifice time points to capture potential acute, adaptive immune, chronic, and/or delayed-onset toxicities, as well as the potential for resolution of toxicities. The time intervals designated for the sacrifice time points will depend on the animal model used, the investigational product, the dosing schedule, the pharmacodynamic and pharmacokinetic response observed, and the proposed patient population.
- (vii) Safety end points that capture potential toxicities should be carefully identified and incorporated.

Standard parameters monitored include mortality (with cause of death determined, if possible), clinical observations, body weights, physical examinations, food consumption/appetite, water consumption (as applicable), clinical pathology (serum chemistry, hematology, coagulation, urinalysis, special biomarkers), organ weights, gross pathology, and histopathology.

- (viii) Additional parameters specific to either the investigational product used or the intended patient population. Examples of product-specific study parameters include humoral or cellular immune responses, biodistribution of cell tissues, product fate, behavioral testing, neurological exams, ophthalmic exams, cardiac assessments, imaging (i.e., MRI, ultrasound, radiography—see Chapter 24 on imaging), presence of abnormal/ectopic growths (i.e., hyperplasia, tumors), putative biomarkers, and specialized histopathology (i.e., immunohistochemistry). The data collected should include both morphological and functional assessment, whenever possible, to determine whether an association exists between nonterminal and terminal findings. Reversibility of any findings should also be addressed. Refer to other sections of this document for guidance that is specific to product class.

Collected preclinical data, considered as an integrated whole, will help guide clinical trial design. For example, data generated from the toxicology studies will potentially establish a no-observed-adverse-effect level (NOAEL), which will help determine the selection of the starting dose level and subsequent dose-escalation scheme for the clinical trial. This information should potentially allow for avoidance or minimization of significant toxicities in patients.

33.1.8 Product Delivery Considerations

The ROA used to deliver the investigational product in the definitive preclinical studies should mimic the ROA to be employed in the clinical setting to the greatest degree possible. If it is not possible to replicate the clinical ROA in the animal model, then alternative routes/methods should be proposed and scientifically justified as a part of the preclinical development plan.

To assess the potential risks associated with the method of product administration, the delivery device system used in the definitive preclinical studies should be identical to the planned clinical product delivery device, if possible. In definitive studies where the planned clinical delivery device system cannot be used, justification for the delivery system that is utilized should be provided. The safety of the delivery device and/or system must also be established. The IND

submission should state whether a device master file (MAF) has been submitted to the Center for Devices and Radiological Health (CDRH) for the delivery device. If a MAF exists, the IND submission should include a letter of authorization from the MAF holder granting permission for FDA to cross-reference specific information in the MAF (a matching copy of the letter must be put in the MAF folder by the holder). CBER will consult with CDRH review staff as necessary to determine whether the information provided in the device MAF is sufficient in detail (e.g., facilities and manufacturing procedures and controls; synthesis, formulation, purification, and specifications for chemicals, materials; biocompatibility (CDRH, 1995), preclinical data; clinical study data) to support use in the clinical trial. If a MAF for the delivery device does not exist or if the information is not sufficient to support the proposed use, CDRH review staff may be consulted to determine the type and extent of information that should be included in the IND submission to support the use of the device in the proposed clinical trial.

The use of a large animal species (healthy animal or a disease/injury model) to test the safety of a delivery device may be appropriate in certain situations, such as assessment of risk associated with use of a previously untested device for intracranial product delivery or assessment of risk associated with use of an investigational delivery system for placement of the product into the heart or the brain. Safety data for the delivery device and delivery procedure may derive from existing active regulatory submissions (i.e., INDs, Investigational Device Exemptions (IDEs), MAFs). As also indicated earlier, in these circumstances, the IND submission for the investigational product should include letters of cross-reference from the sponsors of these existing submissions. Published studies that involve the use of the clinical delivery device and delivery procedure may also provide supportive safety data.

Cell therapy (CT) products vary with respect to characteristics such as formulation (including combination with a scaffold or other noncellular component), the genetic relationship of the cells to the patient (autologous, allogeneic, xenogeneic), and the cell source. CT products can be generally classified as stem cell-derived CT products or mature/functionally differentiated cell-derived CT products. If the CT product is derived from an induced pluripotent stem cell (iPSC), the product has the possibility of expressing characteristics of both stem cell-derived and mature/functionally differentiated cell-derived products; therefore, both fundamental source categories of CT products should be considered during the product development process. The *in vivo* biological activity and safety profile of the investigational CT product is influenced by product origin (donor and tissue source), as well as the level of manipulation and stage of differentiation at the time of administration. Regardless of the type of CT product, if the cells originate from animal

tissue or cells (xenotransplantation products), additional considerations apply:

1. Tissue sources of stem cells include (i) adult (e.g., hematopoietic, neural, mesenchymal, cardiac, adipose, skin), (ii) perinatal (e.g., placental, umbilical cord blood), (iii) fetal (e.g., amniotic fluid, neural), and (iv) embryonic. Stem cell-derived products are characterized by a variable capacity for self-renewing replication through cycles of cell division and the capacity for differentiation into a variety of cell types with specialized properties/functions. Such differentiation and replication are primarily controlled by the physiologic milieu of the host in which the cells reside following *in vivo* administration. Similarly, contamination of a differentiated CT product with undifferentiated stem cells or incompletely differentiated progenitor/precursor cells poses potential safety concerns.
2. Functionally differentiated tissue-derived CT products may be obtained from adult human donors (autologous or allogeneic) or from animal sources (xenogeneic). Source cells can include chondrocytes, pancreatic islet cells, hepatocytes, neuronal cells, and various immune cells. CT products derived from functionally mature tissues typically do not possess the property of self-renewing proliferation and the capacity to differentiate into multiple cell types; however, they may retain some cellular characteristics of their tissue of origin. Additionally, their characteristics may change after *in vivo* administration, based on specific extracellular cues.

Additional considerations for animal model selection for CT products can include:

1. The ability to access the anatomic site for product administration
2. The ability to deliver a specific absolute cell dose to the target site
3. The availability of immunodeficient animals, which may allow for long-term assessment of the safety of the human CT product

Administration of human cells into animals is complicated by the immunogenic responses of healthy immunocompetent animals, potentially resulting in the rejection of the administered human cells. This response may prevent adequate evaluation of the activity and safety of the human cellular product. When conducting preclinical studies to evaluate the activity and safety of a human cellular product, the cross-species immunogenicity may necessitate alteration of the animal model in order to create an *in vivo* immune

tolerant niche for the administered human cells. Various models that have been considered include:

- Immunosuppressive agents in immunocompetent animals
- Genetically immunodeficient animals
- Humanized animals
- Administration into an immune privileged site
- A combination of these scenarios

The administration of analogous cellular products in the preclinical studies is also a potentially acceptable option. The scientific value of this approach is optimized when the analogous CT product is substantially similar to the human CT product. However, preclinical testing using an analogous cellular product can introduce uncertainty regarding the relevance of the data due to potentially different biological activities, molecular regulatory mechanisms, and impurities/contaminants. Therefore, if this preclinical testing pathway is used, the level of analogy of the animal cellular product with the intended human cellular product should be characterized. Examples may include:

- Established procedures for tissue/sample harvest
- Cell identification, isolation, expansion, and *in vitro* culture procedures
- Cell growth kinetics (e.g., cell doubling time, cell growth curve, and time to cell proliferation plateau)
- Phenotype and functional properties (e.g., secretion of growth factors and cytokines, cell population-specific phenotypic/genotypic markers)
- Final product formulation/cell scaffold seeding procedures (as applicable)
- Final product storage conditions and cell viability

The degree of similarity of these parameters for the analogous CT product should be as close to the proposed human CT product as possible in an attempt to maximize the relevance of data derived from the animal studies.

33.1.9 Study Designs

In addition to the general guidance on the preclinical testing program provided earlier, considerations when designing preclinical studies for investigational CT products include the following:

1. The targeted cellular phenotype(s)
2. The source of the cell(s)
3. The extent of *ex vivo* manipulation performed (e.g., selection, purification, expansion, activation)
4. The fate of the cell postadministration (engraftment, migration, differentiation, tumorigenicity)

5. The probability of a host's immune response to the administered cells
6. Administration site reactions
7. Potential inflammatory response in target and/or non-target tissues
8. Unregulated/dysregulated proliferation of the cells within the host

Determination of the fate of the investigational CT product following administration in animals is an important contribution to characterizing the product activity and safety profile. When conducted early in the preclinical testing program, assessment of cell fate can help characterize the putative MOA by determining if engraftment is important and necessary to achieve the desired pharmacological response. Additionally, cell fate can help justify the selection of the animal species/models, justify the duration of the definitive studies, and identify potential target organs of toxicity.

Considerations of cell fate *in vivo* include:

1. Survival/engraftment

Cell viability and subsequent engraftment may be affected by:

- a. The biocompatibility of the cell delivery device and the CT product (considerations include cell shearing and adsorption onto the walls of the catheter/syringe)
- b. The ROA
- c. The genetic relationship of the cells to the host animal (autologous/syngeneic, allogeneic, or xenogeneic)
- d. The immune status of the host animal
- e. The timing of cell administration relative to the onset of the disease/injury (i.e., the pathophysiologic status of the microenvironment)

If long-term cell survival/engraftment is necessary to achieve effectiveness of the CT product, the animals should be followed for an interval sufficient to allow for comprehensive evaluation of *in vivo* cell survival, anatomic engraftment, and biologic activity.

As a consequence of their biologic attributes, CT products administered *in vivo* are not subject to conventional chemical analyses; therefore, standard ADME and pharmacokinetic testing techniques and profiles are not applicable. Although influenced by specifics of the CT product and its ROA, cells have an inherent potential to distribute to sites other than to the target organ/tissue. Various methods, such as imaging modalities used for detection of radioisotope-labeled cells, genetically modified cells (e.g., expressing

green fluorescent protein), nanoparticle-labeled cells (e.g., iron–dextran nanoparticles), or the use of polymerase chain reaction (PCR) analysis and immunohistochemistry to identify cells of human origin or cells of a karyotype different than the host (e.g., gender), have been used to assess distribution. A potential advantage of *in vivo* imaging techniques is that in many instances, the same animal can be evaluated over time, thus decreasing variability and reducing the number of animals used. Data should be provided to support the viability and function of the CT product if the cells are modified to enable use of such imaging techniques.

Cellular differentiation capacity, the plasticity of phenotypic expression attributable to transdifferentiation or fusion with other cell types, and structural and functional tissue integration may all be influenced by physiologic factors within either the local microenvironment into which the CT product is administered or the final location/niche in which the cells ultimately reside. Therefore, conditions found within the local microenvironment into which the cells are placed are likely to have an impact on the safety and/or bioactivity of the CT product. Given the biological attributes of some CT products, the potential for ectopic expression in target and nontarget tissues also exists. Depending on their differentiation status and the extent of manipulation the cells undergo prior to *in vivo* administration, parameters such as cell morphology, phenotype, and level of differentiation following *in vivo* administration should be assessed in the animal studies.

The potential for tumorigenicity, dysplasia, or hyperplasia to occur should be considered and addressed as appropriate for the specific biologic properties of each investigational CT product. Factors that may influence the tumorigenicity assessment include:

- a. The differentiation status profile of cell types within the CT product (ranging from undifferentiated/embryonic to terminally differentiated/specialized)
- b. The extent of cell manipulation employed during manufacture of the product and the resulting growth kinetic profile (e.g., minimal, culture expansion only, culture expansion with/without growth factors, *ex vivo* differentiation, *ex vivo* transduction with or without cell expansion)
- c. The expressed transgene (e.g., various growth factors) of genetically modified cells
- d. The potential to induce or enhance tumor formation from existing subclinical host malignant cells
- e. The target patient population

Studies conducted in animals to assess tumorigenicity should use the intended clinical product, not analogous animal cells. There is currently no scientific consensus regarding the selection of the most relevant animal models to

evaluate tumorigenic potential or the ability of current animal models to predict clinical outcome. However, it is important that animal studies designed to assess this end point for CT products show *in vivo* survival of the cells for a sufficient length of time to allow for potential tumor formation; additional study design considerations include (i) appropriate control groups (e.g., undifferentiated cells, partially differentiated cells, positive controls, vehicle controls); (ii) adequate numbers of animals per group to ensure statistical significance of any biological observations, including any background incidence of tumor formation; (iii) inclusion of at least one dose level that constitutes the maximum absolute amount of cells that can be administered; (iv) delivery of the CT product targeting the planned clinical anatomic site; and (v) sufficient study duration.

33.1.10 CT Products with Implantable Scaffolds

Similar to all CT products, cell characterization should be provided prior to scaffold seeding to support use of the CT component of the multiple-component product.

Any scaffold construct (synthetic or nonsynthetic polymers) used should be identical in composition to the intended clinical scaffold. The scaffold should be adequately characterized for composition, degradation profile, biomechanical performance, and biocompatibility (with respect to host response to the scaffold component and to the cell component of the product). The specific tests that are needed to sufficiently characterize a scaffold are determined by its composition and intended use. The specific testing expectations for scaffold materials will share some features similar to the testing expected if the scaffold were to be used as an implant alone. However, the details of the manufacturing process and the cells used will likely influence the specific tests needed.

Depending on the material(s) that constitutes the intended clinical product, biocompatibility testing may be warranted. Biocompatibility test results in accordance with the ISO 10993-1 (June 2016) “Use of International Standard ISO-10993, Biological Evaluation of Medical Devices Part-1: Evaluation and Testing,” should be consulted for approaches to biocompatibility testing (see Gad and Gad-McDonald, 2015).

Groups should consist of animals administered with the intended clinical product (i.e., scaffold seeded at varying cell densities) and appropriate controls.

Safety and POC of the administered product and product components should be demonstrated via inclusion of biochemical, morphological (i.e., composition and architecture of the tissue), and functional end points. Functional end points may include mechanical testing, which will depend on product design, product components, the method/location of product administration, putative MOA, and disease indication. The mechanical properties of the repaired, replaced, or regenerated tissue should be compared to appropriate concurrent controls.

The optimal dose and length of time needed to assess repair, replacement, or regeneration of clinical lesions (i.e., construct performance) and the durability of the effect should be determined. In addition, the biodegradation profile of the scaffold construct should be evaluated. The study duration will vary based on the product and the clinical indication but should be sufficient to provide data to show durability of effect. For example, study duration of 1 year in a relevant animal injury model is recommended for determination of product performance and assessment of durability for products intended for repair/replacement of knee cartilage.

Local toxicities (e.g., tumorigenicity, altered tissue function at the injection site, inappropriate cellular differentiation, or inflammatory infiltrates) may be due to interactions of the product components with the tissue or to the degradation of product components at the site of administration. Cell migration outside of the target tissue may lead to systemic toxicities, such as ectopic tissue formation and tumorigenicity. The immunogenic potential of the construct (i.e., the scaffold and/or the cells) could also cause toxicity. Both acute and long-term *in vivo* safety of the product should be evaluated.

33.2 NONCLINICAL SAFETY ASSESSMENT OF GENE THERAPY PRODUCTS (GTPS)

GTPs are unusual in that they are subject to regulatory oversight not only by FDA but also by an NIH committee. This dual oversight reflects both the still in flux status of the science and the complication of the process (Takefman and Bi, 2015; Templeton, 2015).

33.2.1 CBER

- Division of Cell and Gene Therapy
 - Manufacturing.
 - Note: 20% of CBER clinical protocols are now for GT.
- Division of Clinical Trial Design and Analysis
 - Preclinical pharmacology and toxicity
 - Clinical trial design, safety, and efficiency

33.2.2 NIH

- Recombinant DNA Advisory Committee (RAC)
 - No authority for approval for clinical trials

Considerations for assessing the safety of GTPs are more complex because of the potential longer-term readouts as to adverse effects and greater difficulty of observing and evaluating tissue distribution of administered product. Such products also have both (therapeutically) active and inactive compounds.

There are two major potential contributors specific to GTPs—the viral vectors for gene delivery (for which acute toxicity, including cytotoxicity immune responses, potentially includes toxic shock and risk of insertional mutagenesis) and the toxicity of expressed proteins arising from the inserted genetic material (Verdier and Descotes, 1999).

GTPs have significant potential to evoke an immune response from the host (either animal model or human in clinical trials or use), which have possible adverse effects and immune neutralization.

Both EMEA (2008) and FDA (2013) have provided specific guidances as to nonclinical evaluation of GTPs.

Specific considerations for the selection of relevant animal species/model for investigational GTPs include:

1. Assessment of the permissiveness/susceptibility of various animal species to infection by, and replication of, the viral vector
2. Pharmacological response of the species to the expressed transgene
3. Sensitivity of the species to the biological actions of the *ex vivo* genetically modified cells
4. Evaluation of DNA/RNA biodistribution
5. Evaluation of gene transfer and biological activity
6. Evaluation of the risk of vertical transmission of the gene
 - a. Assessment of the safety of the vector
 - b. Assess the safety of the product

If the previous parameters relevant to a specific GTP cannot be met using common laboratory animal species, modifications should be considered. For example, genetically modified animals expressing the human receptor target have been used to characterize the biologic activity, and thus the potential pathology, of some viruses. Similarly, immunodeficient animals have been used to evaluate the safety of genetically modified human cells. In instances where the expressed transgene is not biologically active in the animal species, use of the clinical vector expressing an analogous transgene that is active in the laboratory species may suffice, especially if clinical or preclinical data for the expressed protein exist. In such instances, comparison of the intended clinical product and the animal analog should be provided (e.g., sequence, target specificity, expression levels).

33.2.3 Study Designs

Safety concerns for both *ex vivo* and *in vivo* administration of an investigational GTP derive from multiple factors, such as the potential for adverse immune responses to the *ex vivo* genetically modified cells, the vector, or the expressed transgene; vector and transgene toxicities; and the potential risks

of the delivery procedure. For example, administration of genetically modified cells or vector to vital organs, such as the brain or heart, generates concerns for potential toxicity from the product itself, as well as for possible risks associated with the delivery device and the delivery procedure. These issues should be addressed in the preclinical testing program before initiation of clinical trials.

Although assessment of the safety of the *in vivo* administered vector depends on the biological properties of each vector type, concerns that should be addressed include:

- a. Toxicities due to the components of the final formulation (e.g., liposomes and various excipients/contaminants)
- b. Toxicities due to the ROA used
- c. Aberrant localization to nontarget cells/tissues
- d. Level and persistence of vector and expressed transgene
- e. Level of viral replication in nontarget cells/tissues
- f. Immune activation or suppression
- g. Immune response directed against the vector
- h. Phenotype/activation state of target cell(s)
- i. Potential for insertional mutagenesis or oncogenicity
- j. Potential for germline transmission
- k. Potential horizontal transmission of replication-competent vectors from the patient to family members and healthcare providers (i.e., shedding)

Some examples of potential toxicity characteristics of specific vector types include:

- a. Nonviral vectors—Potential for immune response to the DNA or to extraneous bacterial sequences
- b. Replication-deficient viral vectors:
 - i. Adenovirus—Potential for a significant immune response and inflammatory response to the vector and possible adverse effects from any contaminating replication-competent adenovirus.
 - ii. Adeno-associated virus (AAV)—(i) Although AAV remains episomal in the transduced cell, the potential for random integration into host DNA, resulting in insertional mutagenesis and any subsequent adverse biological effects, exists and (ii) potential immune response to the capsid proteins.
 - iii. Retrovirus and lentivirus—(i) Production of a replication-competent retrovirus/lentivirus (RCR/RCL) during manufacturing; (ii) potential for insertional mutagenesis, resulting in oncogene activation; (iii) potential for germline integration; and (iv) potential for altered expression of host genes.

- iv. Poxvirus—(i) Ability to infect and replicate in many types of human tissues and cells, (ii) potential for toxicity in immunocompromised populations such as cancer patients, and (iii) renal/cardiac concerns.
- v. Herpes simplex virus (HSV)—Tropism to the central nervous system and the potential for latency and reactivation.
- c. Replication-competent oncolytic vectors—(i) Potential viral infection and replication in normal cells and (ii) increased viral spread and replication in nontarget tissues in immunosuppressed patients or when administered in combination with radiation, chemotherapy, prodrugs, or other agents
- d. Microbial vectors used for GT—(i) Lack of adequate attenuation of the microbe, (ii) ability to replicate in nontarget tissues, (iii) excessive induction of proinflammatory cytokines, and (iv) lack of antibiotic susceptibility

Although the appropriate duration of clinical follow-up of GT trial participants for adverse events is primarily a trial design issue, vector characteristics and preclinical data are used to inform clinical trial decisions. CBER has issued a guidance for industry entitled “Guidance for Industry: Gene Therapy Clinical Trials—Observing Subjects for Delayed Adverse Events” dated November 2006 (FDA, 2006), which discusses the potential risks of delayed adverse events following exposure to GTPs as a consequence of persistent biological activity of the genetic material or other components of the products used to carry the genetic material. As specified in that guidance, factors that are likely to increase the risk of delayed adverse events in humans include persistence of the viral vector, integration of genetic material into the host genome, prolonged expression of the transgene, and altered expression of the host’s genes. The 2006 guidance should be consulted and, if found to be applicable to the investigational GTPs under study, should be used to guide the design of relevant preclinical studies to address potential long-term safety issues that may result from administration to humans.

When determining the safety of an expressed transgene and/or translated protein, sponsors should consider the following: (i) local versus systemic expression, (ii) level and duration of expression, and (iii) acute versus chronic effects. While persistent transgene expression may be a desired end point for some GTPs, it can also be an undesired outcome for other products due to overexpression, accumulation of transgene protein, or risk of an abnormal immune response. Prolonged expression of transgenes such as growth factors, growth factor receptors, or immunomodulating agents may be associated with long-term risks due to unregulated cell growth, malignant transformation, autoimmune reactions to self-antigens, altered expression of the host’s genes, or other

unanticipated adverse effects. The conduct of long-term preclinical studies should be considered to evaluate these concerns.

In addition, assessment of the *in vivo* transgene expression profile is recommended for vectors expressing a new transgene with an unknown potential to induce toxicity or vectors expressing a transgene with a known or suspected potential to induce toxicity if aberrantly expressed in nontarget tissues. Quantitation of transgene expression using methods such as a quantitative reverse transcriptase PCR (RT-PCR) assay can help determine the threshold level of expression associated with beneficial or deleterious effects for specific tissues/organ systems and correlation of the kinetics of transgene expression with desired activity or undesired toxicity profiles.

In addition, potential immunogenic/neutralization responses directed against the expressed transgene and/or directed against self/endogenous proteins can be a concern. For example, delivery of transgenes that encode various endogenous enzymes, receptors, or structural proteins may elicit antibodies against both the transgene and the endogenous components expressed in normal cells and tissues, resulting in an adverse response. Similarly, transgenes that express fusion or chimera proteins can theoretically be immunogenic due to their foreign (xenogeneic) nature. These concerns should be addressed in the preclinical testing program.

33.2.4 *Ex Vivo* Genetically Modified Cells

The safety assessment of the cellular component of *ex vivo* transduced cells includes end points that are similar to those evaluated for CT products. The significance of the issues described in these sections will depend on the cell type(s), the vector construct, and/or the transgene used. The preclinical study designs should address relevant factors specific to each product. It should be noted that occurrence of anaphylactic responses cannot be evaluated in mice.

33.2.5 Biodistribution Considerations

The characterization of the vector biodistribution profile following *in vivo* administration is an important component of the preclinical development program for GTPs. These data are used to determine the potential for vector presence in desired target tissues/biological fluids (e.g., blood, cerebrospinal fluid), in nontarget tissues/biological fluids, and in the germline. The characterization of the vector presence, persistence, and clearance profile can inform the selection of the GTP dosing schedule, the monitoring schedule for various activity/safety parameters, and the animal sacrifice time points in the definitive preclinical studies (Table 33.1). The biodistribution data, coupled with other preclinical safety end points such as clinical pathology and histopathology, help determine whether vector presence

TABLE 33.1 Pharmacokinetic Considerations for Gene Therapy

| Pharmacokinetic Property | Gene Therapy Property |
|--------------------------|---|
| Absorption | DNA vector distribution |
| Distribution | Vector fraction target cell uptake |
| Distribution | Genetic material traffic in organelles |
| Metabolism | DNA degradation |
| Metabolism | mRNA production |
| Metabolism | Protein production—quantity |
| Excretion | Protein production—stability |
| Metabolism and excretion | Protein production—compartmentalization |
| Excretion | Protein production—secretory fate |

or gene expression correlates with any tissue-specific detrimental effects in the animals.

Prior to administration in humans, biodistribution studies should be considered for:

- Investigational GTPs that belong to a new vector class
- Established vectors with significant changes in the vector backbone
- Established vectors with a significant formulation change
- Established vectors with a significant change in the ROA
- Established vectors with a significant change in the dosing schedule and/or the vector dose levels

Justification should be provided if biodistribution studies are not conducted prior to initiation of early-phase clinical trials.

Tissue/biological fluid analysis should be conducted at the molecular level using a quantitative PCR (qPCR) assay to determine the number of vector copies per microgram of genomic DNA at specified time points postvector administration. Other methods for evaluating biodistribution are radiolabeling and Southern blot. Depending on the ROA and biology of the investigational GTP (vector type and expressed transgene), additional tissues may need to be collected and analyzed. In addition, the presence of a vector sequence in tissues/biological fluids may trigger further analysis to determine the transgene expression levels. Evaluating gene transfer and biological activity would include immunohistochemistry, Western blot, ELISA, and flow cytometry.

Single-dose toxicity studies provide information about systemic and local toxicity. To assess the nonspecific toxicity of the final product, the use of a parenteral route is recommended to maximize exposure to the product. When results of acute studies are the only preclinical toxicology data to be presented prior to single-dose administration to

humans, additional parameters, such as clinical pathology and histology, should be investigated.

Repeated-dose toxicity studies are required when several administrations are scheduled in clinical trials. The design of these studies should be as close as possible to the human dosing regimen. Therefore, daily treatments are often likely to be irrelevant and sequential administrations more appropriate, to mimic the human treatment design. For instance, three administrations every 3 weeks can be used to reproduce exactly the human vaccination schedule for a genetically modified vaccine or, alternatively, two injections per week for 3 weeks in animals versus three administrations every 3 weeks in humans for an anticancer viral vector product. In the latter example, animal exposure is increased while keeping the intermittent treatment design. The immunogenicity (i.e., the potential of the product to trigger an immune response) in the selected animal species is frequently a limiting factor (remember no anaphylactic response in mice). Both humoral and cellular responses can be addressed, especially antibodies against the vector and the expressed protein (using ELISA) and lymphoproliferation and cytotoxic T-lymphocyte (CTL) activity. The latter test requires that the target cells (e.g., adenovirus-infected target cells) express the product antigens, which is technically difficult. Potential adverse reactions in relation to the immune response against the administered product (e.g., immune complex deposition) should also be investigated, as such reactions may occur in human patients. The treated animals have to be examined, even during the postimmunization period. Immunogenicity can be a useful indicator of the efficacy of the product, which may be neutralized by antibodies. The information gained may also be of value for the selection of the less immunogenic vectors. For instance, doubly deleted ($\Delta E1$ and $\Delta E4$) adenovirus vectors have been shown to express fewer viral proteins than singly deleted ($\Delta E1$) adenovirus vectors and to cause diminished hepatotoxicity and lower immune response toward the vector following parenteral administration in the mouse.

33.3 DEFINITIONS

ATMPs—Advanced therapy medicinal products

Gene therapy medicinal products (GTMPs)—Biological products that contain an active substance which contains or consists of a recombinant nucleic acid used in and administered to human being with the intent of regulating, repairing, replacing, adding, or deleting a genetic sequence

Somatic cell therapy—Substantially manipulates cells or tissues with the intent of treating, preventing, or diagnosing a disease through the pharmacologic, immunologic, or metabolic action

Stem cell therapy—The use of pluripotent stem cells induced to treat degenerative diseases, inflammatory conditions, repair of damaged tissues, and cancer

Delivery device—Any material to be used with the GTP or in which the final GTP is prepared, serving to deliver/administer the GTP

REFERENCES

- CDRH (1995) Blue Book Memo #G95-1: Use of International Standard ISO-10993, 'Biological Evaluation of Medical Devices Part 1: Evaluation and Testing' (Replaces #G87-1 #8294); May 1995. Available at: <http://www.fda.gov/MedicalDevices/DeviceRegulationandGuidance/GuidanceDocuments/ucm080735.htm> (accessed May 28, 2016).
- EMA. (2008) *Guideline on the Non-Clinical Studies Required Before First Clinical Use of Gene Therapy Medicinal Products*, May 30, 2008. European Medicines Agency (EMA), London.
- FDA. (1998) *Guidance for Industry: Guidance for Somatic Cell Therapy and Gene Therapy*. Center for Biologics Evaluation and Research (CBER), Rockville, MD.
- FDA. (2006) *Guidance for Industry: Gene Therapy Clinical Trials—Observing Subjects for Delayed Adverse Events*, November, 2006. Center for Biologics Evaluation and Research (CBER), Rockville, MD.
- FDA. (2013) *Guidance for Industry: Preclinical Assessment of Investigational Cellular and Gene Therapy Products*, November, 2013. Center for Biologics Evaluation and Research (CBER), Rockville, MD.
- Gad SC. (2015) *Animal Models in Toxicology*, 3rd Ed. CRC Press, Boca Raton, FL.
- Gad SC, Gad-McDonald SE. (2015) *Biomaterials, Medical Devices and Combination Products; Biocompatibility Testing and Safety Assessment*. CRC Press, Boca Raton, FL.
- Galli MC, Serabian M. (2015) *Regulatory Aspects of Gene Therapy and Cell Therapy Products: A Global Perspective*. Springer, Berlin.
- Gunther C, Hauser A, Hass R. (2016) *Advances in Pharmaceutical Cell Therapy*. World Scientific, Hackensack, NJ.
- Japan (Pharmaceutical and Medical Device Agency—PMDA). (2000–2008) *Notification No. 1314 and Ensuring Quality and Safety of Products Derived from Processed Human Cell/Tissue (No. 208003 for autologous products and No. 0912006 for allogenic products)*.
- Sharper ME, Monton G, Rossi A. (2012) Nonclinical safety strategies for stem cell therapies. *Toxicol. Appl. Pharmacol.* 262:223–231.
- Takefman D, Bi L. (2015) Regulatory aspects of gene therapy. In: *Gene and Cell Therapy; Therapeutic Mechanisms and Strategies* (Templeton NS, Ed.), 4th Ed. CRC Press, Boca Raton, FL.
- Templeton NS. (2015) *Gene and Cell Therapy; Therapeutic Mechanisms and Strategies*, 4th Ed. CRC Press, Boca Raton, FL.
- Verdier F, Descotes J. (1999) Preclinical safety evaluation of human gene therapy products. *Toxicol. Sci.* 47:9–15.

APPENDIX A

SELECTED REGULATORY AND TOXICOLOGICAL ACRONYMS

| | | | |
|--------|--|--------|---|
| 510(k) | Premarket notification for change in a device | CVM | Center for Veterinary Medicine (FDA) |
| AALAS | American Association Laboratory Animal Science | DART | Developmental and Reproductive Toxicology |
| AAMI | Association for the Advancement of Medical Instrumentation | DHHS | Department of Health and Human Services |
| ABT | American Board of Toxicology | DIA | Drug Information Associates |
| ACGIH | American Conference of Governmental Industrial Hygienists | DMF | Drug (or device) master file |
| ACT | American College of Toxicology | DSHEA | Dietary Supplement Health and Education Act |
| ADE | Adverse drug event (of drug substances) | EEC | European Economic Community |
| ADI | Allowable Daily Intake | EFPIA | European Federation of Pharmaceutical Industries Association |
| AIDS | Acquired immune deficiency syndrome | EM | Electron microscopy |
| AIMD | Active implantable medical device | EPA | Environmental Protection Agency |
| ANSI | American National Standards Institute | EU | European Union |
| APHIS | Animal and Plant Health Inspection Service | FCA | Freund's complete adjuvant |
| ASTM | American Society for Testing and Materials | FDA | Food and Drug Administration |
| CAS | Chemical Abstract Service | FDCA | Food, Drug, and Cosmetic Act |
| CBER | Center for Biologic Evaluation and Research (FDA) | FDLI | Food and Drug Law Institute |
| CDER | Center for Drug Evaluation and Research (FDA) | FIFRA | Federal Insecticides, Fungicides, and Rodenticides Act |
| CDRH | Center for Devices and Radiological Health (FDA) | GCP | Good clinical practices |
| CFAN | Center for Food and Nutrition (FDA) | GLP | Good laboratory practices |
| CFR | Code of Federal Regulations | GMP | Good manufacturing practices |
| CIIT | Chemical Industries Institute of Toxicology | GPMT | Guinea pig maximization test |
| CPMP | Committee on Proprietary Medicinal Products (United Kingdom) | HEW | Department of Health, Education, and Welfare (no longer existent) |
| CFR | Code of Federal Regulations | HIMA | Health Industry Manufacturer's Association |
| CSE | Control Standard Endotoxin | HSDB | Hazardous Substances Data Bank |
| CSM | Committee on Safety of Medicines (United Kingdom) | IARC | International Agency for Research on Cancer |
| CTC | Clinical trial certificate (United Kingdom) | ICH | International Conference on Harmonization |
| CTX | Clinical trial certificate exemption (United Kingdom) | ID | Intradermal |
| | | IDE | Investigational device exemption |
| | | IND(A) | Investigational new drug application |

| | | | |
|------------------|---|-------|---|
| INN | International nonproprietary names | NLM | National Library of Medicine |
| IP | Intraperitoneal | NOEL | No Observable Effect Level |
| IRAG | Interagency Regulatory Alternatives Group | NOAEL | No Observable Adverse Effect Level |
| IRB | Institutional Review Board | NTP | National Toxicology Program |
| IRLG | Interagency Regulatory Liaison Group | ODE | Office of Device Evaluation |
| ISO | International Standards Organization | OECD | Organization for Economic Cooperation and Development |
| IUD | Intrauterine device | | |
| IV | Intravenous | PDI | Primary dermal irritation |
| JECFA | Joint Expert Committee for Food Additives | PDN | Product development notification |
| JMAFF | Japanese Ministry of Agriculture, Forestry, and Fishery | PEL | Permissible exposure limit |
| JPMA | Japanese Pharmaceutical Manufacturers Association | PhRMA | Pharmaceutical Research and Manufacturers of America |
| LA | Licensing authority (United Kingdom) | PL | Produce license (United Kingdom) |
| LAL | <i>Limulus</i> ameobocyte lysate | PLA | Product license application |
| LD ₅₀ | Lethal dose 50: the dose calculated to kill 50% of a subject population, median lethal dose | PMA | Premarket approval applications |
| LOEL | Lowest-Observed-Effect Level | PO | Per os (orally) |
| LOAEL | Lowest Observed Adverse Effect Level | PTC | Points to consider |
| MAA | Marketing authorization application (EEC) | QAU | Quality Assurance Unit |
| MedDRA | Medical Dictionary for Drug Regulatory Activities | RAC | Recombinant DNA Advisory Committee |
| MCA | Medicines Control Agency | RCRA | Resources Conservation and Recovery Act |
| MD | Medical device | RTECS | Registry of Toxic Effects of Chemical Substances |
| MHW | Ministry of Health & Welfare (Japan) | SARA | Superfund Amendments and Reauthorization Act |
| MID | Maximum implantable dose | SC | Subcutaneous |
| MOE | Margin of exposure | SCE | Sister chromatid exchange |
| MOU | Memorandum of understanding | SNUR | Significant new use rule |
| MRL | Maximum residue limits | SOP | Standard Operating Procedure |
| MSDS | Material safety data sheet | SOT | Society of Toxicology |
| MTD | Maximum tolerated dose | SRM | Standard Reference Materials (Japan) |
| NAS | National Academy of Sciences | STEL | Short-term exposure limit |
| NCTR | National Center for Toxicological Research | TLV | Threshold limit value |
| NDA | New drug application | USAN | United States Adopted Name Council |
| NIH | National Institutes of Health | USDA | United States Department of Agriculture |
| NIOSH | National Institute of Occupational Safety and Health | USEPA | United States Environmental Protection Agency |
| NK | Natural killer | USP | United States Pharmacopeia |
| | | VAERS | Vaccine Adverse Event Reporting System |
| | | VSD | Vaccine Safety Datalink |
| | | WHO | World Health Organization |

APPENDIX B

DEFINITION OF TERMS AND LEXICON OF “CLINICAL” OBSERVATIONS IN NONCLINICAL (ANIMAL) STUDIES

Movement

| | |
|------------------|--|
| Anesthetized | The absence of or reduced response to external stimuli, accompanied with a loss of righting reflex |
| Ataxia | In coordination of muscular action involving locomotion, including loss of coordination and unsteady gait |
| Catalepsy | A condition characterized by a waxy rigidity of the muscles such that the animal tends to remain in any position in which it is placed |
| Hyperactivity | An abnormally high level of motor activity |
| Hypersensitivity | An abnormally strong reaction to external stimuli such as noise or touch |
| Lethargy | A state of deep and prolonged depression stupor from which it is possible to be aroused, followed by an immediate relapse |
| Low carriage | The animal's torso is carried very close to the ground during movement |
| Paralysis | Inhibition or loss of motor function; may be characterized by affected portion of the body |
| Prostrate | Animal assumes a recumbent position due to loss of strength or exhaustion and may show intermittent uncoordinated movements |
| Righting reflex | The ability of an animal, when placed on its back, to regain a position on all fours |
| Unsteady gait | An erratic manner or style of walking |

Respiration

| | |
|---------------------------|---|
| Audible respiration | An abnormal respiratory sound heard while listening to the breathing of the animal (e.g., wheezing and rales) |
| Bradypnea | An abnormal slowness of the respiration rate |
| Cheyne–Stokes respiration | Breathing characterized by rhythmic waning and waxing of the depth of respiration, with regularly recurring periods of apnea: seen especially in coma resulting from affection of the nervous centers |
| Dyspnea | “Shortness of breath,” difficult or labored breathing |
| Gasping | Spasmodic breathing with the mouth open or laborious respiration with the breath caught convulsively |
| Hyperpnea | Deep and rapid breathing |
| Hypopnea | Shallow and slow breathing |

(Continued)

(Continued)

| | |
|----------------------------------|---|
| Irregular respiration | No definite cycle or rate of breathing |
| Labored respiration | Forced or difficult, usually irregular breathing |
| Tachypnea | An excessive rapidity of the respiration rate |
| <i>Condition of skin and fur</i> | |
| Alopecia | Deficiency of hair (baldness) |
| Cyanosis | Visible skin and/or mucous membranes turn dusky blue due to lack of oxygenation of the blood |
| Necrosis | Actual tissue destruction, masses of dead/destroyed tissue |
| <i>Urogenital region</i> | |
| Anuria | An absence of or sharp decline in urine excretion |
| Diarrhea | An abnormal frequency and liquidity of fecal discharge |
| Polyuria | An abnormally sharp increase in the amount of urine excretion |
| <i>Convulsions and tremors</i> | |
| Clonic | This is often seen as a “paddling; motion of the forelegs of the animal” |
| Convulsions | Transient, self-sustaining electrical dysrhythmias which have a tendency to recur. Convulsions are generally associated with a finite period of unconsciousness and have a muscular involvement manifested as disorganized limb movements |
| Fasciculation | Rapid, often continuous contraction of a bundle of skeletal muscle fibers which does not produce a purposeful movement (twitching) |
| Tonic | Muscular contraction, keeping limbs in a fixed position, generally extended to the rear |
| Torsion | Postural in coordination or rolling. This is generally associated with the vestibular (ear canal) system |
| Tremor | Fine oscillating muscular movements which may or may not be rhythmic |
| <i>Condition of eyes</i> | |
| Blepharospasm | A twitching or spasmodic contraction of the orbicularis oculi muscle |
| Chemosis | Edema of conjunctiva(e)—the conjunctival tissue responds to noxious stimuli by swelling |
| Chromodacryorrhea | The response of reddish conjunctival exudate; no blood cells present in exudate (i.e., not true “bloody tears”) |
| Conjunctivitis | Inflammation of conjunctiva (mucous membrane which lines the eyelids and is reflected into the eyeball) |
| Exophthalmos | An abnormal protrusion of the eyeball from the orbit |
| Lacrimation | The secretion of tears |
| Miosis | Constriction of the pupil |
| Mydriasis | Dilation of the pupil |
| Nystagmus | An abnormal involuntary movement of the eyes. It may be rotational or horizontal or vertical plane |
| Ocular exudate | Secretion (usually transparent and yellow) directly from the eye |
| Opacity | A loss of transparency of the eyeball |
| Pinpoint pupils | Ultimate state of miosis |
| Ptois | Refers to a dropping of the upper eyelid thought to be due to impaired conduction in the third cranial nerve |
| <i>Miscellaneous</i> | |
| Analgesia | The absence of (or reduced response to) painful stimuli |
| Hunched posture | The drawing-in of both ends of the body and extremities with a sharp arching of the back |
| Kyphosis | Humpback—an abnormal curvature and dorsal prominence of the vertebrae column |
| Nasal discharge | Fluid secretion from the nostrils |
| Piloerection | Body hair stands on end; dilation of the pupils usually accompanies piloerection |
| Salivation | Excessive secretion of saliva from the mouth |
| Straub tail | Condition, especially in mice, in which the animal carries its tail in an erect (vertical or nearly vertical) position. This sign is commonly associated with chemicals (e.g., morphine) that bind to opiate receptors |

(Continued)*Reflexes*

| | |
|---|---|
| Corneal reflex | Closure of the eyelids in response to a corneal touch (e.g., with a soft brush bristle) |
| Grip strength (or screen grip) | Measure of the grip strength of the forelimbs or hindlimbs; may be evaluated quantitatively or by subjective estimate or impairment (rodents only) |
| Pinna reflex | Twitch of the outer ear in response to a gentle touch |
| Preyer's reflex (auditory startle response) | Involuntary movement of the outer ears produced by an auditory stimulus (especially in rats) |
| Pupillary reflex | Contraction of the pupil in response to light stimulation of the retina |
| Righting reflex | The ability to land on (when dropped) or regain normal stance on all four limbs |
| Startle reflex | Response to sharp sound, touch, or other startling stimulus; response may range from "absent" to "normal" to "hyperreactive," including exaggerated jerking, jumping, frantic attempts to escape, and even convulsion |

APPENDIX C

NOTABLE REGULATORY INTERNET ADDRESSES

| Organization or Publication | Web Address (URL) | Sample Main Topics |
|--|---|--|
| ABPI | http://www.abpi.org.uk/ | |
| Adverse Reactions Bulletin | http://journals.lww.com/adversedruginform/pages/default.aspx | |
| Agency for Toxic Substances and Disease Registry | www.atsdr.cdc.gov | |
| Association of Clinical Biochemists | http://www.leeds.ac.uk/acb/ | Items of general medical interest and an assay finder to help researcher find methods or labs to measure a wide variety of hormones, metals, enzymes, and drugs in body fluids |
| Australian Therapeutic Goods Administration | http://www.tga.gov.au/ | Medical devices; GMP codes; parliamentary secretary's working; status document; party on complementary medicines; medical releases; publications; site map; related sites |
| Canadian Health Protection Board | http://www.hwc.ca/hpb | |
| Canadian Health Protection Branch | http://www.hc-sc.gc.ca/hpb | Medical devices; chemical hazards; food; product safety; Science Advisory Board; diseases; radiation protection; drugs; HPB transition policy, planning, and coordination |
| Centre for Medicines Research | http://www.cmr.org/ | |
| ChemInfo | www.indiana.edu/~cheminfo/ca_csti.html | SirCH: chemical safety or toxicology information |
| Clinical Pharmacology Drug Monograph Service | http://clinicalpharmacology.com/# | |
| ClinWeb | http://www.ohsu.edu/clinweb | Oregon Health Sciences University |
| CNN Interactive (Health) | http://www.cnn.com/HEALTH/index.html | Up-to-date information on health issues including drug safety concerns and withdrawals |

(Continued)

(Continued)

| Organization or Publication | Web Address (URL) | Sample Main Topics |
|--|--|--|
| Code of Federal Register | http://www.access.gpo.gov/nara/cfr/index.html or http://www.access.gpo.gov/su_docs/aces/aces140.html | For proposed rules and regulations |
| Code of Federal Regulations | http://www.access.gpo.gov/nara/cfr/cfr-table-search.html | NARA code sections |
| Commission on Human Medicines (CHM) | https://www.gov.uk/government/organisations/commission-on-human-medicines | |
| Cornell Legal Library | http://www.law.cornell.edu | Code of Federal Regulations; supreme court decisions; US code; circuit courts of appeal |
| DIA home page | http://www.diahome.org | Home page of the Drug Information Association |
| Doctor's guide to the internet | http://www.psigroup.com | |
| Documents for clinical research | http://www.ams.med.uni-goettingen.de/links-en.shtml | <i>Declaration of Helsinki</i> , other documents and collection of related sites |
| Drug InfoNet | http://www.druginfonet.com | |
| EMBASE | https://www.embase.com/login | |
| EPA | https://www3.epa.gov/ | |
| European Medicines Agency (EMA) | http://www.ema.europa.eu/ema/ | |
| Europa | http://europa.eu/index_en.htm | Official website of the European Union |
| European pharmacovigilance research group | http://cordis.europa.eu/project/rcn/30911_en.html | |
| Food and Drug Administration (FDA) | www.fda.gov | Foods; human drugs; biologics; animal drugs; cosmetics; medical devices; radiological health |
| FDA—CBER (Center for Biologics Evaluation and Research) | http://www.fda.gov/cber | |
| CBER what's new | http://www.fda.gov/cber/whatsnew.htm | |
| FDA—CDER (Center for Drug Evaluation and Research) | http://www.fda.gov/cder | |
| FDA adverse events database | http://www.fda.gov/cder/adr | |
| CDER what's new | http://www.fda.gov/cder/whatsnew.htm | |
| FDA—CDRH | www.fda.gov/cdrh/index.html | Home page |
| Search site | www.fda.gov/cdrh/search.html | Search CDRH site |
| Comment | www.fda.gov/cdrh/comment4.html | Comment on CDRH site |
| Device Advice | www.fda.gov/cdrh/devadvice/32.html | |
| PDF reader | www.fda.gov/cdrh/acrobat.html | |
| FDA—CFRAN (Center for Food Safety and Applied Nutrition) | http://www.fda.gov/Food/ | |
| FDA—Center for Toxicological Research | http://www.fda.gov/nctr | |
| FDA—CVM (Center for Veterinary Medicine) | http://www.fda.gov/cvm | |
| FDA—bioengineered food | http://www.fda.gov/oc/biotech/default.htm | |

(Continued)

| Organization or Publication | Web Address (URL) | Sample Main Topics |
|---|--|---|
| FDA—breast implants | http://www.fda.gov/cdrh/breastimplants/index.html | |
| FDA—cosmetics | http://www.fda.gov/Cosmetics/ | |
| FDA—dietary supplements | http://www.fda.gov/Food/DietarySupplements/ | |
| FDA's Electronic Freedom of Information Act | http://www.fda.gov/foi/foia2.htm | |
| FDA—field operations | www.fda.gov/ora/ | What's new; import program; inspectional, science and compliance references; federal/state relations |
| The common technical document for the registration of pharmaceuticals for human use: 08-24-00 | http://www.fda.gov/cder/guidance/4022dfts.htm | |
| Design controls | www.fda.gov/ora/inspect_ref/qsreq/dcrpgd.html | Design control report and guidance text |
| | www.fda.gov/ora/inspect_ref/igs/elec_med_dev/emcl.html | Guide to inspections of electromagnetic compatibility aspects of medical device quality systems text |
| Guide to Inspections of Quality Systems | www.fda.gov/ora/inspect_ref/igs/qsit/qsitguide.htm | QSIT inspection handbook text |
| Guide to Inspections of Quality Systems | www.fda.gov/ora/inspect_ref/igs/qsit/QSITGUIDE.PDF | PDF version of QSIT Inspection Handbook Text |
| Photosafety Testing 07-05-00 | http://www.fda.gov/cder/guidance/3281dft.htm | |
| Skin Irritation and Sensitization Testing of Generic Transdermal Drug Products 06:01:00 | http://www.fda.gov/cder/guidance/2887fml.htm | |
| FDA—MedWatch | http://www.fda.gov/medwatch/ | USFDA drug adverse event reporting system |
| FDA—tampons | http://www.fda.gov/oc/opacpm/topicindexes/tampons.html | |
| Food and Drug Law Institute | http://www.fdli.org | Special interest; publications; multimedia; order products; academic programs; directory of lawyers and consultants; contact us |
| Health On the Net | http://www.hon.ch | |
| Health Information on the Internet | https://wellcome.ac.uk/education-and-learning-newsletter | Bimonthly newsletter from the Wellcome Trust and the RSM |
| International Classification of Disease (ICD)-10 | http://www.who.int/classifications/icd/en/ | |
| International Conference on Harmonization (ICH) 3 home page | http://www.ich.org/home.html | Official ICH website (with guidelines) |
| International Federation of Pharmaceutical Manufacturers and Associations (IFPMA) | http://www.ifpma.org/ | |
| International Federation of Pharmaceutical Manufacturers | http://www.ifpharma.com | ICH documents and postings; international pharmaceutical issues |

(Continued)

(Continued)

| Organization or Publication | Web Address (URL) | Sample Main Topics |
|--|---|---|
| International Pharmaceutical Regulatory Monitor | http://www.fdanews.com/publications/19 | International Pharmaceutical Regulatory Monitor |
| International Society of Pharmacoeconomics | http://www.pharmacoepi.org | |
| JAMA | http://www.ama-assn.org/jama | This gives many other useful US sites |
| Japanese Ministry of Health and Welfare | http://www.mhlw.go.jp/english/ | Organization; Y2K problem; statistics; white paper; related sites |
| Library of Congress | https://www.loc.gov/ | Searchable database of federal legislation, congressional record, and committee information |
| Market and exploitation of research | http://www.cordis.lu | |
| Medicines and Healthcare Products Regulatory Agency (MHRA) | https://www.gov.uk/government/organisations/medicines-and-healthcare-products-regulatory-agency | |
| Medical Research Council | http://www.mrc.ac.uk/ | |
| Medscape | http://www.medscape.com | |
| Multilingual glossary of medical terms | http://users.ugent.be/~rvdstich/eugloss/welcome.html | |
| National Archives and Public Records Administration | http://www.access.gpo.gov/su_docs/aces/aces140.html | Code of Federal Regulations; federal register; laws; US Congress information |
| National Institutes of Health (USA) | http://www.nih.gov | |
| National Library Network | www.toxnet.nlm.nih.gov | TOXNET: Toxicology Data Network, a cluster of databases on toxicology, hazardous chemicals, and related areas |
| National Toxicology Program | http://ntp-server.niehs.nih.gov/ | |
| New Quality System (QS) Regulation | www.fda.gov/bbs/topics/ANSWERS/ANS00763.html | FDA talk paper announcing the GMP final rule text |
| Organized Medical Network Information | http://www.omni.ac.uk | |
| Pharmaceutical and Medical Safety Bureau—Japan | http://www.mhlw.go.jp/english | |
| Pharmainfo.net | http://www.pharmainfo.net/ | Information resource for pharmaceutical and health-related information |
| PubMed | http://www.ncbi.nlm.nih.gov/pubmed | Search literature from MEDLINE, life science journals, and online books |
| Regulatory Affairs Professionals Society (RAPS) | http://www.raps.org | Certificates; resource center; publications; chapters; related links; contacting RAPS |
| Reuters Health Information Services | http://www.reutershealth.com/ | |
| SCRIP: World Pharmaceutical News | https://scrip.pharmamedtechbi.com/ | |
| SNOMED | http://snomed.org | Systemized Nomenclature of Human and Veterinary Medicines |
| Swedish Medical Products Agency | http://www.mpa.se | |
| US Department of Agriculture (USDA) | http://www.usda.gov | |
| Food safety | http://www.foodsafety.gov/ | |

(Continued)

| Organization or Publication | Web Address (URL) | Sample Main Topics |
|--|---|---|
| USDA—FMS (Farm Service Agency) | http://www.fsa.usda.gov/ | |
| USDA—FSA (Food and Nutrition Service) | http://www.fns.usda.gov/fns/ | |
| USDA—FSIS (Food Safety and Inspection Service) | http://www.usda.gov/fsis | |
| US Department of Commerce | https://www.commerce.gov/ | Bureau of Export Administration; International Trade Association; Patent and Trademark; National Institute of Standards and Technology |
| US Pharmacopeia | www.usp.org/prn | |
| University of Pittsburgh | www.pitt.edu | |
| World Health Organization | http://www.who.int | Governance; health topics; information sources; reports; director-general; about WHO; international digest of health; legislation (http://www.who.int/pub/dig.html) |
| WHO Collaborating Centre for International Drug Monitoring | http://www.who-umc.org/ | |

APPENDIX D

GLOSSARY OF TERMS USED IN THE CLINICAL EVALUATION OF THERAPEUTIC AGENTS

Abnormality A sign, symptom, or laboratory result not characteristic of normal individuals.

Adverse event Unwanted effects that occur and are detected in populations. The term is used whether there is or is not any attribution to a medicine or other cause. Adverse events may be known parts of a disease that are observed to occur within a period of observation, and they may be analyzed to test for their frequency in a given population or trial. This is done to determine if there is an unexpectedly increased frequency resulting from nondisease factors such as medicine treatment. The term “adverse event” or “adverse experience” is used to encompass adverse reactions plus any injury, toxicity, or hypersensitivity that may be medicine related, as well as any medical events that are apparently unrelated to medicine that occur during the study (e.g., surgery, illness, and trauma). See definition of **adverse reaction**.

Adverse experience See **adverse event**.

Adverse reaction Unwanted effect(s) (i.e., physical and psychological symptoms and signs) resulting from treatment. A less rigid definition of adverse reaction includes the previous definition plus any undesirable effect or problem that is present during the period of treatment and may or may not be a well-known or obvious complication of the disease itself. Thus, many common personality, physical, psychological, and behavioral characteristics that are observed in medicine studies are sometimes characterized as adverse reactions even if they were present during baseline.

Synonyms of adverse reactions generally include adverse medical effects, untoward effects, side effects, adverse drug experiences, and adverse drug reactions. Specific distinctions among some of these terms may be defined operationally. For example, the term “adverse reaction” is used to denote those signs and symptoms at least possibly related to a medicine, whereas the term “adverse experiences” is used to include nonmedicine-related medical problems in a trial such as those emanating from trauma or concurrent illness. Distinctions among side effects, adverse events, and adverse reactions are illustrated in the definitions of the two former terms.

Bias (i) A point of view that prevents impartial judgment on issues relating to that point of view. Clinical trials attempt to control this through double blinding. (ii) Any tendency for a value to deviate in one direction from the true value. Statisticians attempt to prevent this type of bias by various techniques, including randomization.

Clinical significance The quality of a study’s outcome that convinces physicians to modify or maintain their current practice of medicine. The greater the clinical significance, the greater is the influence on the practice of medicine. The assessment of clinical significance is usually based on the magnitude of the effect observed, the quality of the study that yielded the data, and the probability that the effect is a true one. Although this operational definition is presented from the physician’s perspective, the term could operationally be defined from the patient’s perspective. Patients are primarily concerned with results that will lead to an improved quality of life or a lengthening of their life.

In addition, clinical significance may be applied to either positive data or efficacy or negative safety data such as for adverse reactions. Synonyms include clinical importance, clinical relevance, and clinical meaningfulness.

Clinical studies The class of all scientific approaches to evaluate medical disease preventions, diagnostic techniques, and treatments. Investigational and marketed prescription medicine evaluations plus over-the-counter medicines are included.

Clinical trials A subset of those clinical studies that evaluates investigational medicines in phases I, II, and III. Phase IV evaluations of marketed medicines in formal clinical trials using the same or similar types of protocols to those used in phases I and III are also referred to as clinical trials.

Compliance (i) Adherence of patients to following medical advice and prescriptions. Primarily applied to taking medicine as directed but also applies to following advice on diet, exercise, or other aspects of a patient's life. (ii) Adherence of investigators to following a protocol and related administrative and regulatory responsibilities. (iii) Adherence of sponsors to following regulatory, legal, and other responsibilities and requirements relating to a clinical trial.

Compound A chemical synthesized or prepared from natural sources that is evaluated for its biological activities in preclinical tests.

Development of medicines The term "development" as applied to medicines is used in several different contexts, even within the pharmaceutical industry. This often leads to confusion and misunderstanding. No single definition is preferred, but the particular meaning intended should be made clear by all people using the term. Three operational definitions are presented, from the broadest to the narrowest:

1. All stages and processes involved in discovering, evaluating, and formulating a new medicine until it reaches the market (i.e., commercial sale).
2. All stages involving the evaluation and formulation of a new medicine (after the medicine has been discovered and has gone through preclinical testing) until it reaches the market.
3. Those stages after the preclinical discovery and evaluation that involve technical development. These processes include formulation work, stability testing, scaling-up the compound for larger-scale synthesis, and providing analytical support. Clinical trials are not included in this definition.

Disease Disorders (e.g., anxiety disorders, seizure disorders), conditions (e.g., obesity, menopause), syndromes, specific illnesses, and other medical problems that are an acquired morbid change in a tissue, organ, or organism. Synonyms are illness and sickness.

Dosage regimen (i) The number of doses per given time period (usually days), (ii) the time that elapses between doses (e.g., dose to be given every 6 h) or the time that the doses are to be given (e.g., dose to be given at 8 a.m., noon, and 4 p.m. each day), or (iii) the quantity of a medicine (e.g., number of tablets, capsules, etc.) that are given at each specific time of dosing.

Efficacy A relative concept referring to the ability of a medicine to elicit a beneficial clinical effect. This may be measured or evaluated using objective or subjective parameters and in terms ranging from global impressions to highly precise measurements. Efficacy is assessed at one or more levels of organization (e.g., subcellular, cellular, tissue, organ, whole body) and may be extrapolated to other levels.

End point An indicator measured in a patient or biological sample to assess safety, efficacy, or another trial objective. Some end points are derived from primary end points (e.g., cardiac output is derived from stroke volume and heart rate). Synonyms include outcome, variable, parameter, marker, and measure. See surrogate end point in the text. Also defined as the final trial objective by some authors.

Incidence rate The rate of occurrence of new cases of a disease, adverse reaction, or other event in a given population at risk (e.g., the incidence of disease X is Y patients per year per 100 000 population).

Interpretation The process whereby one determines the clinical meaning or significance of data after the relevant statistical analyses have been performed. These processes often involve developing an explanation of the data that are being evaluated.

Medicine When a compound or substance is tested for biological and clinical activity in humans, it is considered to be a medicine. Some individuals prefer to define a medicine as a compound that has demonstrated clinically useful properties in patients. This definition, however, would restrict the term to use sometime during or after phase II. Others use the term loosely and apply it to compounds with biological properties during the preclinical period that suggest medical usefulness in humans. The author has adopted the first definition for use in this book.

Patient The term "patient" is used almost exclusively throughout this book in preference to subject or volunteer. Patient is used to cover those cases in which the term "volunteer" would be appropriate.

Pharmacodynamics The processes of the body's responses resulting from treatment with a medicine or compound. The processes include pharmacological, biochemical, physiological, and therapeutic effects. The pharmacodynamics of a response to treatment are presented with the scientific and/or clinical language of the disciplines involved in detecting, measuring, and describing the effects.

Pharmacokinetics The processes of absorption, distribution, metabolism, and excretion of compounds and medicines.

Phase I Initial safety trials on a new medicine, usually conducted in normal male volunteers. An attempt is made to establish the dose range tolerated by volunteers for single and for multiple doses. Phase I trials are sometimes conducted in severely ill patients (e.g., in the field of cancer) or in less ill patients when pharmacokinetic issues are addressed (e.g., metabolism of a new antiepileptic medicine in stable epileptic patients whose microsomal liver enzymes have been induced by other antiepileptic medicines). Pharmacokinetic trials are usually considered phase I trials regardless of when they are conducted during a medicine's development.

Phase IIa Pilot clinical trials to evaluate efficacy (and safety) in selected populations of patients with the disease or condition to be treated, diagnosed, or prevented. Objectives may focus on dose–response, type of patient, frequency of dosing, or numerous other characteristics of safety and efficacy.

Phase IIb Well-controlled trials to evaluate efficacy (and safety) in patients with the disease or condition to be treated, diagnosed, or prevented. These clinical trials usually represent the most rigorous demonstration of a medicine's efficacy. Sometimes referred to as pivotal trials.

Phase IIIa Trials conducted after efficacy of the medicine is demonstrated but prior to regulatory submission of a New Drug Application (NDA) or other dossier. These clinical trials are conducted in patient populations for which the medicine is eventually intended. Phase IIIa clinical trials generate additional data on both safety and efficacy in relatively large numbers of patients in both controlled and uncontrolled trials. Clinical trials are also conducted in special groups of patients (e.g., renal failure patients) or under special conditions dictated by the nature of the medicine and disease. These trials often provide much of the information needed for the packaging insert and labeling of the medicine.

Phase IIIb Clinical trials conducted after regulatory submission of an NDA or other dossier but prior to the medicine's approval and launch. These trials may supplement earlier trials, complete earlier trials, or may be directed toward new types of trials (e.g., quality of life, marketing) or phase IV evaluations. This is the period between submission and approval of a regulatory dossier for marketing authorization.

Phase IV Studies or trials conducted after a medicine is marketed to provide additional details about the medicine's efficacy or safety profile. Different formulations, dosages, durations of treatment, medicine interactions, and other medicine comparisons may be evaluated. New age groups, races, and other types of patients can be studied. Detection and definition of previously unknown

or inadequately quantified adverse reactions and related risk factors are an important aspect of many phase IV studies. If a marketed medicine is to be evaluated for another (i.e., new) indication, then those clinical trials are considered phase II clinical trials. The term “postmarketing surveillance” is frequently used to describe those clinical studies in phase IV (i.e., the period following marketing) that are primarily observational or nonexperimental in nature to distinguish them from well-controlled phase IV clinical trials or marketing studies.

Phases of clinical trials and medicine development Four phases of clinical trials and medicine development exist and are defined in the following text. Each of these definitions is a functional one, and the terms are not defined on a strict chronological basis. An investigational medicine is often evaluated in two or more phases simultaneously in different clinical trials. Also, some clinical trials may overlap two different phases.

Prevalence The total number of people in a population that are affected with a particular disease at a given time. This term is expressed as the rate of all cases (e.g., the prevalence of disease X is Y patients per 100 000 population) at a given point or period of time.

Research (on medicines) Numerous definitions of research are used both in the literature and among scientists. In the broadest sense, research in the pharmaceutical industry includes all processes of medicine discovery, preclinical and clinical evaluation, and technical development. In a more restricted sense, research concentrates on the preclinical discovery phase, where the basic characteristics of a new medicine are determined. Once a decision is reached to study the medicine in humans to evaluate its therapeutic potential, the compound passes from the research to the development phase.

Research and development When research and development are used together, it refers to the broadest definition for research (see previous text). Some people use the term “research” colloquially to include most or all of the scientific and medical areas (discovery, evaluation, and development) covered by the single term “research and development.” “Medicine development” has several definitions and, in its broadest definition, is exactly the same as the broad definition of “research.”

Risk A measure of (i) the probability of occurrence of harm to human health or (ii) the severity of harm that may occur. Such a measure includes judgment of the acceptability of risk. Assessment of safety involves judgment, and there are numerous perspectives (e.g., patients, physicians, company, regulatory authorities) used for judging it.

Safety A relative concept referring to the freedom from harm or damage resulting from adverse reactions or physical, psychological, or behavioral abnormalities that occur as a result of medicine or nonmedicine use. Safety

is usually measured with one or more of the following: physical examination (e.g., vital signs, neurological, ophthalmological, general physical), laboratory evaluations of biological samples (e.g., hematology, clinical chemistry, urinalysis), special tests and procedures (e.g., electrocardiogram, pulmonary function tests), psychiatric tests and evaluations, and determination of clinical signs and symptoms.

Serious adverse reactions Multiple definitions are possible, and no single one is correct in all situations. In general usage referring to patients in clinical trials, a serious adverse reaction may be (i) any bad adverse reaction that is observed, (ii) any bad adverse reaction that one does not expect to observe, (iii) any bad adverse reaction that one does not expect to observe and is not in the label, or (iv) any bad adverse reaction that has not been reported with standard therapy. Definitions also may be based on the degree to which an adverse reaction compromises a patient's function or requires treatment.

Side effect Any effect other than the primary intended effect(s) resulting from medicine or nonmedicine treatment or intervention. Side effects may be negative (i.e., an adverse reaction), neutral, or positive (i.e., a beneficial effect) for the patient. This term, therefore, includes all adverse reactions plus other effects of treatment. See definition of **adverse reaction**.

Site This refers to the place where a clinical trial is conducted. A physician who has offices and sees patients in three separate locations is viewed as having one site.

A physician who is on the staff of four hospitals could be viewed as having one or four sites, depending on how similar or different the patient populations are and whether the data from these four locations will be pooled and considered a single site. For example, a single physician who enrolls groups of patients at a university hospital, private clinic, community hospital, and Veterans Administration Hospital should generally be viewed as having four sites, since the patient populations would be expected to differ at each site. See **satellite site**.

Statistical significance This term relates to the probability that an event or difference occurred by chance alone. Thus, it is a measure of whether a difference is likely to be real, but it does not indicate whether the difference is small or large, important or trivial. The level of statistical significance depends on the number of patients studied or observations made, as well as the magnitude of difference observed.

Therapeutic window This term is applied to the difference between the minimum and maximum doses that may be given to patients to obtain an adequate clinical response and avoid intolerable toxic effects. The greater the value calculated for the therapeutic window, the greater a medicine's margin of safety. Synonyms are "therapeutic ratio" and "therapeutic index."

Volunteer A normal individual who participates in a clinical trial for reasons other than medical need and who does not receive any direct medical benefit from participating in the trial.

APPENDIX E

COMMON VEHICLES FOR THE NONCLINICAL EVALUATION OF THERAPEUTIC AGENTS

TABLE E.1 Index of Vehicles and Excipients: Codex and Details

| Data in Table #s (T#) | | CAS# | Synonyms | Formula | Key Toxicity Review Articles/Sources | Animal Species Evaluated | Administration Routes Evaluated |
|-------------------------------|---------------------------------|------------|--|--|---|---|---------------------------------------|
| Excipient/Vehicle | Combination #s (C: #) | | | | | | |
| Acacia | T2, T111 C:1-7 | 9000-01-5 | <i>Acaciae gummi</i> | Natural product | Anderson (1986), Bachmann et al. (1978), and TOXNET (2015a) | Nonhuman primate, rat | PO |
| Acetate buffer | T111 C:8, 9, 119, 138-141 | 71-50-1 | Acetate ion | C ₂ H ₃ O ₂ | TOXNET (2015b) | Nonhuman primate, minipig, mouse, rat | IM, IV, PO |
| Acetic acid | T3, T111 C:10, 12 | 64-19-7 | Ethanoic acid | C ₂ H ₄ O ₂ | Schonwald (2004a), Szilagyi (2012a), and TOXNET (2015c) | Dog, mouse, rat | IV, PO |
| Acetone | T4, T111 C:11 | 67-64-1 | 2-Propanone | C ₃ H ₆ O | TOXNET (2015d) | Guinea pig, mouse, rabbit, rat | Dermal, PO |
| Acetonitrile | T111 C:12 | 75-05-8 | | C ₂ H ₃ N | TOXNET (2015e) | Dog | PO |
| Acetylated lanolin alcohol | T111 C:174 | 61788-49-6 | | N/A | TOXNET (2015f) | Minipig | Topical |
| Acetylmethylamide | T5 — | 79-16-3 | <i>N</i> -Methylacetamide | C ₃ H ₇ NO | TOXNET (2015g) | Nonhuman primate | PO |
| Alcohol denatured SDA | T111 C:234 | | | | | Minipig | Topical |
| Alginate acid | T6 — | 9005-32-7 | Norgine | | JECFA (1997) and TOXNET (2015h) | Rat | IP |
| Anecortave acetate | T7 — | 7753-60-8 | | C ₂₃ H ₃₀ OS | Jockovich et al. (2006), Talsma (2004), and TOXNET (2015i) | Rat | SC |
| Antifoam 1510-US™ | T111 C:1, 100, 101 | | Silicone emulsion | | Dow Corning (2015a) | Rat | PO |
| Avicel CL-611™ | T8 — | | Microcrystalline cellulose and carboxymethylcellulose sodium, NF, Ph. Eur | | TOXNET (2015j) | Dog | PO |

| | | | | | | |
|-----------------------|------------------------------|------------------------------------|-------------------------------------|-----------------------|---|-----------------------------------|
| Balanced salt saline | T9 — | | | | Rabbit | Intravitreal |
| Basal salt solution | T10 — | | | | Mouse | Subretinal injection |
| Benzoic acid | T11 — | 65-85-0 | Benzoic acid | $C_7H_6O_2$ | Rat | PO |
| | | | | | David et al. (2012), Nair (2001), and TOXNET (2015k) | |
| Benzyl alcohol | T111 C: 8, 9, 13, 14, 31 | 100-51-6 | | C_7H_8O | Nair (2001) and TOXNET (2015l) | Cat, dog, nonhuman primate, rat |
| | | | | | | |
| β -Cyclodextrin | T12 — | 7585-39-9 | β -Dextrin; betadex | $C_{42}H_{70}O_{35}$ | Albers and Muller (1995), Challa et al. (2005), Martin et al. (1998), Rajewski et al. (1995), TOXNET (2015m), Toyoda et al. (1997), and Waner et al. (1995) | Dog, mouse, nonhuman primate, rat |
| | | | | | | IP, IV, PO |
| BHA | T111 C: 73 | 25013-16-5 | Butylated hydroxyanisole | $C_{11}H_{16}O_2$ | TOXNET (2015n) | Minipig |
| | | | | | | Topical |
| BHT | T111 C: 73, 127, 164, 165 | 128-37-0 | Butylated hydroxytoluene | $C_{15}H_{24}O$ | Briggs et al. (1989), Lanigan and Yamarik (2002a), Nakagawa et al. (1984), and TOXNET (2015o) | Cat, dog, minipig, mouse, rat |
| | | | | | | IV, PO, topical |
| Bicarbonate buffer | T13 — | 71-52-3 | Bicarbonate ion | $C-H-O_3$ | TOXNET (2015p) | Mouse |
| | | | | | | PO |
| Calcium chloride | T14 — | 10035-04-8 | | $Ca-Cl_2 \cdot 2H_2O$ | TOXNET (2015q) | Mouse |
| | | | | | | IV, SC |
| Canola oil | T15 — | 120962-03-0 | Canbra oil | Natural product | Evangelista et al. (2004) and TOXNET (2015r) | Dog |
| | | | | | | PO |
| Capmul MCM™ | T111 C: 15 | 26402-22-2, 26402-26-6 | Medium-chain mono- and diglycerides | N/A | Susananta et al. (1995) | Nonhuman primate |
| | | | | | | PO |
| Capmul MCM NF™ | T111 C: 16, 255 | 91744-32-0, 26402-22-2, 26402-26-6 | Glyceryl caprylate/caprate | N/A | | Dog, rat |
| | | | | | | PO |

(Continued)

TABLE E.1 (Continued)

| Excipient/Vehicle | Data in Table #s (T#) | | CAS# | Synonyms | Formula | Key Toxicity Review Articles/Sources | Animal Species Evaluated | Administration Routes Evaluated |
|--------------------------------|-----------------------|--|-------------|---|---|--|--|---------------------------------|
| | Combination #s (C: #) | | | | | | | |
| Capmul PG 8™ | T111 | | | | | | Dog, rat | PO |
| | C:242, 243 | | | | | | | |
| Capryol 90™ | T16 | | 31565-12-5 | Propylene glycol monocaprylate (type II) NF; Capmul PG-8™ | C ₁₁ H ₂₂ O ₃ | Li et al. (2005) and Cho and Gwak (2004) | Dog, rabbit, rat | Dermal, ocular, PO |
| | — | | | | | | | |
| Captisol™ | T17, T111 | | 182410-00-0 | β-Cyclodextrin sulfobutyl ether, sodium salt (CDSBE) | C ₄₂ H ₇₀ - <i>n</i> O ₃₅ ·(C ₄ H ₈ SO ₃ Na) <i>n</i> | Albers and Muller (1995), Challa et al. (2005), Martin et al. (1998), and TOXNET (2015s) | Dog, mouse, nonhuman primate, rat, | IV, PO, SC |
| | C:17, 191 | | | | | | | |
| Carbomer 974P™ | T111 | | 151687-96-6 | Carbomer homopolymer type B (allyl pentaerythritol crosslinked) | N/A | TOXNET (2015t) | Minipig, mouse | Topical |
| | C:234, 235, 272 | | | | | | | |
| Carbopol Ultrez 10™ | T111 | | 195739-91-4 | | N/A | TOXNET (2015u) | Minipig | Topical |
| | C:61 | | | | | | | |
| Carboxymethylcellulose | T18, T111 | | 9000-11-7 | CMC; acetic acid; 2,3,4,5,6-pentahydroxyhexanal | N/A | Gupta et al. (1996), Mehman (2012a), and TOXNET (2015v) | Dog, minipig, mouse, nonhuman primate, rabbit, rat | IA, PO, SC |
| | C:20–30, 301 | | | | | | | |
| Carboxymethylcellulose calcium | T19 | | 9050_04_8 | Calcium CMC; carmellose calcium | N/A | TOXNET (2015w) | Dog | PO |
| | — | | | | | | | |
| Carboxymethylcellulose sodium | T20, T111 | | 9004-32-4 | Sodium CMC; carmellose sodium | N/A | Bachmann et al. (1978), Bar et al. (1995), Cavender (2012d), Freeman et al. (2003), and TOXNET (2015x) | Mouse, rabbit, rat | PO |
| | C:269, 270 | | | | | | | |
| Cavisol™ | T111 | | | | | | Dog, rat | PO |
| | C:185, 186 | | | | | | | |
| Cetostearyl alcohol | T111 | | 67762-27-0 | | N/A | TOXNET (2015y) | Minipig | Topical |
| | C:288 | | | | | | | |
| Cetyl acetate | T111 | | 629-70-9 | | C ₁₈ -H ₃₆ -O ₂ | TOXNET (2015z) | Minipig | Topical |
| | C:174 | | | | | | | |

| | | | | | | | |
|--------------------|--|------------|--|------------------------|---|---|--|
| Cetyl alcohol | T21, T111 C:163 | 36653-82-4 | Hexadecan-1-ol; 1-hexadecanol | $C_{16}H_{34}O_1$ | Bevan (2012c) and TOXNET (2015aa) | Minipig, mouse | IP, topical |
| Citrate buffer | T22, T111 C:18, 19, 126, 147, 229 | 77-92-9 | Sodium citrate/citric acid buffer | $C_6H_8O_7$ | Schonwald (2004b) | Dog, nonhuman primate, rat | IV, PO, SC |
| Citric acid buffer | T23, T111 C:13, 38, 58, 95, 302 | 77-92-9 | | $C_6H_8O_7 \cdot H_2O$ | Szilagyi (2012b) | Cat, mouse, nonhuman primate, rat | PO, topical |
| Coconut oil | T111 C:288 | 8001-31-8 | | Natural product | National Toxicology Program (2001), Shadnia et al. (2005), and TOXNET (2015ab) | Minipig | Topical |
| Collagen matrix | T24 — | 9007-34-5 | Collagen human | Natural product | McCarthy et al. (2002) and Clark et al. (1989) | Nonhuman primate, rabbit | Implantation in humerus bone, implantation |
| Corn oil | T25, T111 C:31–33 | 8001-30-7 | Corn germ oil, glyceric oil | Natural product | DeWitt et al. (2005), Dupont et al. (1990), TOXNET (2015ac), and Wu et al. (2004) | Chicken embryo, dog, mouse, nonhuman primate, rabbit, rat | Injection into egg, PO |
| Cottonseed oil | T26 — | 8001-29-4 | | Natural product | TOXNET (2015ad) | Dog | SC |
| Cyclodextrin | T111 C:34–37 | 12619-70-4 | | | TOXNET (2015ae) | Rat | PO |
| Cyclohexane | T27, T111 C:11 | 110-82-7 | Hexahydrobenzene; hexamethylene; hexanaphthene | C_6H_{12} | Gad (2014), Kreckmann et al. (2000), Malley et al. (2000) and TOXNET (2015af) | Rabbit, rat | PO, dermal |
| Cyclomethicone NF | T111 C:174, 288 | 69430-24-6 | | N/A | TOXNET (2015ag) | Minipig | Topical |
| DAM | T111 C:179–182 | 57-71-6 | Diacetylmonoxime; 2,3-butanedione 2-oxime | $C_4H_7N \cdot O_2$ | TOXNET (2015ah) | Dog, rat | IV |
| Dextrose | T28, T111 C:17, 39, 40, 108–111, 184, 277 | 50-99-7 | Glucose; D-glucose, anhydrous; dextrosol | $C_6H_{12}O_6$ | Buard et al. (2003), Robertson et al. (2003), and TOXNET (2015ai) | Cat, dog, minipig, nonhuman primate, rabbit, rat | IV, IV/PO, perivascular, PO, SC |

(Continued)

TABLE E.1 (Continued)

| Data in Table #s (T#) | | | | | Administration Routes Evaluated | | |
|--|--|---------------------------------|---|--|--|---|--------------------------------------|
| Excipient/Vehicle | Combination #s (C : #) | CAS# | Synonyms | Formula | | Key Toxicity Review Articles/Sources | Animal Species Evaluated |
| Dichlorvos | T29 — | 62-73-7 | DDVP; dichlorophos; dichlorphos; divipan; 2,2- dichloroethenyl dimethyl phosphate | C ₄ H ₇ Cl ₂ O ₄ P | TOXNET (2015aj) | Nonhuman primate | IV |
| Diethylacetamide | T30, T111 C:41 | 685-91-6 | N,N-Diethylacetamide | C ₆ -H ₁₃ -N-O | Budden et al. (1978), Caujolle et al. (1970), ChemIDPlus (2015), and TOXNET (2015ak) | Cat, chicken, dog, mouse, rabbit, rat | IP, IV |
| Diethylene glycol monoethyl ether | T31 | 111-90-0 | 2-(2-Ethoxyethoxy)ethanol; DEGEE; carbitol | C ₆ -H ₁₄ -O ₃ | Hardin (1983), Hardin et al. (1984), and TOXNET (2015al) | Nonhuman primate | IV |
| Dimethicone | T111 C:301 | 9006-65-9 | | N/A | TOXNET (2015am) | Rabbit | PO |
| Dimethiconol blend 20 | T111 C:38 | 70131-67-8 and 63148-62-9 | Dimethicone and dimethiconol | | Dow Corning (2015b) | Minipig | Topical |
| Dimethylacetamide | T32, T111 C:187–188 | 127-19-5 | DMA; N,N- dimethylacetamide; acetdimethylamide | C ₄ -H ₉ -N-O | Leadscope Data Portal (2008a) and TOXNET (2015an) | Chicken, dog, mouse, rabbit, rat | Dermal, inhalation, IP, IV, PO |
| Dimethyl sulfoxide | T33, T111 C:2–5, 43–57, 91, 94, 157, 158, 193–197 | 67-68-5 | DMSO | C ₂ H ₆ OS | Ali (2001), Augustine et al. (2000), Bartsch et al. (1976), Pestel et al. (2006), PharmPK (2002a), RTECS (2015), Ruble et al. (2006), Schonwald (2004a), Sodicoff et al. (1990), TOXNET (2015ao), White et al. (1984), and Wood et al. (1971) | Dog, guinea pig, minipig, mouse, nonhuman primate, rabbit, rat | IP, IV, PO, SC |
| Disodium hydrogen phosphate dihydrate | T111 C:42 | 10140-65-5 | Sodium phosphate, dibasic | H ₃ -O ₄ -P- <i>x</i> - H ₂ -O-2Na | TOXNET (2015ap) | Nonhuman primate | SC |
| Docosanol | T111 C:288 | 661-19-8 | 1-Docosanol | C ₂₂ -H ₄₆ -O | TOXNET (2015aq) | Minipig | Topical |

| | | | | | |
|--------------------------|---|--|----------------------|--|---|
| Dulbecco's modified PBS | T34 — | Dulbecco's modified phosphate-buffered saline solution | | Rat | IV (tail vein) |
| Dulbecco's PBS | T35 — | Dulbecco's phosphate-buffered saline solution | | Rat | PO |
| EDTA | T111 C: 58–61 | Ethylenediaminetetraacetic acid; edetic acid | $C_{10}H_{16}N_2O_8$ | Cavender (2012e), Heimbach et al. (2000), Lanigan and Yamarik (2002c), and TOXNET (2015ar) | Minipig, nonhuman primate, rat Dermal, IV, topical |
| Ethanol | T36, T111 C: 32, 33, 38, 40, 61–73, 98, 99, 108–111, 124, 148–156, 160, 161, 164, 165, 175, 184, 191, 196–205, 211, 212, 223, 224, 239–250, 253, 254, 259, 260, 283, 285, 286, 290 | Ethyl alcohol | C_2H_6O | Bartsch et al. (1976), Bevan (2012a), Church and Witting (1997), Fort et al. (1984), Moorman et al. (2001), Rowe et al. (2012), Ruble et al. (2006), Sivilotti (2004), and TOXNET (2015as) | Dog, guinea pig, minipig, mouse, nonhuman primate, rat Dermal, IP, IV, PO, SC, topical |
| Fumaric acid | T111 C: 90 | 2-Butenedioic acid | $C_4H_4O_4$ | TOXNET (2015at) | Dog PO |
| Gelatin | T37, T111 C: 74, 120 | | | TOXNET (2015au) | Dog, nonhuman primate PO |
| Gelatin phosphate buffer | T28 — | | | Minipig | Topical |
| Gelucire™ | T111 C: 75 | | | Minipig | PO |
| Gelucire 44/14™ | T39, T111 C: 76, 77 | PEG-32 glyceryl laurate | | Cavender (2012b), Dordunoos et al. (1996), Kawakami et al. (2004), Ratsimbazafy et al. (1999), TOXNET (2015av), and Working et al. (1997) | Dog, minipig, mouse rabbit, rat Dermal, ocular, PO |

(Continued)

TABLE E.1 (Continued)

| Excipient/Vehicle | Data in Table #s (T#) | | CAS# | Synonyms | Formula | Key Toxicity Review Articles/Sources | Animal Species Evaluated | Administration Routes Evaluated |
|---------------------------------------|---------------------------|-------------|---------------------|----------|--|---|--|---------------------------------------|
| | Combination #s (C : #) | | | | | | | |
| Gelucire 50/13™ | T40 | 121548-05-8 | G-50-13 | | | Fini et al. (2005), Passerini et al. (2002), Ratsimbazafy et al. (1999), Gattefossé (2004), Sharma et al. (2006), TOXNET (2015aw), Gattefossé (2004), | Rat | PO |
| | — | | | | | | | |
| Gluconic acid | T41 | 133-42-6 | | | C ₆ -H ₁₂ -O ₇ | TOXNET (2015ax) | Dog, rat | PO |
| | — | | | | | | | |
| Glycerol | T42, T111 | 56-81-5 | Glycerine; glycerin | | C ₃ H ₈ O ₃ | Anderson et al. (1950), Bartsch et al. (1976), Cosmetic Ingredient Review (2004), and TOXNET (2015ay) | Dog, guinea pig, minipig, mouse, nonhuman primate, rabbit, rat | IP, IV, PO, SC, topical |
| | C : 61, 90, 239, 240 | | | | | | | |
| Glyceryl stearate SE | T111 | 11099-07-3 | | | C ₁₈ -H ₃₆ -O ₂ -x- C ₃ -H ₈ -O ₃ | Cosmetic Ingredient Review (2004) | Minipig | Topical |
| | C : 163 | | | | | | | |
| Glycofurool | T43 | 31692-85-0 | | | (C ₂ -H ₄ -O) mult- C ₅ -H ₁₀ -O ₂ | Ruble et al. (2006) and TOXNET (2015az) | Dog | IV |
| | — | | | | | | | |
| Glycol dimethacrylate crosspolymer | T111 | | | | | | Rat | Topical |
| | C : 127 | | | | | | | |
| Gum tragacanth | T44 | 9000-65-1 | | | Natural product | Anderson (1989), Bachmann et al. (1978), Hagiwara et al. (1992), and TOXNET (2015ba) | Mouse | PO |
| | — | | | | | | | |
| Gum xanthan | T45, T111 | 11138-66-2 | Keltrol | | (C ₃₅ H ₄₉ O ₂₉) _n | TOXNET (2015bb) | Minipig, rabbit, rat | PO, topical |
| | C : 163, 305 | | | | | | | |
| Hexylene glycol | T111 | 107-41-5 | | | C ₆ -H ₁₄ -O ₂ | | Minipig | Topical |
| | C : 38 | | | | | | | |
| Histidine | T111 | 71-00-1 | | | C ₆ -H ₉ -N ₃ -O ₂ | TOXNET (2015bc) | Dog, rat | IV, SC |
| | C : 78-80, 287 | | | | | | | |

| | | | | | | | |
|---|---------------------------|-------------|--|-----------------------|--|--|----------------------------|
| Hydrochloric acid | T46 — | 7647-01-0 | Hydrogen chloride; muriatic acid; chlorhydric acid; chlorine | HCl | TOXNET (2015bd) | Dog, rat | PO |
| Hydrogenated castor oil | T111 C:288 | 8001-78-3 | | | TOXNET (2015be) | Minipig | Topical |
| Hydroxyethylcellulose | T47, T111 C:100, 101 | 9004-62-0 | Natrosol; 2-hydroxyethyl ether cellulose | | Pestel et al. (2006) and TOXNET (2015bf) | Rat | PO |
| Hydroxypropyl β -cyclodextrin | T48, T111 C:91–99 | 128446-35-5 | 2-Hydroxypropyl- β -cyclodextrin; HP β CD; hydroxypropyl betadex; Cavasol W7 HP™ | $C_{54}H_{102}O_{39}$ | Albers and Muller (1995), Challa et al. (2005), Coussement et al. (1990), Gerloczy et al. (1994), Gould and Scott (2005), Leadscope Data Portal (2008b), Martin et al. (1998), Pestel et al. (2006), Ruble et al. (2006), Stella and He (2008), Thackaberry et al. (2009), and TOXNET (2015bg) | Dog, mouse, nonhuman primate, rabbit, rat | Intranasal, IP, IV, PO, SC |
| Hydroxypropyl cellulose | T49, T111 C:38 | 9004-64-2 | Hyprolose (INN), 2-hydroxypropylcellulose | N/A | Cavender (2012c) and TOXNET (2015bh) | Minipig, rat | PO, topical |
| Hydroxypropyl methylcellulose | T50, T111 C:81–87, 90 | 9004-65-3 | Benecel MHPc, hypromellose, HPMC; Methocel E50 Premium LV™ | N/A | Feitoza et al. (2003), Geerling et al. (2001), Maki et al. (2000), Mehlman (2012b), Obara et al. (1992), Rosen et al. (1987), Thackaberry et al. (2009), and TOXNET (2015bi) | Dog, minipig, mouse, nonhuman primate, rat | IP, PO |
| Hydroxypropyl methylcellulose acetate succinate | T111 C:88, 89, 291–294 | 71138-97-1 | Hypromellose acetate succinate; HPMC AS | N/A | TOXNET (2015bj) | Dog, minipig, mouse, rabbit, rat | PO |
| Hymetellose | T111 C:102 | 9032-42-2 | Methyl hydroxyethylcellulose; Tylose MH50 | N/A | TOXNET (2015bk) | Rabbit | PO |
| Hypotonic PBS | T51 — | | | | | Dog, rat | IV |
| Imwitor 742™ | T111 C:103–105 | | Caprylic/capric glycerides; glycerol monocaprylocaprate, type I | | Susananta et al. (1995) | Hamster, nonhuman primate, rat | PO |

(Continued)

TABLE E.1 (Continued)

| Data in Table #s (T#) | | | | | | | Administration Routes Evaluated |
|---|---|------------|--|--|--|-----------------------------------|---------------------------------|
| Excipient/Vehicle | Combination #s (C: #) | CAS# | Synonyms | Formula | Key Toxicity Review Articles/Sources | Animal Species Evaluated | |
| Isopropyl alcohol | T52, T111 C:14 | 67-63-0 | 2-Propanol; isopropanol; sec-propyl alcohol | C ₃ H ₈ O | Allen et al. (1998), Bevan (2012b), Burleigh-Flayer et al. (1998), Church and Witting (1997), Sivilotti (2004), TOXNET (2015bl), and Tyl et al. (1994) | Dog, rabbit | Dermal, topical |
| Isopropyl myristate | T53, T111 C:163, 234, 235 | 110-27-0 | 1-Methylethyl tetradecanoate; Crodamol IPM | C ₁₇ H ₃₄ O ₂ | Campbell and Bruce (1981), Komatsu et al. (1979), and TOXNET (2015bm) | Minipig, mouse, rabbit | Dermal, topical |
| Isotonic saline | T111 C:117 | | | | | Rat | PO |
| Kolliphor™ | T111 C:43, 48 | | Cremophor; polyoxyl castor oil | | | Nonhuman primate | IV, PO |
| Kolliphor EL™ | T54, T111 C:15, 16, 62, 63, 106, 108–111, 170–173, 209 | 61791-12-6 | Cremophor EL™; polyoxyl castor oil; polyoxyl 35 castor oil | N/A | Gelderblom et al. (2001), Gupta et al. (1996), Lorenz et al. (1977), PharmPK (2002b), Ramadan et al. (2001), Stokes et al. (2013), and TOXNET (2015bn) | Dog, mouse, nonhuman primate, rat | IV, PO |
| Kolliphor ELP™ | T55 — | 61791-12-6 | Cremophor EL™; polyoxyl castor oil; polyoxyl 35 castor oil | N/A | TOXNET (2015bn) and Xcess Biosciences, Inc. (n.d.) | Dog | |
| Kolliphor HS15™ (see also Solutol HS15) | T111 C:113 | | | | | Minipig | PO |
| Kolliphor RH40™ | T56, T111 C:107, 207–209, 231 | 61788-85-0 | Cremophor RH40™; PEG-40 hydrogenated castor oil; polyoxyl 40 hydrogenated castor oil; macrogolglycerol hydroxystearate | N/A | Gupta et al. (1996), Stokes et al. (2013), and TOXNET (2015bo) | Dog, rat | IV, PO |
| Labrafil M1944™ | T57, T111 C:112, 114 | 62563-68-2 | Labrafil™ | N/A | Beckwith-Hall et al. (2002), Gattefossé (2003a) and TOXNET (2015bp) | Dog, rabbit, rat | Dermal, PO |

| | | | | | | | |
|--------------------|---|------------------------|---|--|--|--|------------------------|
| Labrasol™ | T58, T111 C:113–115, 209 | 85536-07-8 | Polyglycolized glycerides | N/A | Hu et al. (2001) and Hu et al. (2002) | Dog, minipig, rabbit, rat | Dermal, IV, ocular, PO |
| Lactated Ringer's | T111 C:116 | 8026-79-7 | Sodium chloride, sodium lactate, potassium chloride and calcium chloride; compound sodium lactate injection | $C_3H_6O_3 \cdot CaCl_2 \cdot Cl \cdot K \cdot Na$ | TOXNET (2015bq) | Nonhuman primate | IV |
| Lactic acid | T111 C:296–298 | 50-21-5 | 2-Hydroxypropanoic acid | $C_3H_6O_3$ | TOXNET (2015br) | Dog, rat | IA, IV |
| Lactose | T59, T111 C:261–264, 278–281 | 63-42-3 (anhydrous) | O-β-D-Galactopyranosyl-(1->4)-α-D-glucopyranose | $C_{12}H_{22}O_{11}$ (anhydrous) | Ahmad et al. (2004), Baldrick and Bamford (1997), and TOXNET (2015abs) | Dog, nonhuman primate, rat | Inhalation, IV, SC |
| Lanolin | T60 — | 8006-54-0 | Wool wax | N/A | Kligman (1983) and TOXNET (2015bt) | Minipig, rabbit | Dermal, topical |
| Lanolin alcohol NF | T111 C:174 | 8027-33-6 | Eucerin | Natural product | TOXNET (2015bu) | Minipig | Topical |
| L-Arginine HCl USP | T111 C:163, 289 | 15595-35-4 | L-Arginine hydrochloride | $C_6H_{14}N_4O_2 \cdot x \cdot Cl \cdot H$ | TOXNET (2015bv) | Minipig, nonhuman primate | SC, topical |
| L-Ascorbic acid | T111 C:117 | 50-81-7 | Cevatine, cevex, cevital | $C_6H_8O_6$ | Bendich and Cohen (1990), Dykes and Meier (1975), Temple (2004), and TOXNET (2015bw) | Rat | PO |
| Lauroglycol 90™ | T61 — | 27194-74-7 | Propylene glycol monolaurate; lauric acid, monoester with propane-1,2-diol | $C_{15}H_{30}O_3$ | Bartsch et al. (1976), Liu et al. (2006a), and TOXNET (2015bx) | Rabbit, rat | Dermal, ocular, PO |
| Maltitol solution | T62 — | 9053-46-7 | Liquid maltitol; Lycasin | $C_{12}H_{24}O_{11} + C_6H_{14}O_6$ | Modderman (1993), Walker and El Harith (1978), and TOXNET (2015by) | Rat | IP |
| Maltol | T63 — | 118-71-8 | 2-Methyl pyrromeconic acid; 2-methyl-3-hydroxy-4-pyrone | $C_6H_6O_3$ | Hironishi et al. (1996), Murakami et al. (2006), and TOXNET (2015bz) | Guinea pig, rabbit | PO |
| Mannitol | T64, T111 C:43, 48, 74, 118–123, 159, 233, 274, 275, 282 | 69-65-8 | D-Mannitol | $C_6H_{14}O_6$ | Horvath et al. (1982), Lina et al. (1996), and TOXNET (2015ca) | Dog, minipig, mouse, nonhuman primate, rabbit, rat | IM, IV, PO, SC |

(Continued)

TABLE E.1 (Continued)

| Excipient/Vehicle | Data in Table #s (T#) | | CAS# | Synonyms | Formula | Key Toxicity Review Articles/Sources | Animal Species Evaluated | Administration Routes Evaluated |
|----------------------|--------------------------|--|------------|--|--|---|---|---------------------------------------|
| | Combination #s (C: #) | | | | | | | |
| Methanesulfonic acid | T111 C:124 | | 75-75-2 | Methylsulfonic acid | CH ₃ O ₃ S | Shertzer (2012) and TOXNET (2015cb) | Rat | PO |
| Methocel™ | T111 C:125 | | | | | | Nonhuman primate, rat | PO |
| Methyl methacrylate | T111 C:127 | | 80-62-6 | | C ₅ -H ₈ -O ₂ | TOXNET (2015cc) | Rat | Topical |
| Methylcellulose | T65, T111 | | 9004-67-5 | Cellulose methyl ether; Methocel A4M Premium | N/A | Bachmann et al. (1978), Gupta et al. (1996), Mehlman (2012c), Sellers et al. (2005), and TOXNET (2015cd) | Dog, guinea pig, mouse, nonhuman primate, rabbit, rat | IV, PO, topical |
| Methylparaben | T111 C:128, 147, 174 | | 99-76-3 | 4-Hydroxybenzoic acid, methyl ester | C ₈ -H ₈ -O ₃ | TOXNET (2015ce) | Minipig | SC, topical |
| Methylpyrrolidone | T66, T111 | | 872-50-4 | N-Methyl-2-pyrrolidone, 1- methyl-2-pyrrolidone, Pharmasolv, NMP | C ₅ H ₉ NO | Bartsch et al. (1976), Kennedy (2012), Lee et al. (1987), Ruble et al. (2006), Solomon et al. (1996), and TOXNET (2015cf) | Dog, minipig, mouse | IV, PO |
| Mineral oil | T67, T111 C:288 | | 8012-95-1 | Liquid paraffin | Natural product | Carlton et al. (2001), Dalbey and Biles (2003), Nash et al. (1996), TOXNET (2015cg), and Trimmer et al. (2004) | Cat, dog, guinea pig, minipig, mouse, rat | PO, topical |
| Myristyl alcohol | T111 C:288 | | 112-72-1 | 1-Tetradecanol | C ₁₄ -H ₃₀ -O | TOXNET (2015ch) | Minipig | Topical |
| Neobee 1053 Oil™ | T111 C:164, 165 | | 73398-61-5 | Medium-chain triglycerides | N/A | Bellantone et al. (1999), Traul et al. (2000), Susananta et al. (1995), and Wieland et al. (1993) | Mouse, rat | IV, PO |
| Octoxynol-40 | T111 C:167–169 | | 9002-93-1 | | (C ₂ -H ₄ -O) mult-C ₁₄ ⁻ H ₂₂ -O (C ₂ -H ₄ -O) mult-C ₁₄ -H ₂₂ -O C ₃₄ -H ₆₂ -O ₁₁ | TOXNET (2015ci) | Dog, rabbit | Ocular |

| | | | | | | | |
|----------------------|----------------------------------|------------|---|-----------------------------|---|--|-----------------------|
| Oleic acid NF | T111 C:163, 170–174 | 112-80-1 | 9-Octadecenoic acid | $C_{18}-H_{34}-O_2$ | TOXNET (2015cj) | Minipig | Topical |
| Oleyl alcohol NF | T111 C:174 | 143-28-2 | | $C_{18}-H_{36}-O$ | TOXNET (2015ck) | Minipig | Topical |
| Olive oil | T68, T111 C:174 | 8001-25-0 | | Natural product | Evangelista et al. (2004) and TOXNET (2015cl) | Minipig, rat | PO, topical |
| Ora-Plus™ suspension | T111 C:34–37 | | Purified water, microcrystalline cellulose, carboxymethylcellulose sodium, xanthan gum, carrageenan, calcium sulfate, trisodium phosphate, citric acid and sodium phosphate as buffers, dimethicone antifoam emulsion. Preserved with methylparaben and potassium sorbate | | Paddock Laboratories (2015) | Rat | PO |
| Panthenol | T111 C:61 | 16485-10-2 | Dexpanthenol | $C_9-H_{19}-N-O_4$ | TOXNET (2015cm) | Minipig | Topical |
| Peanut oil | T69, T111 C:175 | 8002-03-7 | Arachis oil; Fletcher's | Natural product | Cosmetic Ingredient Review (2001), Patel et al. (2005), and TOXNET (2015cn) | Dog, rat | PO, SC |
| Peccol | T111 C:176, 177 | 25496-72-4 | Glycerol monooleate NF; monoolein | $C_{21}-H_{40}-O_4$ | TOXNET (2015co) | Dog, rat | PO |
| PEG 200 | T70, T111 C:131, 183, 184 | 25322-68-3 | Polyethylene glycol 200 | (C_2-H_4-O) mult- H_2-O | Cavender (2012b), Dordunoos et al. (1996), Smyth et al. (1945), Smyth et al. (1950), Smyth et al. (1955), Quadbeck (1950), TOXNET (2015cp), and Working et al. (1997) | Minipig, nonhuman primate, rabbit, rat | IP, IV, PO |
| PEG 300 | T71, T111 C:106, 185–190, 290 | 25322-68-3 | Polyethylene glycol 300 | (C_2-H_4-O) mult- H_2-O | Carpenter and Shaffer (1952), Cavender (2012b), Dordunoos et al. (1996), Patel et al. (2005), Rowe and Wolf (1982), Smyth et al. (1945), Smyth et al. (1950), Smyth et al. (1955), TOXNET (2015cp), and Working et al. (1997) | Cat, dog, guinea pig, mouse, rabbit, rat | IP, IV, PO, PO mucosa |

(Continued)

TABLE E.1 (Continued)

| Data in Table #s (T#) | | | | Administration Routes Evaluated | | | |
|--------------------------|---|------------|--------------------------|---|--|--|-----------------------------|
| Excipient/Vehicle | Combination #s (C: #) | CAS# | Synonyms | | | | |
| PEG 400 | T72, T111 C:44–47, 50–52, 75–77, 107, 115, 151, 152, 170–173, 176, 177, 191–231, 254, 276 | 25322-68-3 | Polyethylene glycol 400 | (C ₂ -H ₄ -O) mult-H ₂ -O | Key Toxicity Review Articles/Sources | Animal Species Evaluated | |
| | | | | | Bartsch et al. (1976), Cavender (2012b), Dordunoos et al. (1996), Fort et al. (1984), Gupta et al. (1996), Gutiérrez-Cabano (2000), Hermansky et al. (1995), Li et al. (2011), Patel et al. (2005), Rowe and Wolf (1982), Ruble et al. 2006), Shideman and Procita (1951), Smyth et al. (1945), Smyth et al. (1950), Smyth et al. (1955), Stokes et al. (2013), Strickley (2008), Thackaberry (2014), TOXNET (2015cp), and Working et al. (1997) | Dog, guinea pig, minipig, mouse, nonhuman primate, minipig, rabbit rat | Dermal, IP, IV, PO, topical |
| PEG 600 | T73 — | 25322-68-3 | Polyethylene glycol 600 | (C ₂ -H ₄ -O) mult-H ₂ -O | Pfordt (1971), Rowe and Wolf (1982), Smyth et al. (1955), and TOXNET (2015cp) | Rat | IP, IV, PO |
| PEG 810 | T74 — | 25322-68-3 | Polyethylene glycol 810 | (C ₂ -H ₄ -O) mult-H ₂ -O | Kärber (1951) and TOXNET (2015cp) | Rat | IV, SC |
| PEG 1000 | T75 — | 25322-68-3 | Polyethylene glycol 1000 | (C ₂ -H ₄ -O) mult-H ₂ -O | Shideman and Procita (1951), Smyth et al. (1950), and TOXNET (2015cp) | Mouse, rabbit, rat | IP, IV, PO |
| PEG 1500 | T76 — | 25322-68-3 | Polyethylene glycol 1500 | (C ₂ -H ₄ -O) mult-H ₂ -O | Rowe and Wolf (1982), Smyth et al. (1947), Smyth et al. (1950), Smyth et al. (1942), Smyth et al. (1955), and TOXNET (2015cp) | Rat | IP, IV, PO |
| PEG 1540 | T77 — | 25322-68-3 | Polyethylene glycol 1540 | (C ₂ -H ₄ -O) mult-H ₂ -O | Smyth et al. (1950), Smyth et al. (1955), and TOXNET (2015cp) | Dog, rabbit, rat | IP, IV, PO |

| | | | | | | | |
|-------------------------------|-------------------------|------------|---|---|--|--|----------------------------|
| PEG 4000 | T78, T111 C:53, 54 | 25322-68-3 | Polyethylene glycol 4000 | (C ₂ -H ₄ -O) mult-H ₂ -O | Rowe and Wolf (1982), Shideman and Procita (1951), Smyth et al. (1942), Smyth et al. (1947), Smyth et al. (1950), Smyth et al. (1955), and TOXNET (2015cp) | Dog, mouse, rabbit, rat | IP, IV, PO |
| PEG 6000 | T79 — | 25322-68-3 | Polyethylene glycol 6000 | (C ₂ -H ₄ -O) mult-H ₂ -O | Smyth et al. (1950), Smyth et al. (1955), and TOXNET (2015cp) | Rabbit, rat | IP, IV, PO |
| PEG 10000 | T80 — | 25322-68-3 | Polyethylene glycol 10000 | (C ₂ -H ₄ -O) mult-H ₂ -O | Smyth et al. (1950) and TOXNET (2015cp) | Rat | IP, PO |
| PEG 400000 | T81 — | 25322-68-3 | Polyethylene glycol 400000 | (C ₂ -H ₄ -O) mult-H ₂ -O | Smyth et al. (1950), Smyth et al. (1970), and TOXNET (2015cp) | Rat | IV, PO |
| Petrolatum | T82 — | 8009-03-8 | Yellow soft paraffin; petroleum jelly | N/A | TOXNET (2015cq) | Rabbit | Dermal |
| Phenoxyethanol | T111 C:234, 235, 272 | 122-99-6 | 2-Phenoxyethanol | C ₈ -H ₁₀ -O ₂ | TOXNET (2015cr) | Minipig, mouse | Topical |
| Phosal 53 MCT™ | T111 C:231 | | Lecithin in caprylic/capric triglycerides, alcohol, glyceryl stearate, oleic acid, and ascorbyl palmitate | | American Lecithin (2007) and Susananta et al. (1995) | Rat | PO |
| Phosphate | T111 C:59, 60, 232 | 14265-44-2 | Phosphate ion | O ₄ -P | TOXNET (2015cs) | Nonhuman primate | IV |
| Phosphate buffer | T111 C:233 | | | | | Rat | SC |
| Phosphate- buffered saline | T83, T111 C:28–30 | | PBS | | | Dog, minipig, mouse, nonhuman primate, rabbit, rat | IA, IV, PO, SC, topical |
| Polawax | T111 C:234, 235 | | Emulsifying wax | | Carlton et al. (2001) | Minipig, mouse | Topical |
| Poloxamer 124™ | T111 C:225–228, 231 | | | | | Dog, mouse, rat | PO |

(Continued)

TABLE E.1 (Continued)

| Excipient/Vehicle | Data in Table #s (T#) | | CAS# | Synonyms | Formula | Key Toxicity Review Articles/Sources | Animal Species Evaluated | Administration Routes Evaluated |
|--|--|------------|---|---|--|--|---------------------------------|---------------------------------|
| | Combination #s (C: #) | | | | | | | |
| Poloxamer 188™ | T84, T111 | 9003-11-6 | Poloxalene | (C ₃ -H ₆ -O-C ₂ -H ₄ -O) _x | Benita (1999), Curry et al. (2004), Frim et al. (2004), Grindel et al. (2002), Lemieux et al. (2000), Serbest et al. (2005), and TOXNET (2015ct) | Dog, minipig, mouse, nonhuman primate, rabbit, rat | PO, SC | |
| Polyethylene terephthalate | T111 | | PET | | | Rat | IV | |
| | C:162 | | | | | | | |
| Poly(glycolide-co-DL-lactide) microspheres | T85 | 26780-50-7 | | | | Dog | Into periodontal pockets | |
| | — | | | | | | | |
| Polyglyceryl oleate | T86 | 9007-48-1 | 1,2,3-Propanetriol, homopolymer, (Z)-9-octadecenoate; decaglyceryl monooleate | C ₁₈ -H ₃₄ -O ₂ - _x -(C ₃ H ₈ -O ₃) _x | Gattefossé (2003b) and TOXNET (2015cu) | Rabbit, rat | Dermal, ocular, PO | |
| | — | | | | | | | |
| Polyvinylpyrrolidone | T87, T111 | 9003-39-8 | Povidone; PVP; PVP K30 | (C ₆ -H ₉ -N-O) _x | Beji et al. (2006), PharmPK (2002b), and TOXNET (2015cv) | Dog, nonhuman primate, rabbit, rat | IM, PO | |
| | C:216, 217, 256–258, 291–293 | | | | | | | |
| Potassium chloride | T111 | 7447-40-7 | KCl | KCl | TOXNET (2015cw) | Minipig | Topical | |
| | C:163 | | | | | | | |
| Propylene glycol | T88, T111 | 57-55-6 | 1,2-Dihydroxypropane | C ₃ H ₈ O ₂ | Cavender (2012a), Fort et al. (1984), Ruble et al. (2006), Thackaberry et al. (2009), and TOXNET (2015bx) | Dog, minipig, guinea pig, nonhuman primate, mouse, rabbit, rat | Dermal, IP, IV, PO, SC, topical | |
| | C:16, 61, 65–69, 73, 98, 99, 107, 148, 153–156, 163, 190, 212–215, 234, 235, 239–255 | | | | | | | |
| Propylene glycol dicaprylate/dicaprate | T111 | 68583-51-7 | Caprylic, capric acid, propylene glycol diester | C ₁₀ -H ₂₀ -O ₂ -C ₈ -H ₁₆ -O ₂ -C ₃ -H ₈ -O ₂ | Cosmetic Ingredient Review (1999) and TOXNET (2015cx) | Rat | Topical | |
| | C:127 | | | | | | | |
| Propylparaben | T111 | 94-13-3 | 4-Hydroxybenzoic acid, propyl ester | C ₁₀ -H ₁₂ -O ₃ | TOXNET (2015cy) | Minipig | SC, topical | |
| | C:128, 147, 174 | | | | | | | |
| PVP VA 64 | T111 | 25086-89-9 | Vinylpyrrolidone-vinyl acetate copolymer; copovidone; PVP VA | (C ₆ -H ₉ -N-O-C ₄ -H ₆ -O ₂) _x | TOXNET (2015cz) | Dog, rat | PO | |
| | C:223, 224, 294 | | | | | | | |

| RAMEB | T89 — | Randomly methylated- β - cyclodextrins | Challa et al. (2005) | Nonhuman primate | Intranasal |
|-------------------------------------|---|--|---|---|--|
| Safflower oil | T90 — | <i>Carthamus tinctorius</i> oil | TOXNET (2015da) | Dog | SC |
| Salicylic acid | T111 C:61 | Benzoic acid, 2-hydroxy- | TOXNET (2015db) | Minipig | Topical |
| Saline (pH adjusted, pH 4.5) | T91 — | | | Mouse | IM |
| Sesame oil | T92, T111 C:259, 260 | Gingilli oil; gingelly oil; til oil | Farber et al. (1976), Genovese et al. (1999), Prasanthi et al. (2005), and TOXNET (2015dc) | Dog, mouse, rabbit, rat | PO |
| Shea butter | T111 C:174 | | TOXNET (2015dd) | Minipig | Topical |
| Simethicone | T111 C:126, 143 | Silicone antifoam agent S 184; gas-X | TOXNET (2015de) | Rat | PO |
| Sodium acetate | T111 C:92, 93, 261–265, 299, 300 | Acetic acid sodium salt | TOXNET (2015df) | Dog, mouse, nonhuman primate, rat | IV, PO, SC |
| Sodium acetate trihydrate buffer | T93, T111 C:266, 267 | | | Mouse, nonhuman primate | IV, PO |
| Sodium chloride | T94, T111 C:41, 49, 58–60, 64, 80, 96, 97, 117, 121, 148–163, 166, 189, 210, 211, 232, 236, 237, 251–253, 265, 268, 271, 273, 278–281, 285, 289 | Salt; saline; halite | Barrie et al. (2012), Caraccio and McGuigan (2004), Meneely et al. (1953), Meneely and Ball (1958), and Moore et al. (1988) | Cat, dog, minipig, mouse, nonhuman primate, rabbit, rat | ID, IP, IM, IV, ocular, perivascular, PO, SC, topical |
| Sodium citrate | T111 C:251–258, 268, 294, 303 | | TOXNET (2015dg) | Nonhuman primate, rat | IV, PO |

(Continued)

TABLE E.1 (Continued)

| Excipient/Vehicle | Data in Table #s (T#) | | CAS# | Synonyms | Formula | Key Toxicity Review Articles/Sources | Animal Species Evaluated | Administration Routes Evaluated |
|--|---|--|------------|---|---|--|---------------------------------------|---------------------------------|
| | Combination #s (C: #) | | | | | | | |
| Sodium dihydrogen phosphate dihydrate | T95, T111 C: 42 | | 13472-35-0 | Sodium phosphate monobasic dihydrate; SDPD | $\text{H}_2\text{-O}_4\text{-P-Na-2H}_2\text{-O}$ | TOXNET (2015dh) | Mouse, nonhuman primate | PO, SC |
| Sodium hydroxide | T111 C: 147, 163, 234, 235, 271, 272, 277, 304 | | 1310-73-2 | Caustic soda | NaOH | TOXNET (2015di) | Minipig, mouse, nonhuman primate, rat | IV, SC, topical |
| Sodium lauryl sulfate | T111 C: 142–146 | | 151-21-3 | Sodium dodecyl sulfate | $\text{C}_{12}\text{-H}_{26}\text{-O}_4\text{-S-Na}$ | TOXNET (2015dj) | Dog, rat | PO |
| Sodium metabisulfite | T96 — | | 7681-57-4 | | | TOXNET (2015dk) | Mouse, nonhuman primate, rat | PO |
| Sodium methylparaben | T111 C: 269, 270 | | 5026-62-0 | Benzoic acid, 4-hydroxy-, methyl ester, sodium salt | $\text{C}_8\text{-H}_7\text{-Na-O}_3$ | TOXNET (2015dl) | Minipig, mouse, rat | PO, topical |
| Sodium phosphate buffer | T97, T111 C: 121, 273–277, 289 | | 7558-80-7 | | | Jefferson (2012), Moore et al. (1988), and TOXNET (2015dm) | Dog, mouse, nonhuman primate, rat | IV, PO |
| Sodium propylparaben | T111 C: 269, 270 | | 35285-69-9 | 4-Hydroxybenzoic acid, propyl ester, sodium salt | $\text{C}_{10}\text{-H}_{11}\text{-Na-O}_3$ | TOXNET (2015dn) | Mouse, rat | PO |
| Sodium succinate | T111 C: 118, 122, 123, 278–282 | | 150-90-3 | Succinic acid sodium salt; succinic acid; disodium butanedioate | $\text{C}_4\text{H}_4\text{Na}_2\text{O}_4$ | Szilagyi (2012c) and TOXNET (2015do) | Dog, mouse, rat | IV, SC |
| Sodium sulfite | T98 — | | 7757-83-7 | Sulfurous acid, disodium salt | $\text{H}_2\text{-O}_3\text{-S-2Na}$ $\text{O}_3\text{-S-2Na}$ | TOXNET (2015dp) | Rabbit | Ocular (topical) |
| Solutol HS15™ (see also Kolliphor HS15™) | T99, T111 C: 55, 70, 71, 218–222, 283–286 | | 61909-81-7 | Polyethylene glycol-15-hydroxystearate; polyethylene glycol 660 hydroxystearate | $(\text{C}_2\text{-H}_4\text{-O})_{\text{mult}}\text{-C}_{18}\text{-H}_{36}\text{-O}_3$ | Cavender (2012b), Coon et al. (1991), Dordunoos et al. (1996), Ruchatz (1998), Stokes et al. (2013), and TOXNET (2015dq) | Dog, mouse, nonhuman primate, rat | IV, IP, PO, any |
| Sorbitan tristearate | T111 C: 174 | | 26658-19-5 | Sorbitan, trioctadecanoate | $\text{C}_{60}\text{-H}_{114}\text{-O}_8$ | Lanigan and Yamarik (2002b) and TOXNET (2015dr) | Minipig | Topical |

| | | | | | | | |
|--|-------------------------------------|-------------|---|---------------------------|--|---|--|
| Sorbitol | T111 C:287, 299, 300 | 50-70-4 | D-Sorbitol | $C_6H_{14}O_6$ | TOXNET (2015ds) | Dog, nonhuman primate, rat | IV |
| Soybean oil | T100, T111 C:288 | 8001-22-7 | | Natural product | Cambridge MedChem Consulting (2012), Earl et al. (2002), Farber et al. (1976), Kawashima et al. (2009), and TOXNET (2015dt) | Dog, minipig, rat | PO, topical |
| Squalane NF | T111 C:163 | 111-02-4 | | $C_{30}H_{50}$ | TOXNET (2015du) | Minipig | Topical |
| Stearic acid | T111 C:288 | 57-11-4 | <i>n</i> -Octadecanoic acid | $C_{18}H_{36}O_2$ | TOXNET (2015dv) | Minipig | Topical |
| Stearyl alcohol | T111 C:288 | 112-92-5 | 1-Octadecanol | $C_{18}H_{38}O$ | TOXNET (2015dw) | Minipig | Topical |
| Sucrose | T111 C:42, 78, 79, 274, 275, 289 | 57-50-1 | Sugar | $C_{12}H_{22}O_{11}$ | TOXNET (2015dx) | Cat, dog, nonhuman primate, rabbit, rat | IV, PO mucosa, PO, SC |
| Sucrose acetate isobutyrate | T111 C:290 | 27216-37-1 | SAIB | $C_{40}H_{62}O_{19}$ | TOXNET (2015dy) | Cat | PO mucosa |
| Sulfobutyl ether β -cyclodextrin | T101 — | 182410-00-0 | SBECD | N/A | Albers and Muller (1995), Challa et al. (2005), Kim et al. (1998), Martin et al. (1998), TOXNET (2015dz), and Ueda et al. (1998) | Mouse | PO |
| Tartaric acid | T102 — | 87-69-4 | D-Tartaric acid; 2,3-dihydroxybutanedioic acid | $C_4H_6O_6$ | Sourkes and Koppanyi (1950), Szilagyi (2012d), and TOXNET (2015ea) | Rabbit, rat | PO |
| Terbinafine HCL™ placebo nail lacquer | T103 — | 78628-80-5 | Terbinafine hydrochloride | $C_{21}H_{25}NCl \cdot H$ | TOXNET (2015eb) | Pig | Dermal |
| Tetraglycol | T111 C:56, 157, 158 | 15826-19-4 | | | TOXNET (2015ec) | Minipig, rat | IV |
| Transcutol™ | T104, T111 C:75-77, 113, 114 | 111-90-0 | Diethylene glycol monoethyl ether; DEGEE; 2-(2-ethoxyethoxy)ethanol | $C_6H_{14}O_3$ | Liu et al. (2006b), Sullivan et al. (2014), and TOXNET (2015al) | Cat, dog, guinea pig, minipig, mouse, rabbit, rat | Dermal, IM, IP, IV, inhalation, ocular, PO, SC |

(Continued)

TABLE E.1 (Continued)

| Excipient/Vehicle | Data in Table #s (T#) | | CAS# | Synonyms | Formula | Key Toxicity Review Articles/Sources | Animal Species Evaluated | Administration Routes Evaluated |
|-----------------------------|-----------------------|-----------|---|--|---|---|---|---------------------------------|
| | Combination #s (C: #) | | | | | | | |
| Trehalose | T111 | 99-20-7 | α -D-Trehalose | $C_{12}H_{22}O_{11}$ | TOXNET (2015ed) | Dog, rat | IA, IV | |
| Tris buffer | T111 | 77-86-1 | Tromethamine | $C_4H_{11}N-O_3$ | TOXNET (2015ee) | Minipig, nonhuman primate | IV, topical | |
| Trisodium citrate dihydrate | T105 | 6132-04-3 | Trisodium citrate; sodium citrate | $C_6H_5Na_3O_7 \cdot 2H_2O$ | TOXNET (2015dg) | Dog, hamster, mouse, rat | PO | |
| Tween 20™ | T106, T111 | 9005-64-5 | Polysorbate 20 NF | N/A | Bartsch et al. (1976) and TOXNET (2015ef) | Dog, minipig, mouse, nonhuman primate, rat | IV/SC/IP, IV, PO, SC, topical | |
| Tween 80™ | T107, T111 | 9005-65-6 | Polysorbate 80; Armotan pmo-20, Tween 80™; polyoxyethylene (20) sorbitan monooleate | N/A | Daher et al. (2003), Fisherman and Cohen (1974), Gelperina et al. (2002), National Toxicology Program (1992), O'Sullivan et al. (2004), Oz et al. (2004), Sellers et al. (2005), Thackaberry et al. (2009), and TOXNET (2015eg) | Dog, guinea pig, hamster, minipig, mouse, nonhuman primate, rabbit, rat | IA, IV/SC/IP, intranasal, IV, IP, PO, SC, topical | |
| Vitamin E | T111 | 59-02-9 | α -Tocopherol | $C_{29}H_{50}O_2$ | TOXNET (2015eh) | Dog, rabbit, rat | Ocular, PO | |
| Vitamin E TPGS | T108, T111 | 9002-96-4 | Tocophersolan (USAN) | $(C_2H_4-O)_n$ mult- $C_{33}H_{54}-O_5$ $(C_2H_4-O)_n$ - $C_{33}H_{54}-O_5$ | TOXNET (2015ei) | Dog, mouse, rabbit, rat | PO | |

| | | | | | |
|--|---|-----------|---|--|--|
| Water (Note: Only combination formulations expressly listing water in their ingredients are listed here) | T109, T111 | 7732-18-5 | H ₂ O | Dog, guinea pig, minipig, mouse, nonhuman primate, pig, rabbit, rat | Dermal, IA/SC, IP, IV, PO, SC, topical |
| | C:1, 17, 23-27, 42, 53-55, 58, 61-63, 65-72, 74, 80, 86, 87, 90-94, 98-101, 108-111, 118, 121, 122-125, 138-141, 143-146, 163, 174, 190, 191, 200-205, 212-215, 234, 235, 241-250, 261-264, 273-275, 278-283, 286, 287, 289, 291-295, 301, 304, 305 | | | | |
| White wax | T111 C: 288 | 8012-89-3 | Beeswax | Minipig | Topical |
| Xylitol | T110 — | 87-99-0 | Xylite | Nonhuman primate | Intranasal |
| | | | C ₅ H ₁₂ O ₅ | Carlton et al. (2001) and TOXNET (2015ej) Takahashi et al. (1999) and TOXNET (2015ek) | |

TABLE E.2 Acacia

| Species | Route | Duration | Dose | Adverse Reactions/Toxicity | Notes |
|------------------|-------|----------|---|--|----------------------------------|
| Nonhuman primate | PO | Efficacy | 100 mg kg ⁻¹ day ⁻¹ | Well tolerated; some weight loss, reduction in food intake | Arabic gum; 3% solution in water |
| Rat | PO | 1 month | 500 mg kg ⁻¹ | Well tolerated | |
| | PO | 90 days | 10 mL kg ⁻¹ | Well tolerated | 20% solution |

TABLE E.3 Acetic Acid

| Species | Route | Duration | Dose | Adverse Reactions/Toxicity | Notes |
|---------|-------------|----------|---|----------------------------|-----------------------------|
| Mouse | PO | 1 month | 5 mL kg ⁻¹ | Well tolerated | 15% solution |
| | PO (gavage) | 90 days | 5 mL kg ⁻¹ | Well tolerated | 3% solution |
| Rat | PO | Acute | 5 mL kg ⁻¹ | Well tolerated | 10% solution |
| | PO | 1 month | 10 mL kg ⁻¹ | Well tolerated | 20% solution |
| | PO (gavage) | 90 days | 5 mL kg ⁻¹ day ⁻¹ | Well tolerated | 3% in purified water (92/8) |

TABLE E.4 Acetone

| Species | Route | Duration | Dose | Adverse Reactions/Toxicity | Notes |
|------------|--------|----------|--|---|--|
| Guinea pig | Dermal | 1 month | 1 mL | Well tolerated | |
| Mouse | Dermal | 2 years | 0.5 mL | Well tolerated | |
| | PO | 2 weeks | 3 mL kg ⁻¹ | Higher doses cause acidosis; transient neurobehavioral effects at this dose | |
| Rabbit | Dermal | 90 days | 1 mL | Defatting of application site | |
| Rat | Dermal | 30 days | 5 mL kg ⁻¹ | Well tolerated | |
| | Dermal | 90 days | 1.5 mL kg ⁻¹ , 6 h daily, 5 days per week | Well tolerated | Sham treatment group included, vehicle similar to sham treatment; 100% acetone; age 60 days; ♂/♀ |
| | PO | 2 weeks | 5 mL kg ⁻¹ | Higher doses cause acidosis; transient neurobehavioral effects at this dose | |

TABLE E.5 Acetylmethylamide

| Species | Route | Duration | Dose | Adverse Reactions/Toxicity | Notes |
|------------------|-------|----------------|------|----------------------------|----------|
| Nonhuman primate | PO | 1 month (ADME) | | Well tolerated | In water |

TABLE E.6 Alginic Acid

| Species | Route | Duration | Dose (mg kg ⁻¹) | Adverse Reactions/Toxicity | Notes |
|---------|-------|----------|-----------------------------|----------------------------|-------|
| Rat | IP | 1 month | 100 | Well tolerated | |

TABLE E.7 Anecortave Acetate

| Species | Route | Duration | Dose (mL kg ⁻¹) | Adverse Reactions/Toxicity | Notes |
|---------|------------|----------|-----------------------------|----------------------------|-------|
| Rat | SC (bolus) | 4 doses | 2 | Well tolerated | |

TABLE E.8 Avicel CL-611

| Species | Route | Duration | Dose (mL kg ⁻¹) | Adverse Reactions/Toxicity | Notes |
|---------|-------------|-------------|-----------------------------|----------------------------|--|
| Dog | PO (gavage) | Single dose | 1 | Soft feces | 2.4% in sterile water; age 5 months; ♂/♀ |

TABLE E.9 Balanced Salt Saline

| Species | Route | Duration | Dose | Adverse Reactions/Toxicity | Notes |
|---------|--------------|----------|--------------------|----------------------------|------------------------------|
| Rabbit | Intravitreal | 43 days | 50 µL per eye q14d | None | Non-GLP; age 5 months; 2♂/2♀ |

TABLE E.10 Basal Salt Solution

| Species | Route | Duration | Dose (µL) | Adverse Reactions/Toxicity | Notes |
|---------|----------------------|-----------------|-----------|----------------------------|-----------------------------|
| Mouse | Subretinal injection | SD for 9 months | 2.0 | None | Non-GLP; age 5–7 weeks; 44♂ |

TABLE E.11 Benzoic Acid

| Species | Route | Duration | Dose (mg kg ⁻¹) | Adverse Reactions/Toxicity | Notes |
|---------|-------|----------|-----------------------------|----------------------------|-------|
| Rat | PO | Acute | 100 | Well tolerated | |

TABLE E.12 β-Cyclodextrin

| Species | Route | Duration | Dose | Adverse Reactions/Toxicity | Notes |
|------------------|-------|-----------|-------------------------|--|-------|
| Dog | IP | 1 month | 50 mg kg ⁻¹ | Well tolerated | |
| | IV | 1 month | 100 mg kg ⁻¹ | Well tolerated | |
| | PO | 1 month | 200 mg kg ⁻¹ | Well tolerated | |
| Mouse | IP | 1 month | 10 mg kg ⁻¹ | Well tolerated | |
| Nonhuman primate | PO | 12 months | | Tubular hypertrophy at doses above 100 mg kg ⁻¹ day ⁻¹ at 3 months or longer | |
| Rat | IV | ≥3 months | | Tubular hypertrophy at doses above 100 mg kg ⁻¹ day ⁻¹ at 3 months or longer | |
| | PO | 12 months | 500 g kg ⁻¹ | Hepatitis, nephrosis, acute tubular necrosis at dose levels above 20 g kg ⁻¹ | |

TABLE E.13 Bicarbonate Buffer, pH 9.5

| Species | Route | Duration | Dose (mL kg ⁻¹) | Adverse Reactions/Toxicity | Notes |
|---------|-------------|-----------------|-----------------------------|----------------------------|---------------------|
| Mouse | PO (gavage) | QD for 2 months | 10 | None | Age 8–10 weeks; ♂/♀ |

TABLE E.14 Calcium Chloride

| Species | Route | Duration | Dose (mL kg ⁻¹) | Adverse Reactions/Toxicity | Notes |
|---------|-------|-------------------------|-----------------------------|----------------------------|-------------------------|
| Mouse | IV | 1× per week for 4 weeks | 1.61 | None | 0.5 M; age 5 weeks; ♂/♀ |
| | SC | 1× per week for 4 weeks | 1.61 | None | 0.5 M; age 5 weeks; ♂/♀ |

TABLE E.15 Canola Oil

| Species | Route | Duration | Dose (mL kg ⁻¹) | Adverse Reactions/Toxicity | Notes |
|---------|-------|----------|-----------------------------|----------------------------|-------|
| Dog | PO | 1 month | 2 | Well tolerated | |

TABLE E.16 Capryol 90™

| Species | Route | Duration | Dose | Adverse Reactions/Toxicity | Notes |
|---------|--------|----------|---|--|-------|
| Dog | PO | 28 days | 1000 mg kg ⁻¹ | Nontoxic | |
| | PO | 28 days | 2500 mg kg ⁻¹ | Nontoxic | |
| Rabbit | Dermal | Acute | No dilution | Mildly irritant | |
| | Ocular | Acute | No dilution | Moderately irritant | |
| Rat | PO | Acute | | Nontoxic LD ₅₀ > 5 g kg ⁻¹ | |
| | PO | 7 days | 300, 1000, and 2500 mg kg ⁻¹ | Well tolerated | |
| | PO | 28 days | 500, 1500, and 2500 mg kg ⁻¹ | NOAEL of 2500 mg kg ⁻¹ | |

TABLE E.17 Captisol™

| Species | Route | Duration | Dose | Adverse Reactions/Toxicity | Notes |
|------------------|---------------|---------------------------|--------------------------|----------------------------|--------------------------------------|
| Dog | PO (gavage) | 2× per week for 28 days | 5 mL kg ⁻¹ | None | 15% in DI water; age 5–6 months; ♂/♀ |
| | IV (bolus) | | 1 mL kg ⁻¹ | Well tolerated | 12% solution in water pH 3–11 |
| | IV (infusion) | | 2 mL kg ⁻¹ | Well tolerated | 12% solution in water pH 3–11 |
| Mouse | PO | 1 month | 500 mg kg ⁻¹ | Well tolerated | 10% solution |
| | SC | 90 days | 1200 mg kg ⁻¹ | NOEL | |
| | SC | 6 months | 1200 mg kg ⁻¹ | NOAEL | |
| Nonhuman primate | PO | 9 months | 1 g kg ⁻¹ | Well tolerated | 10% solution |
| | SC | 3× per week for 12 months | 120 mg kg ⁻¹ | Well tolerated | |
| Rat | IV | 1 month | 4 mL kg ⁻¹ | Well tolerated | 12% in water |
| | IV (bolus) | | 2 mL kg ⁻¹ | Well tolerated | 12% solution in water pH 3–11 |
| | IV (infusion) | | 5 mL kg ⁻¹ | Well tolerated | 12% solution in water pH 3–11 |
| | PO | 1 month | 10 mL kg ⁻¹ | Well tolerated | 12% in water |
| | PO (gavage) | 2× per week for 28 days | 10 mL kg ⁻¹ | None | 15% in DI water; age 7–8 weeks; ♂/♀ |

TABLE E.18 Carboxymethylcellulose (CMC)

| Species | Route | Duration | Dose | Adverse Reactions/Toxicity | Notes |
|------------------|-------------|----------|----------------------------|----------------------------|---|
| Minipig | PO | 14 days | 8.0 mL kg ⁻¹ SD | None | 0.5% CMC in water; GLP; age 3 months; 1♂/1♀ |
| | PO | 28 days | 8.0 mL kg ⁻¹ QD | None | 0.5% CMC in water; GLP; age 3–4 months; 4♂/4♀ |
| Nonhuman primate | PO | 30 days | | Well tolerated | 5% in water |
| | SC | Acute | 10 mL kg ⁻¹ | Well tolerated | |
| Rat | PO | | 20 mg kg ⁻¹ | NOEL | 5% in water |
| | PO | 14 days | 8.0 mL kg ⁻¹ SD | None | 0.5% CMC in water; GLP; age 8 weeks; 5♂/5♀ |
| | PO | 28 days | 8.0 mL kg ⁻¹ QD | None | 0.5% CMC in water; GLP; age 8 weeks; 10♂/10♀ |
| | PO (gavage) | 93 weeks | 10 mL kg ⁻¹ QD | None | 1% CMC (medium viscosity) in DI water; age 6 weeks; ♂/♀ |

TABLE E.19 Carboxymethylcellulose Calcium (Calcium CMC)

| Species | Route | Duration | Dose (mL kg ⁻¹) | Adverse Reactions/Toxicity | Notes |
|---------|-------|----------|-----------------------------|----------------------------|-------------|
| Dog | PO | 90 days | 1 | Well tolerated | 1% solution |

TABLE E.20 Carboxymethylcellulose Sodium (Sodium CMC)

| Species | Route | Duration | Dose (mL kg ⁻¹) | Adverse Reactions/Toxicity | Notes |
|---------|-------|----------|-----------------------------|----------------------------|-------------|
| Rabbit | PO | 1 month | 0.5 | Well tolerated | 1% solution |

TABLE E.21 Cetyl Alcohol

| Species | Route | Duration | Dose (mg kg ⁻¹) | Adverse Reactions/Toxicity | Notes |
|---------|-------|----------|-----------------------------|----------------------------|-------|
| Mouse | IP | 1 month | 100 | Well tolerated | |

TABLE E.22 Citrate Buffer

| Species | Route | Duration | Dose | Adverse Reactions/Toxicity | Notes |
|---------|-------------|----------|--|--------------------------------------|----------------|
| Dog | IV infusion | 8 doses | 30 mL kg ⁻¹ day ⁻¹ | Well tolerated | 0.1 M, aqueous |
| | SC | 30 days | 10 mL kg ⁻¹ QD | Well tolerated | |
| Rat | IV | 4 weeks | nd | Hypoactivity, pain at injection site | 100 mM at pH 5 |
| | PO | 2 weeks | 10 mL kg ⁻¹ | Well tolerated | 50 mM |
| | PO | 2 weeks | 15 mL kg ⁻¹ | Well tolerated | 50 mM |

TABLE E.23 Citric Acid Buffer

| Species | Route | Duration | Dose (mL kg ⁻¹ QD) | Adverse Reactions/Toxicity | Notes |
|------------------|-------------|----------|-------------------------------|----------------------------|--|
| Mouse | PO (gavage) | 182 days | 2 | None | 0.015 M at pH 4.50; age 6–7 weeks; ♂/♀ |
| Nonhuman primate | PO (gavage) | 39 weeks | 7.5 | None | 10 mM; age 3–3.5 years; ♂/♀ |
| Rat | PO | 2 weeks | 10 | Well tolerated | 50 mM |
| | PO | 2 weeks | 15 | Well tolerated | 50 mM |

TABLE E.24 Collagen Matrix

| Species | Route | Duration | Dose | Adverse Reactions/Toxicity | Notes |
|------------------|------------------------------|----------|---|----------------------------|----------------------------------|
| Nonhuman primate | Implantation in humerus bone | 6 months | 10 mL kg ⁻¹ day ⁻¹ | Well tolerated | Bovine type I and hydroxyapatite |
| Rabbit | Implantation | 6 months | Single application, 5 mL kg ⁻¹ | Well tolerated | |

TABLE E.25 Corn Oil

| Species | Route | Duration | Dose | Adverse Reactions/Toxicity | Notes |
|------------------|--------------------|-------------|--------------------------|--|------------------|
| Chicken embryo | Injection into egg | Single dose | 0.1 µL g ⁻¹ | Less mortality than 1.0 µL g ⁻¹ egg | |
| | Injection into egg | Single dose | 1.0 µL g ⁻¹ | Increased mortality, decreased activity during righting reflex, running time, visual discrimination, and olfactory aversion test | |
| Dog | PO | 1 month | 3.0 mL kg ⁻¹ | Well tolerated | |
| Mouse | PO | 1 month | 2.5 mL kg ⁻¹ | Well tolerated | |
| Nonhuman primate | PO | 1 month | 1 mL kg ⁻¹ | Well tolerated | |
| Rabbit | PO | 1 month | 1 mL kg ⁻¹ | Well tolerated | |
| Rat | PO (gavage) | Single dose | 10 mL kg ⁻¹ | Well tolerated | |
| | PO (gavage) | 20 doses | 5 mL kg ⁻¹ | Well tolerated | |
| | PO (gavage) | 90 days | 5 mL kg ⁻¹ QD | None | Age 6 weeks; ♂/♀ |

TABLE E.26 Cottonseed Oil

| Species | Route | Duration | Dose (mL) | Adverse Reactions/Toxicity | Notes |
|---------|-------|-------------|-----------|---|---------------|
| Dog | SC | Single dose | 1 | Well tolerated; no evidence of irritation macroscopically or histologically | 100% solution |

TABLE E.27 Cyclohexane

| Species | Route | Duration | Dose (mL kg ⁻¹ day ⁻¹) | Adverse Reactions/Toxicity | Notes |
|---------|-------------|----------|---|--|-------|
| Rabbit | PO | 30 days | 0.5 | Well tolerated | |
| Rat | Dermal | 30 days | 1 | Well tolerated | |
| | PO (gavage) | 4 weeks | 5 | Intermittent convulsive after dosing, piloerection round back and emaciated appearance | |

TABLE E.28 Dextrose

| Species | Route | Duration | Dose | Adverse Reactions/Toxicity | Notes |
|------------------|-----------------|-------------|--|----------------------------|------------------------------------|
| Cat | Oral mucosa | 24h | 0.6 mL SD | None | 5%; non-GLP; age >6 months; 3♂/3♀ |
| Dog | IV | Single dose | 150 mL h ⁻¹ | Well tolerated | 5%, USP |
| | IV/PO | ADME | 2/10 mL kg ⁻¹ day ⁻¹ | Well tolerated | 5% solution |
| Nonhuman primate | PO (gavage) | 13 weeks | 0.78–9.3 mL kg ⁻¹ day ⁻¹ | Well tolerated | 10% solution (w/w) |
| | PO (gavage) | ADME | 5 mL kg ⁻¹ day ⁻¹ | Well tolerated | 5% solution |
| | PO (gavage) | Card. Vas. | 5 mL kg ⁻¹ day ⁻¹ | Well tolerated | 5% solution |
| Rabbit | IV (slow bolus) | 12 doses | | Well tolerated | 5%, USP |
| Rat | IV | Single dose | 1.4 mL per animal | Well tolerated | 5%, USP |
| | IV | 7 days | 5 mL kg ⁻¹ SD | None | 5%; non-GLP; age 7–10 weeks; 2♂/2♀ |
| | PO (gavage) | 26 weeks | 0.71–8.6 mL kg ⁻¹ day ⁻¹ | Well tolerated | 10% solution (w/w) |
| | PO (gavage) | Prelim | 5 mL kg ⁻¹ day ⁻¹ | Well tolerated | 5% solution |
| | SC | 2 weeks | 0.75 mL kg ⁻¹ day ⁻¹ | Well tolerated | 5% solution |

TABLE E.29 Dichlorvos

| Species | Route | Duration | Dose | Adverse Reactions/Toxicity | Notes |
|------------------|-------------|----------|--|----------------------------|---|
| Nonhuman primate | IV infusion | 2 weeks | 2 mL kg ⁻¹ 10 min 3× per week | Well tolerated | 10 mg mL ⁻¹ dichlorvos; age 3–6.5 years; ♀ |

TABLE E.30 Diethylacetamide

| Species | Route | Duration | Dose | Adverse Reactions/ Toxicity | Notes |
|---------|-------|---------------|---|--------------------------------|--|
| Cat | IV | Single dose | 1 g kg ⁻¹ (1000 mg kg ⁻¹) | LD _{Lo} | Behavioral: altered sleep time (including change in righting reflex) |
| Chicken | IV | Single dose | 3900 mg kg ⁻¹ | LD _{Lo} | |
| Dog | IV | Single dose | 1 g kg ⁻¹ (1000 mg kg ⁻¹) | LD _{Lo} | Behavioral: altered sleep time (including change in righting reflex) |
| Mouse | IP | Single dose | 1600 mg kg ⁻¹ | LD ₅₀ | Sense organs and special senses: mydriasis (pupillary dilation) |
| | IV | Range finding | MTD: 1.4 g kg ⁻¹ ; NOEL: 468 mg kg ⁻¹ | | Published LD ₅₀ = 2.3–3.2 g kg ⁻¹ |
| Rabbit | IV | Single dose | 1920 mg kg ⁻¹ | LD _{Lo} | |
| Rat | IP | Single dose | 1840 mg kg ⁻¹ | LD ₅₀ | |
| | IV | Single dose | 1 g kg ⁻¹ (1000 mg kg ⁻¹) | LD ₅₀ | Behavioral: altered sleep time (including change in righting reflex) |

TABLE E.31 Diethylene Glycol Monoethyl Ether (DEGEE)

| Species | Route | Duration | Dose | Adverse Reactions/Toxicity | Notes |
|------------------|-------|----------------|--|----------------------------|---------------------|
| Nonhuman primate | IV | 1 month (ADME) | 0.355 mL kg ⁻¹ single injection | Well tolerated | Into saphenous vein |

TABLE E.32 Dimethylacetamide (DMA)

| Species | Route | Duration | Dose | Adverse Reactions/Toxicity | Notes |
|---------|---------------------|----------|--|--|--|
| Chicken | IV | Acute | 12 000 mg kg ⁻¹ | LD _{Lo} | Lowest published lethal dose |
| Dog | Dermal | 6 weeks | 2 690.476 mg kg ⁻¹ QD | TD _{Lo} ; fatty liver degeneration; chronic death (nutritional and gross metabolic); reproductive | |
| Mouse | IP | Acute | 2 800 mg kg ⁻¹ | LD ₅₀ | |
| | IV | Acute | 3 020 mg kg ⁻¹ | LD ₅₀ | |
| | IV (into tail vein) | Acute | 469 mg kg ⁻¹ (dose vol. 5 mL kg ⁻¹) | NOEL | 10% solution |
| | IV (into tail vein) | Acute | 1 405 mg kg ⁻¹ (dose vol. 5 mL kg ⁻¹) | MTD | 30% solution; mild to moderate hypoactivity for up to 6 min is typical |
| | PO | Acute | 4 620 mg kg ⁻¹ | LD ₅₀ | Intragastric feeding or introduction with drinking water |
| Rabbit | Dermal | Acute | 2 240 mg kg ⁻¹ | LD ₅₀ | Application directly onto the skin, either intact or abraded |
| | IV | Acute | 8 340 mg kg ⁻¹ | LD _{Lo} | Lowest published lethal dose |
| | PO | 13 days | 3 900 mg kg ⁻¹ | TD _{Lo} ; specific developmental abnormalities (eye, ear, craniofacial including nose and tongue, musculoskeletal system); postimplantation mortality; fetotoxicity (except death, e.g., stunted fetus) | Intragastric feeding or introduction with drinking water; 6–18 days of pregnancy |

TABLE E.32 (Continued)

| Species | Route | Duration | Dose | Adverse Reactions/Toxicity | Notes |
|---------|------------|-----------------------------------|------------------------------|--|---|
| Rat | Inhalation | 6 h day ⁻¹ for 10 days | 281 ppm | TD _{Lo} | Inhalation in chamber by cannulation or through mask; 6–15 days of pregnancy |
| | IP | | 2 mg kg ⁻¹ | TD _{Lo} ; postimplantation mortality; fetotoxicity (except death, e.g., stunted fetus) | 1-day pregnant |
| | IP | Acute | 2 750 mg kg ⁻¹ | LD ₅₀ | |
| | IV | Acute | 2 640 mg kg ⁻¹ | LD ₅₀ | |
| | PO | 14 days | 5 600 mg kg ⁻¹ | TD _{Lo} ; effects on fertility, postimplantation mortality; fetal death; specific developmental abnormalities (craniofacial, including nose and tongue, musculoskeletal system, cardiovascular system, homeostasis); fetotoxicity (except death, e.g., stunted fetus) | Intragastric feeding or introduction with drinking water; 6–19 days of pregnancy |
| | PO | 14 days | 2 240 mg kg ⁻¹ | TD _{Lo} | Intragastric feeding or introduction with drinking water; 6–19 days of pregnancy |
| | PO | Acute | 4 300 mg kg ⁻¹ | LD ₅₀ | Intragastric feeding or introduction with drinking water |
| | PO | 10 days | 1 500 mg kg ⁻¹ QD | TD _{Lo} ; findings in digestive system and liver; chronic | Intragastric feeding or introduction with drinking water; lowest published toxic dose |
| | PO | 26 weeks | 2 mg kg ⁻¹ QD | TD _{Lo} ; enzyme inhibition, induction, or change in blood or tissue levels; liver | Intragastric feeding or introduction with drinking water; lowest published toxic dose |
| | PO | 90 days | 50 mg kg ⁻¹ QD | TD _{Lo} ; changes in erythrocyte (RBC) and leukocyte (WBC) counts | Intragastric feeding or introduction with drinking water; lowest published toxic dose |

TABLE E.33 Dimethyl Sulfoxide (DMSO)

| Species | Route | Duration | Dose | Adverse Reactions/Toxicity | Notes |
|------------------|---------------------|--|--|--|--|
| Dog | IV | Single dose at 0.2 mL min ⁻¹ for 18.5 min | 0.43–0.46 mL kg ⁻¹ | During dose administration excessive salivation and labored respiration were noted. At 1 and 4 h post dose | <i>Must</i> use IV catheter; 100% solution |
| | IV | | 2 mg kg ⁻¹ | | |
| | IV | | 2.5 g kg ⁻¹ | LD ₅₀ | MUST use IV catheter; 100% solution |
| | IV | Intermittent for 4 weeks | 57 600 mg kg ⁻¹ | LD ₅₀ ; lowest published toxic dose, hematuria, normocytic anemia, death | MUST use IV catheter |
| | IV | | 2 500 mg kg ⁻¹ | LD ₅₀ ; cardiac changes, hematuria | |
| | | Single dose | 1 mL kg ⁻¹ | Well tolerated | 10% solution |
| | IV | | 0.1 mL kg ⁻¹ | | MUST use IV catheter; 100% solution |
| | IV | 1 month | 1.25 mL/ (0.112) × (BW) | Well tolerated | |
| Guinea pig | IV | 1 month | 0.1 mL kg ⁻¹ | Well tolerated | |
| | IP | Acute | 6.5 g kg ⁻¹ | LD ₅₀ | 100% solution |
| Mouse | IP | 1 month | 2.5 g kg ⁻¹ | LD ₅₀ | 100% solution |
| | IP | | 3.82–10.73 g kg ⁻¹ | LD ₅₀ | 100% solution |
| | IP | | 8.2 g kg ⁻¹ | LD ₅₀ | 100% solution |
| | IP | | 20.06 g kg ⁻¹ | LD ₅₀ | 100% solution |
| | IP | 1 month | 100 mg kg ⁻¹ | Well tolerated | |
| | IP | 3 days | 10 mL kg ⁻¹ | Well tolerated | 15% solution |
| | IV (into tail vein) | Acute | 1 650 mg kg ⁻¹ (dose vol. 5 mL kg ⁻¹) | NOEL | 30% solution |
| | IV (into tail vein) | Acute | 2 200 mg kg ⁻¹ (dose vol. 5 mL kg ⁻¹) | MTD; rapid breathing, ataxia, and muscle contractions; full recovery by 1 min is typical | 40% solution |
| | IV | Range finding | MTD: 2.2 g kg ⁻¹ ; NOEL: 1.6 g kg ⁻¹ | | Published LD ₅₀ = 3.8–7.6 g/kg |
| | IV | | 3 100 mg kg ⁻¹ | LD ₅₀ ; eye hemorrhage, conjunctiva irritation | |
| | IV | | 240 g kg ⁻¹ | Lowest published toxic dose. Postimplantation mortality | Age day 1–20 presumed pregnant |
| | PO (gavage) | | 15.0–22 g kg ⁻¹ | LD ₅₀ | 100% solution |
| | PO (gavage) | | 7.9 g kg ⁻¹ | LD ₅₀ | 100% solution |
| | PO | | 5 mL kg ⁻¹ | | |
| | SC | | 13.9–20.5 g kg ⁻¹ | LD ₅₀ | 100% solution |
| Nonhuman primate | PO (gavage) | Efficacy | 3 mL kg ⁻¹ day ⁻¹ | Well tolerated | |
| Rabbit | SC | 1 month | 1 mL kg ⁻¹ | Erythema, inflammation | |

TABLE E.33 (Continued)

| Species | Route | Duration | Dose | Adverse Reactions/Toxicity | Notes |
|---------|-------------|-------------|---|---|---------------|
| Rat | IV | | 4–5 mg kg ⁻¹ | | |
| | IV | | 5.25–5.36 g kg ⁻¹ | LD ₅₀ | 100% solution |
| | IV | | 5.3 g kg ⁻¹ | LD ₅₀ | 100% solution |
| | IV | Single dose | 200 mg kg ⁻¹ | In serum, slightly and transiently changed metabolic parameters including glucose, lactate, triglycerides, free fatty acids, or creatinine as well as electrolytes (Na, Cl, Mg) and osmolality, increased ALT, impeded clinical chemistry measurements of various parameters at 4 h postdose, kidney function—induced loss of protein and albumin | 2% solution |
| | IV | | 5 360 mg kg ⁻¹ | LD ₅₀ ; tremors, muscle weakness, dyspnea | |
| | IV | 1 month | 200 mg kg ⁻¹ | Well tolerated | |
| | SC | | 12 g kg ⁻¹ | LD ₅₀ | 100% solution |
| | PO (gavage) | 7 days | 5 mL kg ⁻¹ day ⁻¹ | Well tolerated | |
| | PO (gavage) | 4 weeks | 5 mL kg ⁻¹ day ⁻¹ | Well tolerated | |
| | PO (gavage) | | 16.0–28.3 g kg ⁻¹ | LD ₅₀ | 100% solution |
| | PO (gavage) | | 14.5 g kg ⁻¹ | LD ₅₀ | 100% solution |
| | PO (gavage) | Single dose | 200 mg kg ⁻¹ | Did not affect stomach emptying, did not reduce intestinal transit time | 2% solution |
| | PO (gavage) | Single dose | 1 000 mg kg ⁻¹ | Did not affect stomach emptying, did not reduce intestinal transit time | 10% solution |
| | IP | 28 doses | 5 mL kg ⁻¹ | Well tolerated | 15% solution |
| | IP | | 6.5–13.621 g kg ⁻¹ | LD ₅₀ | 100% solution |
| | IP | | 8.2 g kg ⁻¹ | LD ₅₀ | 100% solution |

TABLE E.34 Dulbecco's Modified Phosphate-Buffered Saline (PBS)

| Species | Route | Duration | Dose (mL kg ⁻¹ day ⁻¹) | Adverse Reactions/Toxicity | Notes |
|---------|---------------------|----------|---|----------------------------|-------|
| Rat | IV (into tail vein) | 1 month | 1 | Well tolerated | |

TABLE E.35 Dulbecco's Phosphate-Buffered Saline (PBS)

| Species | Route | Duration | Dose (mg kg ⁻¹ day ⁻¹) | Adverse Reactions/Toxicity | Notes |
|---------|-------------|----------|---|----------------------------|-------|
| Rat | PO (gavage) | 1 month | 0.1, 0.8, and 1.2 | Well tolerated | |

TABLE E.36 Ethanol (EtOH)

| Species | Route | Duration | Dose | Adverse Reactions/Toxicity | Notes |
|---------|---------------------|---|--|---|---|
| Dog | IV | | 5 mL kg ⁻¹ at a rate of 0.3 mL kg ⁻¹ | Hemolysis <i>in vitro</i> dog blood | 30% solution |
| | IV | | 5 mL kg ⁻¹ at a rate of 0.3 mL kg ⁻¹ | Hemolysis <i>in vitro</i> dog blood | 40% solution |
| | IV | Single dose | 1 mL kg ⁻¹ | Well tolerated | 10% solution |
| | IV | Single dose | 1 mL kg ⁻¹ | CNS depression, ataxia | 30% solution |
| | IV | 5 days | 1 mL kg ⁻¹ at 2 mL min ⁻¹ | Excessive salivation | 30% solution; MUST use winged infusion set |
| | IV | At 3, 7, 14, and 24 h after ingestion of ethylene glycol | 1584 mg kg ⁻¹ | Remained recumbent or severely ataxic for 36, depressed for 72 | 20% solution |
| Minipig | IV | Every 4 h for five treatments, then every 6 h for four treatments | 5.5 mL kg ⁻¹ | | 20% solution |
| | PO | 1 month | 5 mL kg ⁻¹ | Well tolerated | 7.5% solution |
| | PO | 6 months | 400 mL kg ⁻¹ | Hepatotoxicity, myopathy; CNS changes | |
| | PO | 90 days | 5 mL kg ⁻¹ | Well tolerated | 5% solution |
| | Dermal | ADME | nd (5 mg cm ⁻²) | Well tolerated | 60/40: purified water/ethyl alcohol absolute: v/v |
| | | | | | |
| Mouse | Dermal | 13 weeks | 100 µL per animal per day | Well tolerated | 70% (62% m/m) |
| | Dermal | 7 days | 0.5 mL (fixed vol.) QD | Well tolerated | 80% solution; age 6 weeks; ♂/♀ |
| | IP | Acute | 5 mL kg ⁻¹ | Well tolerated | 5% solution |
| | IV | Range finding | MTD: 986 mg kg ⁻¹ ; NOEL: 197 mg kg ⁻¹ | | Published LD ₅₀ = 1.6–4.3 g kg ⁻¹ |
| | IV (into tail vein) | Acute | 197 mg kg ⁻¹ (dose vol. 5 mL kg ⁻¹) | NOEL | 5% solution |
| | IV (into tail vein) | Acute | 986 mg kg ⁻¹ (dose vol. 5 mL kg ⁻¹) | MTD: ventral recumbency and “swimming” behavior immediately postdose; ataxia for up to 6 min is typical | 25% solution |
| PO | 1 month | | 2.5 mL kg ⁻¹ | Well tolerated | 5% solution |
| PO | 6 month | | 2500 g kg ⁻¹ | Well tolerated | |

| Nonhuman primate | PO | 9 months | 250 g kg ⁻¹ | Behavioral changes |
|------------------|-------------|-----------|--|--|
| Rat | Dermal | ADME | nd (5 mg cm ⁻²) | Well tolerated |
| | Dermal | 91 days | 0.5 mL (fixed vol.) QD | Well tolerated |
| | IP | Acute | 3.75 g kg ⁻¹ | LD ₅₀ |
| | IV | Acute | 1.44 g kg ⁻¹ | LD ₅₀ |
| | IV | Acute | 5 mL kg ⁻¹ at a rate of 0.3 mL kg ⁻¹ | Hematuria |
| | IV | 9 days | 5 mL kg ⁻¹ at 2 mL min ⁻¹ | Ataxia, respiratory depression, and death; is dosed faster than 2 mL min ⁻¹ |
| | IV | 12 months | 250 g kg ⁻¹ | Nephrosis, ATN, bladder changes, weight loss |
| | PO | Acute | 7.06 g kg ⁻¹ | LD ₅₀ |
| | PO (gavage) | 7 days | 0.8, 2, and 5 mL kg ⁻¹ day ⁻¹ | Well tolerated |
| | PO | 7 days | 10 mL kg ⁻¹ | Well tolerated |
| | PO | Acute | 5 mL kg ⁻¹ | Depression |
| | PO (gavage) | 4 weeks | 2 mL kg ⁻¹ QD | Hypokinesia, dyspnea regurgitation, distended lungs/ileum, and swollen abdomen |
| | PO | 28 doses | 175 g kg ⁻¹ | Depression, decreased RBC |
| | PO | 90 day | 8 mL kg ⁻¹ | Well tolerated |
| | PO | 12 month | 1000 mg kg ⁻¹ | Fatty liver |

TABLE E.37 Gelatin Capsules

| Species | Route | Duration (days) | Dose | Adverse Reactions/Toxicity | Notes |
|---------|-------|-----------------|------|----------------------------|-------|
| Dog | PO | 5 | QD | Well tolerated | |
| | PO | 6 | QD | Well tolerated | |
| | PO | 8 | QD | Well tolerated | |
| | PO | 14 | QD | Well tolerated | |
| | PO | 16 | QD | Well tolerated | |

TABLE E.38 Gelatin Phosphate Buffer

| Species | Route | Duration | Dose | Adverse Reactions/Toxicity | Notes |
|---------|---------|----------|----------------------|----------------------------|----------------------------|
| Minipig | Topical | 28 days | 10 mL q14d × 2 doses | None | GLP; age 4–6 months; 5♂/5♀ |

TABLE E.39 Gelucire 44/14™

| Species | Route | Duration | Dose | Adverse Reactions/Toxicity | Notes |
|---------|--------|----------|---|--|-------|
| Dog | PO | 3 months | 400, 1000, and 2500 mg kg ⁻¹ day ⁻¹ | NOAEL: >2500 mg kg ⁻¹ day ⁻¹ | |
| | PO | 14 days | 400, 1000, and 2500 mg kg ⁻¹ day ⁻¹ | NOAEL: >2500 mg kg ⁻¹ day ⁻¹ | |
| Rabbit | Dermal | Acute | 0.5 mL | Not irritant | |
| | Ocular | Acute | 0.1 mL | Slightly irritant | |
| Rat | PO | 28 day | 600, 1500, and 2400 mg kg ⁻¹ day ⁻¹ | NOEL: 2400 mg kg ⁻¹ day ⁻¹ | |
| | PO | 7 day | 600, 1500, and 2400 mg kg ⁻¹ day ⁻¹ | NOEL: 2400 mg kg ⁻¹ day ⁻¹ | |
| | PO | Acute | No dilution | LD ₅₀ : >2004 mg kg ⁻¹ day ⁻¹ | |

TABLE E.40 Gelucire 50/13™

| Species | Route | Duration | Dose | Adverse Reactions/Toxicity | Notes |
|---------|-------|----------|------|----------------------------|---|
| Rat | PO | Acute | | Well tolerated | No dilution; LD ₀ ≥ 20 000 mg kg ⁻¹ day ⁻¹ |

TABLE E.41 Gluconic Acid

| Species | Route | Duration | Dose (mL kg ⁻¹ QD) | Adverse Reactions/Toxicity | Notes |
|---------|-------------|----------|-------------------------------|----------------------------|--|
| Dog | PO (gavage) | 14 days | 2 | Well tolerated | 0.3 M gluconic acid pH 3.0; age 11 months; ♂/♀ |
| Rat | PO (gavage) | 14 days | 10 | Well tolerated | 0.3 M gluconic acid pH 3.0; age 10 weeks; ♂/♀ |

TABLE E.42 Glycerol

| Species | Route | Duration | Dose | Adverse Reactions/Toxicity | Notes |
|------------|-------|----------|--------------------------------------|------------------------------------|---|
| Dog | SC | 28 days | 20 mL day ⁻¹ (fixed vol.) | None | 2% solution in sterile water; age 5–6 months; ♂/♀ |
| Guinea pig | PO | 1 month | 500 mg kg ⁻¹ | Well tolerated | |
| Mouse | IV | 1 month | 100 mg kg ⁻¹ | Well tolerated | |
| | IP | 1 month | 250 mg kg ⁻¹ | Well tolerated | |
| | PO | 90 days | 500 mg kg ⁻¹ | Depression and reduced respiration | |
| | SC | Acute | 10 mg kg ⁻¹ | Well tolerated | |
| Rabbit | IV | Acute | 10 mg kg ⁻¹ | Well tolerated | |
| Rat | PO | Acute | 1000 mg kg ⁻¹ | Well tolerated | |
| | PO | 1 month | 15 g kg ⁻¹ | Reduced adrenal weights | |
| | PO | 1 month | 1000 mg kg ⁻¹ | Well tolerated | |
| | SC | Acute | 10 mg kg ⁻¹ | Well tolerated | |

TABLE E.43 Glycofurool

| Species | Route | Duration | Dose (mL kg ⁻¹) | Adverse Reactions/Toxicity | Notes |
|---------|-------|-------------|-----------------------------|----------------------------|--------------|
| Dog | IV | Single dose | 1 | Well tolerated | 50% solution |

TABLE E.44 Gum Tragacanth

| Species | Route | Duration | Dose (mL kg ⁻¹ QD) | Adverse Reactions/Toxicity | Notes |
|---------|-------------|----------|-------------------------------|----------------------------|--------------------------|
| Mouse | PO (gavage) | 2 weeks | 10 | Well tolerated | In distilled water, 0.5% |

TABLE E.45 Gum Xanthan

| Species | Route | Duration | Dose (mL kg ⁻¹ day ⁻¹) | Adverse Reactions/Toxicity | Notes |
|---------|-------------|-----------|---|----------------------------|-----------------------|
| Rabbit | PO (gavage) | Tolerance | 3 | Well tolerated | 0.4% aqueous solution |
| | PO (gavage) | Segmt. II | 3 | Well tolerated | 0.4% aqueous solution |

TABLE E.46 Hydrochloric Acid (HCl)

| Species | Route | Duration | Dose | Adverse Reactions/Toxicity | Notes |
|---------|-------------|----------|---------------------------|----------------------------|--|
| Dog | PO (gavage) | 4 weeks | 4 mL kg ⁻¹ QD | None | 0.1 N; age 5–6 months; ♂/♀ |
| | PO (gavage) | Daily | 10 mL kg ⁻¹ | None | 0.1–10% in water; beagles; age 5 months; ♂/♀ |
| Rat | PO (gavage) | 26 weeks | 10 mL kg ⁻¹ QD | None | 0.05 M HCl; age 6 weeks; ♂/♀ |
| | PO (gavage) | Daily | 10 mL kg ⁻¹ | None | 0.1–10% in water; age 6–8 weeks; ♂/♀ |

TABLE E.47 Hydroxyethylcellulose

| Species | Route | Duration | Dose (mg kg ⁻¹) | Adverse Reactions/Toxicity | Notes (%) |
|---------|-------|-------------|-----------------------------|--|-----------|
| Rat | PO | Single dose | 20 | Intestinal transit was slightly enhanced, not dose dependent | 0.5 |
| | PO | Single dose | 100 | Intestinal transit was slightly enhanced, not dose dependent | 1 |
| | PO | 28 days | 50 | Easiest and most tolerable formulation for PO administration | |

TABLE E.48 Hydroxypropyl- β -Cyclodextrin (HP β CD)

| Species | Route | Duration | Dose | Adverse Reactions/Toxicity | Notes |
|------------------|-----------------------|-----------|---|--|---|
| Dog | IV | 1 dose | 1 mL kg ⁻¹ | Well tolerated | 40% solution |
| | IV (slow bolus) | 1 dose | 1.2 mL kg ⁻¹ | Well tolerated | 6% solution |
| | IV (2 h infusion) | 1 month | 10 mL kg ⁻¹ | Well tolerated | 40% solution |
| | Intranasal | 14 days | 1 mL per nostril TID 2 h apart | Well tolerated | 45%; beagle dogs (marshall); age ~6 months; ♂/♀ |
| | PO (gavage) | 1 dose | 10 mL kg ⁻¹ | Well tolerated | 6% solution |
| Mouse | PO (gavage) | 90 days | Dose vol. 5 mL kg ⁻¹ ; daily dose 500 and 1 000 mg kg ⁻¹ | Loose/soft stools in high-dose group | Dose concentrations 100 and 200 mg mL ⁻¹ , respectively |
| | PO (gelatin capsules) | 28 doses | | Emesis, fecal changes | 6% solution |
| | SC | 91 days | 0.86 mL kg ⁻¹ QD | None | 10% solution in sterile water; age 5.5–6 months; ♂/♀ |
| | IV | Acute | 5 000 mg kg ⁻¹ | LD | Administration directly into the vein by hypodermic needle; >5 g kg ⁻¹ |
| | PO (gavage) | 90 days | Dose vol. 10 mL kg ⁻¹ ; daily dose 500 and 1 000 mg kg ⁻¹ | Produced elevated transaminase (aspartate and alanine aminotransferase) levels; use with caution | Dose concentrations 50 and 100 mg mL ⁻¹ , respectively |
| Nonhuman primate | PO (gavage) | 104 weeks | 500 mg kg ⁻¹ day ⁻¹ | Well tolerated | |
| | PO (gavage) | 13 doses | 5 mL kg ⁻¹ | None | 11% solution |
| | PO (gavage) | 90 days | Dose vol. 5 mL kg ⁻¹ ; daily dose 500 and 1 000 mg kg ⁻¹ | Loose/soft stools in high-dose group | Dose concentrations 100 and 200 mg mL ⁻¹ , respectively |
| | IP | Acute | 10 000 mg kg ⁻¹ | Well tolerated, LD > 10 000 mg kg ⁻¹ | |
| Rabbit | PO (gavage) | 12 doses | 2 mL kg ⁻¹ | Well tolerated | 11% solution |

| Rat | Intranasal | 14 days | 50 mL per nostril TID 2 h apart | Well tolerated | 45%; Sprague-Dawley (Harlan); age ~8–10 weeks at initiation; ♂/♀; histopathology limited to purulent exudates (minimal to mild) involving the nasal turbinates |
|-----|-------------------|-------------|--|---|--|
| | IP injection | Single dose | 1 000 mg kg ⁻¹ | Increased glucose levels at 4 h; minor transient changes for triglycerides and BUN; no functional changes were observed; only slight enhancement of ALT and AST | 10% solution |
| | IV | Single dose | 1 mL kg ⁻¹ | Well tolerated | 20% solution |
| | IV | Single dose | 10 mL kg ⁻¹ | Death occurred within a few minute of receiving a bolus dose of the vehicle. The rate of administration was slowed to ~2 min, which was tolerated and clinical observations limited to red urine | 45%; Sprague-Dawley (Harlan); age ~8 weeks of age; ♂/♀ |
| | IV (slow bolus) | 10 doses | 4 mL kg ⁻¹ | Well tolerated | 12.5% solution |
| | IV (1 h infusion) | 1 month | 10 mL kg ⁻¹ | Well tolerated | 40% solution |
| | IV (slow bolus) | | 2 mL kg ⁻¹ | Well tolerated | 12.5% solution |
| | PO | Acute | 5 000 mg kg ⁻¹ | LD ₅₀ > 5 000 mg kg ⁻¹ | |
| | PO | Single dose | Up to 2 000 mg kg ⁻¹ | No effect on gastric emptying, modestly inhibited intestinal transit | 20% solution |
| | PO (gavage) | Single dose | 10 mL kg ⁻¹ | Well tolerated | 6% solution |
| | PO (gavage) | 10 doses | 10 mL kg ⁻¹ | Well tolerated | 11% solution |
| | PO (gavage) | 4 weeks | 10 mL kg ⁻¹ QD | None | 20%; age 6 weeks; ♂/♀ |
| | PO (gavage) | 90 days | Dose vol. 5 mL kg ⁻¹ ; daily dose 500 and 1 000 mg kg ⁻¹ | Produced elevated transaminase (aspartate and alanine aminotransferase) levels; use with caution | Dose concentrations 100 and 200 mg mL ⁻¹ , respectively |
| | PO (gavage) | 2 years | 500 mg kg ⁻¹ | No effects | |
| | SC | 28 days | QD | Erosion/ulceration and/or necrosis generally observed. At end of 28-day recovery period, fibrosis and subacute/chronic inflammation at the injection sites persisted but were resolving; injection site necrosis not observed | 10% HPβCD |
| | SC | 91 days | QD | At end of 28-day recovery period, fibrosis and subacute/chronic inflammation at the injection sites persisted but were resolving; injection site necrosis not observed | 10% HPβCD |
| | SC | 91 days | 1.14 mL kg ⁻¹ QD | None | 10% solution in sterile water; age 7 weeks; ♂/♀ |

FDA has informed sponsors that this excipient is no longer permitted for use in subcutaneous clinical formulations due to perceived risk of carcinogenicity (Stella and He, 2008) and has advised that it not be used in repeat-dose nonclinical safety studies.

TABLE E.49 Hydroxypropyl Cellulose

| Species | Route | Duration | Dose (g kg ⁻¹) | Adverse Reactions/Toxicity | Notes |
|---------|-------|----------|----------------------------|----------------------------|-------|
| Rat | PO | 90 days | 1000 | Well tolerated | |

TABLE E.50 Hydroxypropyl Methylcellulose (HPMC)

| Species | Routes | Duration | Dose | Adverse Reactions/Toxicity | Notes |
|------------------|-------------|-----------------|---|----------------------------|--|
| Dog | IP | Acute | 200 mg kg ⁻¹ | Well tolerated | |
| | PO (gavage) | | 5 mL kg ⁻¹ | None | 1% (Methocel E5 Premium LV, 5cp) in DI water; age 6.5–7 months; ♂/♀ |
| | PO (gavage) | 90 days | Dose vol. 5 mL kg ⁻¹ , daily dose 20 mg kg ⁻¹ (0.4% w/v) | Well tolerated | |
| Minipig | PO (gavage) | 7 days | 5 mL kg ⁻¹ QD | None | 0.5% in distilled water; age 6.5–11.5 months; ♂/♀ |
| Mouse | IP | Single dose | 5 mL kg ⁻¹ | Well tolerated | 0.5% |
| | IP | Acute | 50 mg kg ⁻¹ | Well tolerated | |
| | PO | 5 days | 20 mL kg ⁻¹ BID | None | 0.5%; non-GLP; age 9 weeks; 6♀ |
| | PO (gavage) | Single dose | 10 mL kg ⁻¹ | Well tolerated | 0.5% |
| | PO (gavage) | 10 doses | 10 mL kg ⁻¹ | Well tolerated | 0.2% |
| | PO (gavage) | 90 days | Dose vol. 10 mL kg ⁻¹ , daily dose 20 mg kg ⁻¹ (0.4% w/v) | Well tolerated | |
| Nonhuman primate | PO (gavage) | 28 days | 2 mL kg QD | None | 0.5% in distilled water; age 2–3 years; ♂/♀ |
| | PO (gavage) | 28 days | 5 mL kg ⁻¹ QD | Soft feces (nonadverse) | 1% (Methocel E5 Premium LV, 5cp) in DI water; age 2.5–3.5 years; ♂/♀ |
| | PO (gavage) | 90 days | Dose vol. 5 mL kg ⁻¹ , daily dose 20 mg kg ⁻¹ (0.4% w/v) | Well tolerated | |
| | PO (gavage) | 91 days | 5 mL kg ⁻¹ QD | None | 0.2% in distilled water; cynomolgus monkeys; age 2 years; ♂/♀ |
| Rat | IP | Single dose | 5 mL kg ⁻¹ | Well tolerated | 0.5% |
| | PO (gavage) | Single dose | 10 mL kg ⁻¹ | Well tolerated | 0.2% |
| | PO (gavage) | Single dose | 10 mL kg ⁻¹ | Well tolerated | 0.5% |
| | PO (gavage) | 90 days | Dose vol. 5 mL kg ⁻¹ , daily dose 20 mg kg ⁻¹ (0.4% w/v) | Well tolerated | |
| | PO (gavage) | 91 days | 5 mL kg ⁻¹ QD | None | 1% (Methocel E5 Premium LV, 5cp) in DI water; age 6 weeks; ♂/♀ |
| | PO (gavage) | Up to 104 weeks | 10 mL kg ⁻¹ QD | None | 0.2% in distilled water; age 6–11 weeks; ♂/♀ |
| | PO (gavage) | 182 days | 10 mL kg ⁻¹ QD | None | 0.5% in distilled water; age 6 weeks; ♂/♀ |

TABLE E.51 Hypotonic Phosphate-Buffered Saline (PBS)

| Species | Route | Duration (days) | Dose (mL kg ⁻¹ SD) | Adverse Reactions/Toxicity | Notes |
|---------|-------|-----------------|-------------------------------|----------------------------|----------------------------|
| Dog | IV | 2 | 2 | None | GLP; age 6 months; 5♂/5♀ |
| Rat | IV | 2 | 2 | None | GLP; age ≥8 weeks; 10♂/10♀ |

TABLE E.52 Isopropyl Alcohol

| Species | Route | Duration | Dose (g kg ⁻¹) | Adverse Reactions/Toxicity | Notes |
|---------|--------|----------|----------------------------|----------------------------|-------|
| Rabbit | Dermal | 1 month | 1000 | Well tolerated | |

TABLE E.53 Isopropyl Myristate

| Species | Route | Duration | Dose (mg kg ⁻¹) | Adverse Reactions/Toxicity | Notes |
|---------|--------|----------|-----------------------------|----------------------------|-------|
| Rabbit | Dermal | 1 month | 500 | Well tolerated | |

TABLE E.54 Kolliphor EL™

| Species | Route | Duration (month) | Dose | Adverse Reactions/Toxicity | Notes |
|---------|-------|------------------|-------------------------|----------------------------|-------|
| Dog | IV | 1 | 2 mL kg ⁻¹ | Well tolerated | |
| Rat | PO | 1 | 100 mg kg ⁻¹ | Well tolerated | |

TABLE E.55 Kolliphor ELP™

| Species | Route | Duration | Dose | Adverse Reactions/Toxicity | Notes |
|---------|-------|----------|------|----------------------------|---------------------------------------|
| Dog | | Acute | | Anaphylaxis | Up to 50%; never use for dog studies! |

TABLE E.56 Kolliphor RH40™

| Species | Route | Duration | Dose (mL kg ⁻¹) | Adverse Reactions/Toxicity | Notes |
|---------|-------|----------|-----------------------------|----------------------------|-------|
| Dog | IV | 1 month | 2 | Well tolerated | |

TABLE E.57 Labrafil M1944™

| Species | Route | Duration | Dose | Adverse Reactions/Toxicity | Notes |
|---------|--------|-----------|-----------------------|----------------------------|---|
| Dog | PO | 1 month | 2 mg kg ⁻¹ | Well tolerated | |
| Rabbit | Dermal | Acute/PDI | | Non irritant | No dilution; 0.38 dermal irritation index |
| Rat | PO | Acute | 20 g kg ⁻¹ | Well tolerated | No dilution; LD ₀ ≥ 20 000 mg kg ⁻¹ day ⁻¹ |

TABLE E.58 Labrasol™

| Species | Route | Duration | Dose | Adverse Reactions/Toxicity | Notes |
|---------|--------|--------------------------------------|--|--|-------|
| Dog | PO | 14 days | 100, 300, 1000, and 3000 mg kg ⁻¹ day ⁻¹ | In high-dose group, moderate suppurative inflammation of the lungs. No adverse effects on survival and clinical observations | |
| | PO | 13 weeks | 0, 300, 1000, and 3000 mg kg ⁻¹ day ⁻¹ | NOEL: 1000 mg kg ⁻¹ day ⁻¹ ; NOAEL: 3000 mg kg ⁻¹ day ⁻¹ | |
| Rabbit | Dermal | Patch test | 0.5 mL | Well tolerated | |
| | Ocular | Acute | 0.1 mL | Slightly irritant | |
| Rat | Dermal | Acute | | Very well tolerated | |
| | Dermal | Patch test | 0.02 mL per animal | Well tolerated | |
| | IV | ADME | 10 mg kg ⁻¹ day ⁻¹ | | |
| | Ocular | | | Slightly irritant | |
| | PO | Acute | 20, 22.4, 25.1, 28.21, and 31.60 g kg ⁻¹ | LD ₅₀ > 22 g kg ⁻¹ ; well tolerated | |
| | PO | ADME | 10, 150 mg kg ⁻¹ day ⁻¹ | Well tolerated | |
| | PO | Segment II: embryo–fetal development | 1000, 2000, or 3000 mg kg ⁻¹ day ⁻¹ | NOEL: 3000 mg kg ⁻¹ day ⁻¹ with no indication of a teratogenicity | |
| | PO | 14 days | 100, 300, 1000, and 3000 mg kg ⁻¹ day ⁻¹ | NOAEL: 3000 mg kg ⁻¹ day ⁻¹ | |
| | PO | 6 months | 300, 1000, and 3000 mg kg ⁻¹ day ⁻¹ | NOEL: 300 mg kg ⁻¹ day ⁻¹ ; NOAEL: 3000 mg kg ⁻¹ day ⁻¹ | |

TABLE E.59 Lactose

| Species | Route | Duration | Dose (L min ⁻¹ per animal) | Adverse Reactions/Toxicity | Notes |
|------------------|------------|----------|---------------------------------------|----------------------------|---------------------------|
| Nonhuman primate | Inhalation | 2 weeks | 1 | Well tolerated | Lactose 200 μM; anhydrous |

TABLE E.60 Lanolin

| Species | Route | Duration | Dose (mg kg ⁻¹) | Adverse Reactions/Toxicity | Notes |
|---------|--------|----------|-----------------------------|----------------------------|-------|
| Rabbit | Dermal | 90 days | 1000 | Well tolerated | |

TABLE E.61 Lauroglycol 90™

| Species | Route | Duration | Dose | Adverse Reactions/Toxicity | Notes |
|---------|--------|----------|-------------|--|-------|
| Rabbit | Dermal | Acute | No dilution | Moderately irritant | |
| | Ocular | Acute | No dilution | Slightly irritant | |
| Rat | PO | Acute | | LD ₅₀ : >2003 mg kg ⁻¹ day ⁻¹ | |

TABLE E.62 Maltitol Solution

| Species | Route | Duration | Dose (mg kg ⁻¹) | Adverse Reactions/Toxicity | Notes |
|---------|-------|----------|-----------------------------|----------------------------|-------|
| Rat | IP | 1 month | 500 | Well tolerated | |

TABLE E.63 Maltol

| Species | Route | Duration (month) | Dose (mg kg ⁻¹) | Adverse Reactions/Toxicity | Notes |
|------------|-------|------------------|-----------------------------|----------------------------|-------|
| Guinea pig | PO | 1 | 75 | Well tolerated | |
| Rabbit | PO | 1 | 100 | Well tolerated | |

TABLE E.64 Mannitol

| Species | Route | Duration | Dose | Adverse Reactions/Toxicity | Notes |
|------------------|-------------|-----------|---|----------------------------|-------------|
| Minipig | SC | Tolerance | 0.2 mL kg ⁻¹ day ⁻¹ | Well tolerated | 5% solution |
| Nonhuman primate | PO (gavage) | 2 sem | 10 mL kg ⁻¹ day ⁻¹ | Well tolerated | |
| Rabbit | IV | ADME | 0.8 mL kg ⁻¹ | Well tolerated | |
| | PO | ADME | 1.6 mL kg ⁻¹ | Well tolerated | |

TABLE E.65 Methylcellulose

| Species | Route | Duration | Dose | Adverse Reactions/Toxicity | Notes |
|------------------|--------------|----------------------------------|---|---|---|
| Dog | IV | | 40 mL | Anemia, decreased WBC, increased sedimentation rate in 24 h | 0.7–2.8% solution |
| | PO | 14 days | 5.0 mL kg ⁻¹ QD | None | GLP; age 6 months; 6♂/6♀ |
| | PO (dietary) | 90 days | 6% | Well tolerated | |
| | PO (gavage) | 14 doses | 10 mL kg ⁻¹ dose ⁻¹ | Well tolerated | 0.5% |
| | PO (gavage) | 39 weeks | 5 mL kg ⁻¹ QD | Soft/mucoid feces | 0.5% solution (400 cps) in DI water; age 6–10.5 months; ♂/♀ |
| Guinea pig | PO | 12 doses | 4 mL kg ⁻¹ | Well tolerated | 0.5% |
| | Topical | 3 weeks | 0.4 mL once per week | None | 0.5% solution (400 cps) in DI water; age 2–3 months; ♂/♀ |
| Mouse | PO | 90 days | 10 mL kg ⁻¹ | Well tolerated | 0.5% |
| Nonhuman primate | IV | 30 min infusion in a single dose | 1 mL kg ⁻¹ | None | 0.5% solution (400 cps) in DI water; ♂/♀ |
| | PO (gavage) | 14 doses | 5 mL kg ⁻¹ | Well tolerated | 0.5% |
| | PO | 1 month | 10 mL kg ⁻¹ dose ⁻¹ | Well tolerated | 0.5% |
| | PO (gavage) | 28 doses | 5 mL kg ⁻¹ dose ⁻¹ | Well tolerated | 1% |
| | PO (gavage) | 28 doses | 10 mL kg ⁻¹ | Well tolerated | 0.1% |
| Rabbit | PO (gavage) | 12 doses | 4 mL kg ⁻¹ | Well tolerated | 0.5% |

(Continued)

TABLE E.65 (Continued)

| Species | Route | Duration | Dose | Adverse Reactions/Toxicity | Notes |
|---------|--------------|-------------|--|----------------------------|---|
| Rat | PO (dietary) | 90 days | 10% | Well tolerated | |
| | PO (gavage) | | 1020 mg kg ⁻¹ | NOAEL | |
| | PO (gavage) | Single dose | 10 mL kg ⁻¹ | None | 1% |
| | PO (gavage) | Single dose | 10 mL kg ⁻¹ | None | 0.5% |
| | PO | Single dose | 10 mL kg ⁻¹ | Well tolerated | 2% |
| | PO | 3 days | 10 mL kg ⁻¹ QD | None | 1% solution in water; non-GLP; age 11 weeks; 5♂ |
| | PO (gavage) | 5 doses | 10 mL kg ⁻¹ | None | 0.5% |
| | PO (gavage) | 9 doses | 10 mL kg ⁻¹ day ⁻¹ | None | 0.5% |
| | PO (gavage) | 14 days | 10 mL kg ⁻¹ QD | Well tolerated | 1% |
| | PO | 1 month | 10 mL kg ⁻¹ | Well tolerated | 0.5% |
| | PO | 1 month | 5 mL kg ⁻¹ | Well tolerated | 0.5% |
| | PO (gavage) | 28 doses | 5 mL kg ⁻¹ dose ⁻¹ | None | 1% |
| | | 2 years | 120 mg kg ⁻¹ | Well tolerated | 1% |
| | PO (gavage) | 182 days | 10 mL kg ⁻¹ BID | None | 0.5% solution (400 cps) in DI water; age 6 weeks; ♂/♀ |

TABLE E.66 Methylpyrrolidone

| Species | Route | Duration | Dose | Adverse Reactions/Toxicity | Notes |
|---------|---------------------|---------------|---|---|--|
| Dog | IV | Single dose | 0.25 mL kg ⁻¹ | Well tolerated | 50% solution |
| Mouse | IV | Range finding | MTD: 1.3 g kg ⁻¹ ; NOEL: 257 mg kg ⁻¹ | | Published LD ₅₀ = 54–36 000 mg kg ⁻¹ |
| | IV (into tail vein) | Acute | 257 mg kg ⁻¹ (dose vol. 5 mL kg ⁻¹) | NOEL | 5% solution |
| | IV (into tail vein) | Acute | 1285 mg kg ⁻¹ (dose vol. 5 mL kg ⁻¹) | MTD; struggling and vocalization at dosing; rapid breathing, stiff tail, and splayed limbs immediately postdose. Hypoactivity for up to 15 min is typical | 25% solution |

TABLE E.67 Mineral Oil

| Species | Route | Duration | Dose | Adverse Reactions/Toxicity | Notes |
|------------|---------|---------------------------------|--|----------------------------|----------------------------|
| Cat | Topical | 2× per week×2 doses for 28 days | 1.15 mL total: 0.35 mL, then 0.4 mL every 60 min after initial application for 2 doses | None | GLP; age 54–57 days; 6♂/6♀ |
| | Topical | q14d for 56 days | 2.1 mL total: 0.5 mL, then 0.4 mL every 30 min after initial application for four applications | None | GLP; age 9 weeks; 6♂/6♀ |
| Dog | PO | 1 month | 2.5 mL kg ⁻¹ | Well tolerated | |
| Guinea pig | Topical | 28 days | 0.4 mL q7d | None | GLP; age 6 weeks; 11♀ |
| Mouse | PO | 1 month | 250 mg kg ⁻¹ | Well tolerated | |
| Rat | PO | 1 month | 5 mL kg ⁻¹ | Well tolerated | |

TABLE E.68 Olive Oil

| Species | Route | Duration | Dose (mL kg ⁻¹ QD) | Adverse Reactions/Toxicity | Notes |
|---------|-------------|----------|-------------------------------|----------------------------|------------------|
| Rat | PO (gavage) | 28 days | 10 | Well tolerated | Age 6 weeks; ♂/♀ |

TABLE E.69 Peanut Oil

| Species | Route | Duration | Dose | Adverse Reactions/Toxicity | Notes |
|---------|-------------|-----------|---|----------------------------|-------|
| Rat | PO | 1 month | 10 g kg ⁻¹ | Well tolerated | |
| | PO | 12 months | 10 g kg ⁻¹ | Well tolerated | |
| | PO | 90 days | 10 g kg ⁻¹ | Well tolerated | |
| | PO (gavage) | | 5 mL kg ⁻¹ day ⁻¹ | Well tolerated | |
| | SC | | 2 mL kg ⁻¹ day ⁻¹ | Well tolerated | |

TABLE E.70 PEG 200

| Species | Route | Duration | Dose | Adverse Reactions/Toxicity | Notes |
|------------------|---------------------|----------|--|---|--|
| Nonhuman primate | PO (gavage) | 14 days | 5 mL kg ⁻¹ QD | Soft/watery feces (nonadverse) | Age 2–3.5 years; ♂/♀ |
| Rabbit | IV | Acute | >10 g kg ⁻¹ bw | LD ₅₀ | |
| Rat | IP | | 8–9 g kg ⁻¹ bw | LD ₅₀ | |
| | PO (dietary) | 90 days | 20, 40, 80, 160, and 240 g kg ⁻¹ diet QD (1.5, 3, 6, 12, and 18 g kg ⁻¹ bw QD) | NOAEL = 6 g kg ⁻¹ bw QD | At 12, 18 g kg ⁻¹ bw liver and liver/kidney weights increased (respectively); 5♂/5♀ |
| | PO (drinking water) | 90 days | 4.8 g kg ⁻¹ bw QD | NOAEL | 5♂ |
| | PO (drinking water) | 90 days | 10.9 g kg ⁻¹ bw QD | 66% mortality; decreased body weight gain | 5♂ |

TABLE E.71 PEG 300

| Species | Route | Duration | Dose | Adverse Reactions/Toxicity | Notes |
|------------|---------------------|----------|--|---|--|
| Dog | PO (gavage) | 28 days | 1 mL kg ⁻¹ BID | None | Age 5.5–6 months; ♂/♀ |
| Guinea pig | IV | 1 month | 1 mL kg ⁻¹ | Well tolerated | |
| Mouse | PO (gavage) | ADME | 10 mL kg ⁻¹ day ⁻¹ | Well tolerated | |
| Rabbit | PO | 1 month | 500 g kg ⁻¹ | Well tolerated | |
| | IV | Acute | >10 g kg ⁻¹ bw | LD ₅₀ | |
| Rat | IP | Acute | 16–18 g kg ⁻¹ bw | LD ₅₀ | |
| | IP | Acute | 17 g kg ⁻¹ bw | LD ₅₀ | |
| | IV | Acute | 8 g kg ⁻¹ bw | LD ₅₀ | |
| | IV | Acute | 7.1 g kg ⁻¹ bw | LD ₅₀ | |
| | PO (dietary) | 90 days | 20, 40, 80, 160, and 240 g kg ⁻¹ diet QD (1.5, 3, 6, 12, and 18 g kg ⁻¹ bw QD) | NOAEL = 3 g kg ⁻¹ bw QD | At 6, 12, 18 g kg ⁻¹ bw decreased body weight gain, liver and kidney weight increase, and both decreased body weight gain and increased liver weight, respectively; 5 ♂/5 ♀ |
| | PO (drinking water) | 90 days | 5.4 g kg ⁻¹ bw QD | NOAEL | 5 ♂ |
| | PO (drinking water) | 90 days | 20.5 g kg ⁻¹ bw QD | 66% mortality; decreased body weight gain; liver and kidney changes | 5 ♂ |
| | PO (gavage) | 2 weeks | 7.5 mL kg ⁻¹ once or 2.5 mL kg ⁻¹ TID | Well tolerated | 50% |
| | PO (gavage) | 4 weeks | 5 mL kg ⁻¹ day ⁻¹ | Well tolerated | 50% |
| | PO (gavage) | 28 days | 2 mL kg ⁻¹ BID | None | Age 8 weeks; ♂/♀ |

TABLE E.72 PEG 400

| Species | Route | Duration | Dose | Adverse Reactions/Toxicity | Notes |
|------------|---------------|---------------------------|---|---|--|
| Dog | IV | Single dose | Total dose <2 g kg ⁻¹ | | 100% |
| | IV | 28 days | 2–3 g kg ⁻¹ | Decreased blood pressure and reversible depression in respiration. These symptoms increased at doses of 3 g kg ⁻¹ or greater and eventually resulted in complete respiratory arrest. At necropsy dogs were found to have pulmonary edema and small infarcts in the lungs, but no changes in the heart or kidneys | 100% |
| | IV | | | | NS solution; hemolysis occurs >33% (v/v) concentration |
| Guinea pig | IV | Single dose | 1 mL kg ⁻¹ | Well tolerated | 30% |
| | IV | 30 days, +21-day recovery | 8.45 g kg ⁻¹ QD | Dry mouth and dry nasal mucous membrane; histopathological change in the kidney; reversible cloudy swelling of kidney cells and increased glomerular volume | NS solution 25% (v/v) |
| | IV | 30 days, +21-day recovery | 5 mL kg ⁻¹ day ⁻¹ | Dry mouth and dry nasal mucous membrane | NS solution 25% (v/v) |
| | IV | 30 days, +21-day recovery | 4.23 g kg ⁻¹ QD | Increased values of electrolytes (Na ⁺ and Cl ⁻) | NS solution 25% (v/v) |
| | IV (bolus) | | 1 mL kg ⁻¹ | Well tolerated | 60% solution in water pH 3–11 |
| | IV (infusion) | | 2 mL kg ⁻¹ | Well tolerated | 60% solution in water pH 3–11 |
| | IV (infusion) | | 0.5 mL kg ⁻¹ | Well tolerated | 80% solution in water pH 3–11 |
| | PO (dietary) | 1 year | 20 g kg ⁻¹ diet QD (0.5 g kg ⁻¹ bw QD) | NOAEL=0.5 g kg ⁻¹ bw QD | 3♂/1♀ |
| | PO (gavage) | 28 days | 0.25 mL kg ⁻¹ QD | Soft/watery feces (nonadverse) | Age 5–5.5 months; ♂/♀ |
| | PO | 1 month | 1000 mg kg ⁻¹ | Well tolerated | |
| Minipig | Dermal | 2 weeks | 2.5 mL kg ⁻¹ | Well tolerated | |
| | Topical | 90 days | 2 mL kg ⁻¹ QD | Mild dose site inflammation after >30 days of administration | GLP; age 4–6 months; 6♂/6♀ |

(Continued)

TABLE E.72 (Continued)

| Species | Route | Duration | Dose | Adverse Reactions/Toxicity | Notes |
|------------------|---------------------|---------------|---|---|--|
| Mouse | IP | 3 days | 10 mL kg ⁻¹ | Well tolerated | 35% |
| | IP | 1 month | 2.5 mL kg ⁻¹ | Well tolerated | 40% solution |
| | IP | 28 doses | 500 mg kg ⁻¹ | Well tolerated | |
| | IP | | 14.5 g kg ⁻¹ bw | LD ₅₀ | |
| | IP | | 9.2 g kg ⁻¹ bw | LD ₅₀ | |
| | IV | | 8.6 g kg ⁻¹ bw | LD ₅₀ | |
| | IV | Range finding | MTD: 4.5 g kg ⁻¹ ; NOEL: 1.7 g kg ⁻¹ | | Published LD ₅₀ = 8.6–9.7 g kg ⁻¹ |
| | IV (into tail vein) | Acute | 1692 mg kg ⁻¹ (dose vol. 5 mL kg ⁻¹) | NOEL | 30% solution |
| | IV (into tail vein) | Acute | 4512 mg kg ⁻¹ (dose vol. 5 mL kg ⁻¹) | MTD; tremors, ventral recumbency and splayed limbs shortly after dosing; hypoactivity for up to 12 min is typical | 80% solution |
| | PO (gavage) | 4 weeks | 10 mL kg ⁻¹ day ⁻¹ | Well tolerated | |
| | PO | 13 weeks | 5 mL kg ⁻¹ BID | Well tolerated | 100%; CD-1 (Harlan); age ~7 weeks at study initiation; ♂/♀ |
| Nonhuman primate | PO (gavage) | 28 days | 1 mL kg ⁻¹ QD | Soft/watery feces (nonadverse) | In DI water; age 1–2 years; ♂/♀ |
| | PO (gavage) | 28 days | 5 mL kg ⁻¹ QD | Soft/watery feces (nonadverse) | In DI water; age 1–2 years; ♂/♀ |
| Rabbit | IV | | >10 g kg ⁻¹ bw | LD ₅₀ | |
| Rat | Dermal | 13 weeks | 2.5 mL kg ⁻¹ day ⁻¹ | Well tolerated | |
| | Dermal | 104 weeks | 2.5 mL kg ⁻¹ day ⁻¹ | Well tolerated | |
| | IP | | 14.7 g kg ⁻¹ bw | LD ₅₀ | |
| | IP | | 12.3 g kg ⁻¹ bw | LD ₅₀ | |
| | IP | 1 month | 5 mL kg ⁻¹ | Well tolerated | 35% |
| | IV | | 4.7 g kg ⁻¹ bw | LD ₅₀ | |
| | IV | Single dose | 0.5 mL kg ⁻¹ | Well tolerated | |
| | IV (bolus) | | 2 mL kg ⁻¹ | Well tolerated | 60% solution in water pH 3–11 |
| | IV (infusion) | | 5 mL kg ⁻¹ | Well tolerated | 60% solution in water pH 3–11 |

| | | | | |
|---------------------|-------------|--|--|--|
| PO (gavage) | | 2 mL kg ⁻¹ | Well tolerated | |
| PO (gavage) | Single dose | 5 mL kg ⁻¹ | Well tolerated | |
| PO (gavage) | 10 doses | 1.67 mg kg ⁻¹ | Well tolerated | |
| PO | 1 month | 5 mL kg ⁻¹ | Well tolerated | |
| PO (gavage) | 4 weeks | 5 mL kg ⁻¹ day ⁻¹ | Well tolerated | |
| PO (gavage) | 28 days | 3 mL kg ⁻¹ QD | None | Age 6 weeks; ♂/♀ |
| PO (gavage) | 28 days | 10 mL kg ⁻¹ QD | None | In DI water; age 6 weeks; ♂/♀ |
| PO | 13 weeks | 10 mL kg ⁻¹ day ⁻¹ | Loose feces and decreased food consumption; increased water consumption; increases in relative kidney weights; reversible renal toxicity | |
| PO | 26 weeks | 5 mL kg ⁻¹ BID | Well tolerated | 100%; Sprague–Dawley (Harlan) age ~8 weeks at initiation; ♂/♀ |
| PO (dietary) | 90 days | 20, 40, 80, 160, and 240 g kg ⁻¹ diet QD (1.5, 3, 6, 12, and 18 g kg ⁻¹ bw QD) | NOAEL = 6 g kg ⁻¹ bw QD | At 12, 18 g kg ⁻¹ bw decreased body weight gain and liver and kidney weight increase, respectively; 5♂/5♀ |
| PO (dietary) | 2 years | 10, 20, 40, and 80 g kg ⁻¹ diet QD (0.75, 1.5, 3, and 6 g kg ⁻¹ bw QD) | NOAEL = 1.5 g kg ⁻¹ bw QD | From 3 g kg ⁻¹ bw decreased body weight gain (male); 20♂/20♀ |
| PO (drinking water) | 90 days | 4.8 g kg ⁻¹ bw QD | NOAEL | 5♂ |
| PO (drinking water) | 90 days | 16.4 g kg ⁻¹ bw QD | 66% mortality, decreased body weight gain | 5♂ |

TABLE E.73 PEG 600

| Species | Route | Duration | Dose | Adverse Reactions/Toxicity | Notes |
|---------|--------------|----------|--|------------------------------------|--|
| Rat | IP | | 14.1 g kg ⁻¹ bw | LD ₅₀ | |
| | IV | | 7.7 g kg ⁻¹ bw | LD ₅₀ | |
| | PO (dietary) | 90 days | 20, 40, 80, 160, and 240 g kg ⁻¹ diet QD (1.5, 3, 6, 12, and 18 g kg ⁻¹ bw QD) | NOAEL = 6 g kg ⁻¹ bw QD | From 12 g kg ⁻¹ bw decreased body weight gain and increased kidney weights; 5♂/5♀ |

TABLE E.74 PEG 810

| Species | Route | Duration | Dose (g kg ⁻¹ bw) | Adverse Reactions/Toxicity | Notes |
|---------|-------|----------|------------------------------|----------------------------|-------|
| Rat | IV | | 13 | LD ₅₀ | |
| | SC | | 16 | LD ₅₀ | |

TABLE E.75 PEG 1000

| Species | Route | Duration | Dose | Adverse Reactions/Toxicity | Notes |
|---------|--------------|----------|--|------------------------------------|---|
| Mouse | IP | | 2 g kg ⁻¹ bw | LD ₅₀ | |
| Rabbit | IV | Acute | >10 g kg ⁻¹ bw | LD ₅₀ | |
| Rat | IP | Acute | 15.6 g kg ⁻¹ bw | LD ₅₀ | |
| | PO (dietary) | 90 days | 20, 40, 80, 160, and 240 g kg ⁻¹ diet QD (1.5, 3, 6, 12, and 18 g kg ⁻¹ bw QD) | NOAEL = 6 g kg ⁻¹ bw QD | From 12 g kg ⁻¹ bw decreased body weight gain; 5♂/5♀ |

TABLE E.76 PEG 1500

| Species | Route | Duration | Dose | Adverse Reactions/Toxicity | Notes |
|---------|---------------------|----------|--|---|---|
| Rat | IP | | 17.7 g kg ⁻¹ bw | LD ₅₀ | |
| | IV | | 8.5 g kg ⁻¹ bw | LD ₅₀ | |
| | PO (dietary) | 90 days | 0.88, 4.05, 8.1, and 22.9 g kg ⁻¹ bw QD | NOAEL = 2 g kg ⁻¹ bw QD | From 4.05 g kg ⁻¹ bw kidney damage |
| | PO (drinking water) | 90 days | 20, 40, 80, 160, and 240 g kg ⁻¹ diet QD (1.5, 3, 6, 12, and 18 g kg ⁻¹ bw QD) | NOAEL = 3 g kg ⁻¹ bw QD | From 6 g kg ⁻¹ decreased body weight gain; at 18 g kg ⁻¹ bw increased kidney weights; 5♂/5♀ |
| | PO (drinking water) | 2 years | 0.2, 0.8, 4, and 20 g L ⁻¹ QD (0.015, 0.059, 0.27, and 1.69 g kg ⁻¹ bw QD) | 1.69 g kg ⁻¹ bw QD: no effects on fertility, survival, hematology, or histopathology | 8♂/8♀ |

TABLE E.77 PEG 1540

| Species | Route | Duration | Dose | Adverse Reactions/Toxicity | Notes |
|---------|--------------|----------|---|--------------------------------------|---|
| Dog | PO (dietary) | 1 year | 20 g kg ⁻¹ diet QD (0.5 g kg ⁻¹ bw QD) | NOAEL = 0.5 g kg ⁻¹ bw QD | |
| Rabbit | IV | Acute | >10 g kg ⁻¹ bw | LD ₅₀ | |
| Rat | IP | Acute | 15.4 g kg ⁻¹ bw | LD ₅₀ | |
| | PO (dietary) | 90 days | 20, 40, 80, 160, and 240 g kg ⁻¹ diet QD (1.5, 3, 6, 12, and 18 g kg ⁻¹ bw QD) | NOAEL = 3 g kg ⁻¹ bw QD | From 6 g kg ⁻¹ bw decreased body weight gain; from 18 g kg ⁻¹ bw increased kidney weights; 5♂/5♀ |
| | PO (dietary) | 2 years | 0.2, 0.8, 4, 20, 40, and 80 g kg ⁻¹ diet QD (0.015, 0.06, 0.3, 1.5, 3, and 6 g kg ⁻¹ bw QD) | NOAEL = 3 g kg ⁻¹ bw QD | From 6 g kg ⁻¹ QD cloudy swelling in the liver; 35♂/35♀ |

TABLE E.78 PEG 4000

| Species | Route | Duration | Dose | Adverse Reactions/Toxicity | Notes |
|---------|---------------------|------------------------------|---|---|---|
| Dog | PO (dietary) | 1 year | 20 g kg ⁻¹ diet QD (0.5 g kg ⁻¹ bw QD) | NOAEL = 0.5 g kg ⁻¹ bw QD | |
| Mouse | IP | | 8.0 g kg ⁻¹ bw | LD ₅₀ | |
| Rabbit | IV | Acute | >10 g kg ⁻¹ bw | LD ₅₀ | |
| | PO (gavage) | 5 weeks (6 days per week) | 5, 10, 20 g kg ⁻¹ bw QD | From 5 g kg ⁻¹ bw decreased body weight gain, decreased glycogen storage; from 20 g kg ⁻¹ bw decreased body weights | |
| Rat | IP | Acute | 11.6–13 g kg ⁻¹ bw | LD ₅₀ | |
| | IP | | 9.7 g kg ⁻¹ bw | LD ₅₀ | |
| | IV | | 7.5 g kg ⁻¹ bw | LD ₅₀ | |
| | PO (dietary) | 90 days | 20, 40, 80, 160, and 240 g kg ⁻¹ diet QD (1.5, 3, 6, 12, and 18 g kg ⁻¹ bw QD) | NOAEL = 3 g kg ⁻¹ bw QD | From 6 g kg ⁻¹ bw decreased body weight gain; from 12 g kg ⁻¹ bw increased kidney weights; 5♂/5♀ |
| | PO (dietary) | 90 days | 1.6 g kg ⁻¹ bw QD | NOAEL = 1.6 g kg ⁻¹ bw QD | |
| | PO (dietary) | 2 years | 0.375, 0.75, 1.5, 3, and 6 g kg ⁻¹ bw QD | NOAEL = 3 g kg ⁻¹ bw QD | At 6 g kg ⁻¹ bw decreased body weight gain; 20♂/20♀ |
| | PO (drinking water) | 90 days | 0.04–19 g kg ⁻¹ bw QD | NOAEL = 0.8 g kg ⁻¹ bw QD | At 0.23 g kg ⁻¹ bw degeneration of the testis tubules, degenerated sperm; from 7 g kg ⁻¹ bw decreased body weight gain; at 19 g kg ⁻¹ bw kidney damage; 5♂ |
| | PO (drinking water) | 2 years | 0.00085, 0.0036, 0.017, and 0.062 g kg ⁻¹ bw QD | NOAEL = 0.062 g kg ⁻¹ bw QD | 8♂/8♀ |

TABLE E.79 PEG 6000

| Species | Route | Duration | Dose | Adverse Reactions/ Toxicity | Notes |
|---------|--------------|----------|--|-------------------------------------|---|
| Rabbit | IV | Acute | >10 g kg ⁻¹ bw | LD ₅₀ | |
| Rat | IP | Acute | 6.8 g kg ⁻¹ bw | LD ₅₀ | |
| | PO (dietary) | 90 days | 20, 40, 80, 160, and 240 g kg ⁻¹ diet QD (1.5, 3, 6, 12, and 18 g kg ⁻¹ bw QD) | NOAEL = 12 g kg ⁻¹ bw QD | At 18 g kg ⁻¹ bw kidney weights increased, decreased body weight gain; 5♂/5♀ |

TABLE E.80 PEG 10000

| Species | Route | Duration | Dose | Adverse Reactions/Toxicity | Notes |
|---------|--------------|----------|------------------------------|-------------------------------|-------|
| Rat | IP | Acute | 12.6 g kg ⁻¹ bw | LD ₅₀ | |
| | PO (dietary) | 90 days | 1.6 g kg ⁻¹ bw QD | NOAEL | |

TABLE E.81 PEG 4000000

| Species | Route | Duration | Dose | Adverse Reactions/Toxicity | Notes |
|---------|--------------|----------|---------------------------------------|---|---------|
| Rat | IV | Acute | >10 g kg ⁻¹ bw | LD ₅₀ | |
| | PO (dietary) | 90 days | 8.0 and 18.4 g kg ⁻¹ bw QD | From 8 g kg ⁻¹ bw cloudy swelling in the renal tubules; at 18.4 g kg ⁻¹ bw decreased body weight gain (males), decreased relative liver weights | 10♂/10♀ |
| | PO (dietary) | 2 years | Up to 2.76 g kg ⁻¹ bw QD | NOAEL = 2.76 g kg ⁻¹ bw QD | 36♂/36♀ |
| | PO (dietary) | 2 years | Up to 0.56 g kg ⁻¹ bw QD | NOAEL = 0.56 g kg ⁻¹ bw QD | 4♂/2♀ |

TABLE E.82 Petrolatum

| Species | Route | Duration | Dose (g kg ⁻¹) | Adverse Reactions/Toxicity | Notes |
|---------|--------|----------|----------------------------|----------------------------|-------|
| Rabbit | Dermal | 1 month | 1 | Well tolerated | |

TABLE E.83 Phosphate-Buffered Saline (PBS)

| Species | Route | Duration | Dose | Adverse Reactions/ Toxicity | Notes |
|------------------|----------------------|--|--|--------------------------------|---|
| Dog | IV (slow bolus) | 28 days | 2.5 mL kg ⁻¹ dose ⁻¹ (5 mL kg ⁻¹ day ⁻¹) BID | Well tolerated | pH 7.2; age 5–6 months; ♂/♀ |
| | 1 h infusion | 1 h infusion single dose | 10 mL kg ⁻¹ | Well tolerated | 0.01 M PBS (powder) in sterile water for injection; beagle dog; age 5 months; ♂/♀ |
| | PO (gavage) | 28 days | 5 mL kg ⁻¹ QD | Well tolerated | pH 7.2; age 5–6 months; ♂/♀ |
| | Topical | q7d × 2 doses for 7 weeks | 1.4 mL kg ⁻¹ divided into 3 doses 60 min apart | Well tolerated | Non-GLP; age 7 weeks; 2♂/2♀ |
| Minipig | IV | 14 days | 5 mL kg ⁻¹ QD | Well tolerated | GLP; pH 6.0; age 17–21 days; 5♂/5♀ |
| | IV | 14 days | 1 mL kg ⁻¹ QD | Well tolerated | GLP; pH 6.0; age 5–8 months; 5♂/5♀ |
| | PO | 7 days | 5 mL kg ⁻¹ BID | Well tolerated | GLP; age 4 days; 4♂/4♀ |
| | PO | 28 days | 5 mL kg ⁻¹ BID | Well tolerated | GLP; pH 6.0; age 4–5 days; 10♂/10♀ |
| Mouse | SC (infusion) | Continuous (24 h day ⁻¹) infusion for 7 days | 1.0 mL h ⁻¹ (+0.15 mL h ⁻¹) | Well tolerated | pH 7.2; age 10 weeks; ♂/♀ |
| | SC (bolus) | Every 2 days for 1 week, then weekly for 26 weeks | 10–11.83 mL kg ⁻¹ | Well tolerated | pH 7.2; age 6 weeks; ♂/♀ |
| | SC | 6 months | 10 mL kg ⁻¹ | Well tolerated | |
| Nonhuman primate | IV (slow bolus) | 28 days | 10 mL kg ⁻¹ QD | Well tolerated | pH 7.2; age 2–3 years; ♂/♀ |
| | PO (gavage) | 2 weeks | 10 mL kg dose ⁻¹ | Well tolerated | |
| | PO (gavage) | 2 weeks | 1.6 mL kg ⁻¹ | Well tolerated | |
| | SC | 1 week | 0.2 mL kg ⁻¹ | Well tolerated | |
| | SC | 9 months | 1 mL kg ⁻¹ | Well tolerated | |
| Rabbit | IV | 12 weeks | 1.0 mL kg ⁻¹ q7d | Well tolerated | Non-GLP; pH 6.5; age 7–8 months; 9♂ |
| Rat | IV | Single dose | 1 mL kg ⁻¹ | Well tolerated | |
| | IV (slow bolus) | 28 days | 5 mL kg ⁻¹ once weekly | Well tolerated | pH 7.2; age 6 weeks; ♂/♀ |
| | PO (gavage) | 5 doses | 10 mL kg ⁻¹ | Well tolerated | |
| | PO | 28 doses | 10 mL kg ⁻¹ | Well tolerated | |
| | PO (gavage) | 28 days | 10 mL kg ⁻¹ dose ⁻¹ QD | Well tolerated | pH 7.2; age 6 weeks; ♂/♀ |
| | SC | 1 month | 1 mL kg ⁻¹ QD | Well tolerated | |
| | Slow bolus injection | 11 doses | 1 mL kg ⁻¹ | Well tolerated | |

TABLE E.84 Poloxamer 188™

| Species | Route | Duration | Dose | Adverse Reactions/Toxicity | Notes |
|---------|-------|----------|--------------------------|--|---|
| Dog | SC | 4 weeks | 5 mL kg ⁻¹ QD | Dog vehicle changed following a single administration due to animal distress in response to the injections. The distress was attributed to the vehicle | 2% in DI water; pH 5 ± 0.2; age 5–6 months; ♂/♀ |
| Mouse | PO | 1 month | 10 mL kg ⁻¹ | Well tolerated | 5% solution |
| Rat | PO | 1 month | 10 mL kg ⁻¹ | Well tolerated | 7.5% solution |
| | SC | 4 weeks | 5 mL kg ⁻¹ QD | None | 2% in DI water; pH 5 ± 0.2; age 6 weeks; ♂/♀ |

TABLE E.85 Poly(glycolide-co-DL-Lactide) Microspheres

| Species | Route | Duration | Dose (mg per pocket SD) | Adverse Reactions/Toxicity | Notes |
|---------|--------------------------|----------|-------------------------|----------------------------|---------------------|
| Dog | Into periodontal pockets | 28 days | 3.5 | Well tolerated | GLP; age 6–10 years |

TABLE E.86 Polyglyceryl Oleate

| Species | Route | Duration | Dose | Adverse Reactions/Toxicity | Notes |
|---------|--------|----------|------|---|-------------|
| Rabbit | Dermal | Acute | | Moderately irritant (irritation and corrosion test) | No dilution |
| | Ocular | Acute | | Slightly irritant (irritation and corrosion test) | No dilution |
| Rat | PO | Acute | | LD ₅₀ ≥ 2005 mg kg ⁻¹ day ⁻¹ | No dilution |

TABLE E.87 Polyvinylpyrrolidone (PVP)

| Species | Route | Duration | Dose | Adverse Reactions/Toxicity | Notes |
|---------|-------|-------------|------|--|-------|
| Dog | PO | Acute | | Causes histamine release in dogs. The reaction is highly variable ranging from no discernible effect to reddening of extremities to total collapse | |
| Rat | IM | Single dose | 1 mL | Well tolerated | 1% |

TABLE E.88 Propylene Glycol

| Species | Route | Duration | Dose | Adverse Reactions/Toxicity | Notes |
|------------------|---------------------|-------------------|--|--|---|
| Dog | IV | Single dose | 1 mL kg ⁻¹ | NOAEL; some emesis and diarrhea | 30% |
| | IV | 14 days | 5 mL kg ⁻¹ at a rate of 0.3 mL kg ⁻¹ | Hemolysis <i>in vitro</i> dog blood | 60% solution in water |
| | IV (bolus) | | 1 mL kg ⁻¹ | Well tolerated | 40% solution in water pH 3–11; 50% PG causes hemolysis |
| | IV (infusion) | | 2 mL kg ⁻¹ | Well tolerated | 40% solution in water pH 3–11; 50% PG causes hemolysis |
| | PO | 28 days | 1.5 mL kg ⁻¹ | Well tolerated | 100% |
| | PO | 1 month | 2.5 mL kg ⁻¹ | Well tolerated | |
| | PO (gavage) | Up to 7 days | 2 mL kg ⁻¹ day ⁻¹ | Well tolerated | |
| | PO (gavage) | 90 days | Dose vol. 5 mL kg ⁻¹ , daily dose 1000 mg kg ⁻¹ (20% w/v) | Well tolerated | Dose concentration 200 mg mL ⁻¹ |
| Minipig | Dermal | 26 weeks | 2.5 mL kg ⁻¹ | Well tolerated | |
| Nonhuman primate | PO (gavage) | 90 days | Dose vol. 5 mL kg ⁻¹ , daily dose 1000 mg kg ⁻¹ (20% w/v) | Well tolerated | Dose concentration 200 mg mL ⁻¹ |
| Mouse | IP | 1 month | 2.5 mL kg ⁻¹ | Well tolerated | 40% solution |
| | IV | Range finding | MTD: 1.5 g kg ⁻¹ ; NOEL: 1 g kg ⁻¹ | | Published LD ₅₀ = 5.0–8.6 g kg ⁻¹ |
| | IV (into tail vein) | Acute | 1036 mg kg ⁻¹ (dose vol. 5 mL kg ⁻¹) | NOEL | 20% solution |
| | IV (into tail vein) | Acute | 1554 mg kg ⁻¹ (dose vol. 5 mL kg ⁻¹) | MTD; tremors and hind limb ataxia; full recovery by 1 min is typical | 30% solution |
| | PO | 1 month | 10 mL kg ⁻¹ | Well tolerated | 50% solution |
| | PO (gavage) | 90 days | Dose vol. 10 mL kg ⁻¹ , daily dose 1000 mg kg ⁻¹ (20% w/v) | Well tolerated | Dose concentration 100 mg mL ⁻¹ |
| Rat | IV (bolus) | | 2 mL kg ⁻¹ | Well tolerated | 40% solution in water pH 3–11; 50% PG causes hemolysis |
| | IV (infusion) | | 5 mL kg ⁻¹ | Well tolerated | 40% solution in water pH 3–11; 50% PG causes hemolysis |
| | PO | 1 month | 2.5 mL kg ⁻¹ | Well tolerated | |
| | PO (gavage) | Prelim/segment II | 5 mL kg ⁻¹ day ⁻¹ | Well tolerated | 60/40: purified water/propylene glycol: w/w |
| | PO (gavage) | 2 weeks | 5 mL kg ⁻¹ day ⁻¹ | Well tolerated | 60/40: purified water/propylene glycol: w/w |
| | PO (gavage) | Segment II | 5 mL kg ⁻¹ day ⁻¹ | Well tolerated | 60/40: purified water/propylene glycol: w/w |
| | PO (gavage) | 2 weeks | 2 mL kg ⁻¹ day ⁻¹ | Well tolerated | |
| | PO (gavage) | 90 days | Dose vol. 5 mL kg ⁻¹ , daily dose 1000 mg kg ⁻¹ (20% w/v) | Well tolerated | Dose concentration 200 mg mL ⁻¹ |
| | SC | 4 weeks | 2.5 mL kg ⁻¹ day ⁻¹ | Well tolerated | |

TABLE E.89 Randomly Methylated β -Cyclodextrins (RAMEB)

| Species | Route | Duration | Dose | Adverse Reactions/Toxicity | Notes |
|------------------|------------|----------|---|----------------------------|-------|
| Nonhuman primate | Intranasal | 1 month | 82.8 mg mL ⁻¹ (with treatment) 74.7 mg mL ⁻¹ (placebo) TID | Well tolerated | 7.5% |

TABLE E.90 Safflower Oil

| Species | Route | Duration | Dose | Adverse Reactions/Toxicity | Notes |
|---------|-------|-------------|------|---|-------|
| Dog | SC | Single dose | 1 mL | Well tolerated, no evidence of irritation macroscopically or histologically | 100% |

TABLE E.91 Saline (pH Adj.)

| Species | Route | Duration | Dose (mL kg ⁻¹ SD) | Adverse Reactions/Toxicity | Notes |
|---------|-------|----------|-------------------------------|----------------------------|---------------------------------------|
| Mouse | IM | 14 days | 4 | None | (pH 4.5); non-GLP; age 6 weeks; 8♂/8♀ |

TABLE E.92 Sesame Oil

| Species | Route | Duration | Dose | Adverse Reactions/Toxicity | Notes |
|---------|-------------|----------|--------------------------|----------------------------|---------------------|
| Dog | PO | 1 month | 5 mL kg ⁻¹ | Well tolerated | |
| | PO (gavage) | 28 days | 5 mL kg ⁻¹ QD | None | Age 7–8 months; ♂/♀ |
| | PO (gavage) | 9 months | 1 mL kg ⁻¹ QD | None | Age 7–8 months; ♂/♀ |
| Mouse | PO | 1 month | 0.25 mL kg ⁻¹ | Well tolerated | |
| Rabbit | PO | 1 month | 0.5 mL kg ⁻¹ | Well tolerated | |
| Rat | PO | 1 month | 1 mL kg ⁻¹ | Well tolerated | |
| | PO (gavage) | 26 weeks | 1 mL kg ⁻¹ QD | None | Age 6 weeks; ♂/♀ |

TABLE E.93 Sodium Acetate Trihydrate Buffer

| Species | Route | Duration | Dose (mL kg ⁻¹) | Adverse Reactions/Toxicity | Notes |
|------------------|-------|----------|-----------------------------|----------------------------|-------|
| Nonhuman primate | IV | | 1 | Well tolerated | |

TABLE E.94 Sodium Chloride

| Species | Route | Duration | Dose | Adverse Reactions/ Toxicity | Notes |
|---------|---------------|-------------|--|--|---|
| Cat | SC | 9 days | 0.1 mL kg ⁻¹ QD | None | 0.9% saline; non-GLP; age 4 months; 2♂/2♀ |
| | SC | 9 days | 0.1 mL kg ⁻¹ QD | None | Non-GLP; 2♂/2♀ |
| | IV | 8 days | 1.9 mL kg ⁻¹ SD | None | 0.9% saline; non-GLP; age 8–9 months; 2♂/2♀ |
| | IV | 19 days | 1.32 mL kg ⁻¹ 3 × per week × 2 week | None | 0.9% saline; non-GLP; age 8–9 months; 2♂/2♀ |
| Dog | IV | Single dose | 10 mL kg ⁻¹ | Well tolerated | 0.9% |
| | IV | Single dose | 2 mL kg ⁻¹ | Well tolerated | 0.9% |
| | IV | 2 weeks | 5 mL kg ⁻¹ day ⁻¹ | Well tolerated | 0.9% |
| | IV (bolus) | Single dose | 0.3 mL kg ⁻¹ | Well tolerated | 0.9% |
| | IV | 14 days | 2.5 mL kg ⁻¹ QD | None | 0.9% saline and water to make 0.8% saline; GLP; age 8 weeks; 10♂/10♀ |
| | IV | 14 days | 2 mL kg ⁻¹ QD | None | 0.9% saline and water to make 0.8% saline; GLP; age 10–11 months; 12♂/12♀ |
| | IV (infusion) | 8 weeks | 1 mL kg ⁻¹ 30 min infusion 3 × per week | None | Non-GLP; age 8–15 months; 2♀ |
| | IV (infusion) | 48 h | 2 mL kg ⁻¹ 20 min infusion SD | None | 0.9% saline; non-GLP; age 8–13 months; 2♂/2♀ |
| | IV (infusion) | 4 h | 2 mL kg ⁻¹ 20 min infusion SD | None | 0.9% saline; non-GLP; age 8–9 months; 1♀ |
| | Ocular | 6 months | 1 drop per eye TID | None | 0.9% saline; GLP; age 6–7 months; 4♂/4♀ |
| | PO | Single dose | 0.282 mL kg ⁻¹ | Well tolerated | 0.9% |
| | SC | 1 month | 0.025 mL | NOEL | 0.9% |
| | SC | 6 months | 0.5 mL kg ⁻¹ total: 0.1 mL kg ⁻¹ at five separate locations q21d | Occasional transient injection site erythema | 0.9% saline; GLP; age 5–6 months; 4♂/4♀ |
| Minipig | SC | 28 days | 0.0225 mL kg ⁻¹ QD | None | 0.9% saline; GLP; age 3 months; 4♂/4♀ |
| | IM | 49 days | 0.5 mL q28d × 2 doses | None | 0.9% saline; non-GLP; age 4 months; 3♂ |
| | ID | 49 days | 0.5 mL q28d × 2 doses | None | 0.9% saline; non-GLP; age 4 months; 3♂ |
| Mouse | IM | 30 doses | | Well tolerated | 0.9% |
| | IV | Single dose | 10 mL kg ⁻¹ | Well tolerated | 0.9% |
| | IV | 2 days | 10 mL kg ⁻¹ SD | None | 0.9% saline and water to make 0.8% saline; GLP; age 4–7 weeks; 11♂/11♀ |
| | SC | Single dose | 10 mL kg ⁻¹ | Well tolerated | 0.9% |
| | Topical | 7 days | 0.10 mL QD | None | 0.9% saline; non-GLP; age 8 weeks; 28♂ |

(Continued)

TABLE E.94 (Continued)

| Species | Route | Duration | Dose | Adverse Reactions/ Toxicity | Notes |
|------------------|--------------|-------------|--|--------------------------------|--|
| Nonhuman primate | SC | 28 doses | 0.67 mL kg ⁻¹ | Well tolerated | 0.9% |
| | SC | 56 doses | 0.5 mL kg ⁻¹ dose ⁻¹ | Well tolerated | 0.9% |
| | Slow bolus | 9 doses | 10 mL kg ⁻¹ dose ⁻¹ | None | 0.9% |
| Rabbit | IM | 33 days | 0.2 mL q14d × 3 doses | None | 0.9% saline; GLP; age 6 months; 6♀ |
| | IV | Single dose | 0.1 mL kg ⁻¹ | Well tolerated | 0.9% |
| | Perivascular | Single dose | 0.1 mL kg ⁻¹ | Well tolerated | 0.9% |
| | SC | Single dose | 0.5 mL kg ⁻¹ | Well tolerated | 0.9% |
| Rat | IP | 90 days | 10 mL kg ⁻¹ QD | None | 0.9% saline; GLP; age 9 weeks; 10♂/10♀ |
| | IV | 6 h | 4 µL g ⁻¹ SD | None | 0.9% saline; non-GLP; age 8–9 weeks; 2♂/2♀ |
| | IV | Single dose | 1 mL kg ⁻¹ | Well tolerated | 0.9% |
| | IV | Single dose | 2 mL kg ⁻¹ | None | 0.9% |
| | IV | Single dose | 10 mL kg ⁻¹ | Well tolerated | 0.9% |
| | IV | 3 doses | 4 mL kg ⁻¹ | Well tolerated | 0.9% |
| | IV | 3 days | 0.376 mL kg ⁻¹ QD | None | 0.9% saline; GLP; age 4 months; 5♂ |
| | IV | 7 doses | 1 mL kg ⁻¹ | None | 0.9% |
| | IV | 2 weeks | 10 mL kg ⁻¹ day ⁻¹ | Well tolerated | 0.9% |
| | IV infusion | 13 weeks | 10 mL kg ⁻¹ 30 min infusion q7d | None | 0.9% saline; GLP; age 11 weeks; 20♂/20♀ |
| | SC | Single dose | 0.1–0.4 mL | Well tolerated | 0.9% |
| | SC | 14 days | 0.5 mL kg ⁻¹ QD | None | 0.9% saline; GLP; age 71–72 days; 15♂/15♀ |
| | SC | 28 doses | 4 mL kg ⁻¹ | Well tolerated | 0.9% |
| | SC | 56 doses | 2 mL kg ⁻¹ dose ⁻¹ | Well tolerated | 0.9% |
| | Slow bolus | Single dose | 1 mL kg ⁻¹ | Well tolerated | 0.9% |
| | Slow bolus | Single dose | 5 mL kg ⁻¹ | Well tolerated | 0.9% |
| | Slow bolus | Single dose | 10 mL kg ⁻¹ | Well tolerated | 0.9% |
| | Slow bolus | 3 doses | 2 mL dose ⁻¹ | Well tolerated | 0.9% |

TABLE E.95 Sodium Dihydrogen Phosphate Dihydrate (SDPD)

| Species | Route | Duration | Dose (mL kg ⁻¹ QD) | Adverse Reactions/Toxicity | Notes |
|------------------|-------------|----------|-------------------------------|----------------------------|--|
| Mouse | PO (gavage) | 91 days | 10 | None | 0.5 M SDPD in DI water; age 5 weeks; ♂/♀ |
| Nonhuman primate | PO (gavage) | 91 days | 10 | None | 0.5 M SDPD in DI water; age 2.5–3.5 years; ♂/♀ |

TABLE E.96 Sodium Metabisulfite

| Species | Route | Duration (days) | Dose (mL kg ⁻¹ QD) | Adverse Reactions/Toxicity | Notes |
|------------------|-------------|-----------------|-------------------------------|----------------------------|--|
| Mouse | PO (gavage) | 91 | 20 | Well tolerated | 10% in distilled water; age 6 weeks; ♂/♀ |
| Nonhuman primate | PO (gavage) | 91 | 10 | Well tolerated | 10% in distilled water; age 2–4.5 years; ♂/♀ |
| Rat | PO (gavage) | 91 | 10 | Well tolerated | 10% in distilled water; age 6 weeks; ♂/♀ |

TABLE E.97 Sodium Phosphate Buffer

| Species | Route | Duration | Dose | Adverse Reactions/Toxicity | Notes |
|------------------|-------------|----------|---|----------------------------|-----------------------------------|
| Dog | PO | 14 doses | 10 mL kg ⁻¹ dose ⁻¹ | Well tolerated | 70 mM |
| Mouse | PO (gavage) | 90 days | 10 mL kg ⁻¹ QD | Well tolerated | 0.1 M; pH 7.0; age 6 weeks; ♂/♀ |
| Nonhuman primate | PO (gavage) | 90 days | 5 mL kg ⁻¹ QD | Well tolerated | 0.1 M; pH 7.0; age 2–3 years; ♂/♀ |
| | PO (gavage) | 91 days | 5 mL kg ⁻¹ QD | Well tolerated | 0.1 M; pH 7.0; age 2–3 years; ♂/♀ |
| Rat | PO | 2 weeks | 10 mL kg ⁻¹ QD | Well tolerated | 70 mM |
| | PO (gavage) | 90 days | 10 mL kg ⁻¹ QD | None | 0.1 M; pH 7.0; age 6 weeks; ♂/♀ |
| | PO (gavage) | 91 days | 10 mL kg ⁻¹ QD | None | 0.1 M; pH 7.0; age 6 weeks; ♂/♀ |

TABLE E.98 Sodium Sulfite

| Species | Route | Duration | Dose (mL dose ⁻¹ QID) | Adverse Reactions/Toxicity | Notes |
|---------|------------------|----------|----------------------------------|----------------------------|--|
| Rabbit | Ocular (topical) | 28 days | 50 | None | 10% in reverse osmosis DI water; age 7 months; ♂/♀ |

TABLE E.99 Solutol HS15™

| Species | Route | Duration | Dose | Adverse Reactions/Toxicity | Notes |
|---------|-------------|----------|--|---|---|
| Dog | Any | 1 dose | Varies | Poorly tolerated in significant amounts | Solutol® HS15/purified water |
| Mouse | IP | 2 weeks | 10 mL kg ⁻¹ 3× per week | None | 10% Solutol; non-GLP; age 4–5 weeks; 5♂/5♀ |
| | IV | 2 weeks | 10 mL kg ⁻¹ 3× per week | None | 10% Solutol; non-GLP; age 4–5 weeks; 5♂/5♀ |
| Rat | PO (gavage) | 2 months | 10 mL kg ⁻¹ day ⁻¹ | Well tolerated | 10% Solutol HS15 in purified water; Sprague–Dawley rats |

TABLE E.100 Soybean Oil

| Species | Route | Duration | Dose | Adverse Reactions/Toxicity | Notes |
|---------|-------|----------|--------------------------|---|-------|
| Dog | PO | Acute | | Well tolerated | 50% |
| Rat | PO | 13 weeks | 2000 mg kg ⁻¹ | Well tolerated; reduced food consumption at 500 mg kg ⁻¹ and above | |
| | PO | 13 weeks | 20% in diet | Well tolerated | |

TABLE E.101 Sulfobutyl Ether β -Cyclodextrin (SBECD)

| Species | Route | Duration | Dose (mL kg ⁻¹ BID) | Adverse Reactions/Toxicity | Notes |
|---------|-------|----------|--------------------------------|----------------------------|----------------------------------|
| Mouse | PO | 7 days | 10 | None | 10%; non-GLP; age 7 weeks; 3♂/3♀ |

TABLE E.102 Tartaric Acid

| Species | Route | Duration | Dose | Adverse Reactions/Toxicity | Notes |
|---------|-------------|-------------------|---|----------------------------|-------|
| Rabbit | PO (gavage) | Prelim./Segmt. II | 3 mL kg ⁻¹ day ⁻¹ | Well tolerated | |
| | PO (gavage) | Segmt. II | 3 mL kg ⁻¹ day ⁻¹ | Well tolerated | |
| Rat | PO (gavage) | 39 weeks | 0.5 mL kg ⁻¹ | Well tolerated | |
| | PO (gavage) | Sys. nerv | 3 mL kg ⁻¹ day ⁻¹ | Well tolerated | |

TABLE E.103 Terbinafine HClTM Placebo Nail Lacquer

| Species | Route | Duration | Dose (μ L kg ⁻¹) | Adverse Reactions/Toxicity | Notes |
|---------|--------|----------|-----------------------------------|-----------------------------------|-------|
| Pig | Dermal | 1 month | 600 | Erythema, peeling or flaking skin | |

TABLE E.104 **Transectol™**

| Species | Route | Duration | Dose | Adverse Reactions/Toxicity | Notes |
|------------|-------------|-------------------------|---|--|---|
| Cat | IV | 1 month | 2 mL kg ⁻¹ single dose | Well tolerated | No evidence of hemolysis or hematotoxicity |
| | IV | Acute | 1000 mg kg ⁻¹ | LD ₅₀ | |
| | IV | Acute | 3000 mg kg ⁻¹ | LD ₅₀ | |
| Dog | PO | 90 days | 1500 mg kg ⁻¹ day ⁻¹ | NOAEL | |
| | PO (gavage) | Acute (dose escalating) | 500, 1000, 1500, and 2000 mg kg ⁻¹ | MTD > 2000 mg kg ⁻¹ | Undiluted |
| | PO (gavage) | Subacute (7 days DRF) | 0, 500, 1000, and 2000 mg kg ⁻¹ day ⁻¹ | MTD > 2000 mg kg ⁻¹ day ⁻¹ | |
| | PO (gavage) | Subchronic (13 weeks) | 0, 400, 1000, and 2000/1500 mg kg ⁻¹ day ⁻¹ | NOAEL = 1000 mg kg ⁻¹ day ⁻¹ | |
| | PO | Acute | 300 mg kg ⁻¹ | LD ₅₀ | |
| Guinea pig | PO | Acute | 300 mg kg ⁻¹ | LD ₅₀ | |
| Minipig | PO | 90 days | 0, 167, 500, and 1500 mg kg ⁻¹ day ⁻¹ | NOAEL = 167 mg kg ⁻¹ day ⁻¹ | Uremia, death at 1500 mg kg ⁻¹ day ⁻¹ ; high dose reduced to 1000 mg kg ⁻¹ day ⁻¹ after 21 days; histopathologies in doses > 500 mg kg ⁻¹ day ⁻¹ include hydropic degeneration of liver and proximal kidney tubules; at > 1000 mg kg ⁻¹ day ⁻¹ increased relative kidney weight and decreased RBC (males) |

(Continued)

TABLE E.104 (Continued)

| Species | Route | Duration | Dose | Adverse Reactions/Toxicity | Notes |
|---------|-----------------------|---|--|--|---|
| Mouse | IP | Acute | 3900 mg kg ⁻¹ | LD ₅₀ | |
| | IV | Acute | 4300 mg kg ⁻¹ | LD ₅₀ | |
| | IV (bolus, tail vein) | Acute (dose escalating) | 25, 50, 100, 200, 400, 800, 1600, 6400, 3200, and 4800 mg kg ⁻¹ | MTD (IV): 3200 mg kg ⁻¹ | Males; physiological saline solution |
| | IV (bolus, tail vein) | Acute (dose escalating) | 25, 50, 100, 200, 400, 800, 1600, 8000, 6400, 4800, and 3200 mg kg ⁻¹ | MTD (IV): 3200 mg kg ⁻¹ | Females; physiological saline solution |
| | PO | Acute | 6.6 g kg ⁻¹ | Tested toxic | |
| | PO | Acute | 7250 mg kg ⁻¹ | LD ₅₀ | |
| | PO (gavage) | Developmental (dosed GD 7–14, littered and reared to PND 3) | 5500 mg kg ⁻¹ day ⁻¹ | No developmental toxicity | >99% pure; 50 mated CD1 mice; 14% maternal mortality, maternal weight gain decreased, no external malformations on pups |
| | PO | 90 days | 300, 900, 2700, and 8100 mg kg ⁻¹ bw (0, 0.2, 0.6, 1.8, and 5.4% in diet, respectively) | NOAEL = 850–1000 mg kg ⁻¹ bw (0.6% in diet) | At 8100 mg kg ⁻¹ bw intracellular edema of the kidney, increased organ weights, decreased RBC (males), liver cell enlargement, protein inclusions in bladder lumen (males), tubular degeneration, and atrophy. At 2700 mg kg ⁻¹ bw increased relative kidney weights in males were seen |
| | PO | Chronic (12 months) | 850–1000 mg kg ⁻¹ | NOEL | |
| | SC | Acute | 5500 mg kg ⁻¹ | LD ₅₀ | |
| | Dermal | Skin irritation | 0.5 mL over 2 cm ² area | Nonirritant | 50% |
| | Dermal | 28 days | 0, 300, 1000, and 3000 mg kg ⁻¹ day ⁻¹ | NOEL > 1000 mg kg ⁻¹ day ⁻¹ | Undiluted |
| Rabbit | IM | 14 days | 0, 0.62, 0.82, and 1.6 mL kg ⁻¹ day ⁻¹ | NOAEL = 1.6 mL kg ⁻¹ day ⁻¹ | No treatment-related effects |
| | Ocular | Eye irritation | 0.1 mL | Slight irritation | 30% |
| | Ocular | Eye irritation | 0.1 mL | Slight irritation | Undiluted |
| | PO | Acute | 3620 mg kg ⁻¹ | LD ₅₀ | |
| | SC | Acute | 2000 mg kg ⁻¹ | LD ₅₀ | |
| | | | | | |

| | | | | | |
|-----|-------------------------|---|---|--|---|
| Rat | Inhalation (nasal) | 28 days (6h day ⁻¹ , 5 days per week) | 0, 16, 49, and 200 ppm (0, 90, 270, and 1100 mg m ⁻³ , respectively) | NOAEL = 1100 mg m ⁻³ | No systemic effects; mild local irritation; focal necrosis in the larynx (males); 1100 mg m ⁻³ was higher than max concentration at which only vapor present |
| | Inhalation (whole body) | Developmental (7h day ⁻¹ from GD 7–15) | 0 and 102 ppm | No maternal or fetal toxicity | 98–99.5% pure; Sprague–Dawley rats |
| | IP | Acute | 6300 mg kg ⁻¹ | LD ₅₀ | |
| | IV | Acute | 4000 mg kg ⁻¹ | LD ₅₀ | |
| | PO | Acute | 7500 mg kg ⁻¹ | LD ₅₀ | |
| | PO | Acute | 5.0 g kg ⁻¹ | LD ₅₀ > 5000 mg kg ⁻¹ | Undiluted |
| | PO | 90 days | 0, 0.25, 1, and 5% | NOEL = 1% | |
| | PO (gavage) | 90 days | 250 and 2500 mg kg ⁻¹ bw (0.5 and 5.0% in diet, respectively) | NOAEL = 250 mg kg ⁻¹ bw | High-dose group saw reduction in growth rate and food consumption, decreased hemoglobin, RBC decreased (females), oxalate crystals in urine (females), increased organ weights, calcification of the renal cortex; CFE rats |
| | PO (gavage) | Fertility (segment I) | 0, 300, 1000, and 2000 mg kg ⁻¹ day ⁻¹ | NOAEL (oral) = 2000 mg kg ⁻¹ day ⁻¹ | In sterile water |
| | PO (gavage) | 6 weeks | 1340, 2680, and 5360 mg kg ⁻¹ day ⁻¹ | NOAEL = 1340 mg kg ⁻¹ day ⁻¹ | Death, hematological/clinical signs in intermediate- and high-dose groups; lethargy during first week; 10♂; Sprague–Dawley rats |
| | PO (gavage) | Embryo–fetal development study (segment II) | 0, 300, 1000, and 2000 mg kg ⁻¹ day ⁻¹ | NOAEL(dev, mat) 1000 mg kg ⁻¹ day ⁻¹ | In sterile water |
| | PO | Fertility and embryotoxicity range-finding study | 500, 1000, 2000, and 4000 mg kg ⁻¹ day ⁻¹ | NOEL > 500 mg kg ⁻¹ day ⁻¹ | |
| | SC | Acute | 6000 mg kg ⁻¹ | LD ₅₀ | |

TABLE E.105 Trisodium Citrate Dihydrate

| Species | Route | Duration | Dose (mL kg ⁻¹ day ⁻¹) | Adverse Reactions/Toxicity | Notes |
|---------|-------------|-----------|---|----------------------------|------------------------|
| Dog | PO (gavage) | 52 weeks | 10 | Well tolerated | 2.65% aqueous (pH 6.4) |
| Hamster | PO (gavage) | 13 weeks | 10 | Well tolerated | 2.65% aqueous (pH 6.4) |
| Mouse | PO (gavage) | 13 weeks | 10 | Well tolerated | 2.65% aqueous (pH 6.4) |
| Rat | PO (gavage) | 4 weeks | 10 | Well tolerated | 2.65% aqueous (pH 6.4) |
| | PO (gavage) | Segm. III | 10 | Well tolerated | 2.65% aqueous (pH 6.4) |
| | PO (gavage) | 39 weeks | 10 | Well tolerated | 2.65% aqueous (pH 6.4) |

TABLE E.106 Tween 20™

| Species | Route | Duration | Dose | Adverse Reactions/Toxicity | Notes |
|------------------|-----------------|-------------|---------------------------------------|--|---|
| Dog | IV, SC, IP | Single dose | | Not tolerated in any significant amount | |
| | PO | Single dose | | Poorly tolerated in significant amounts | |
| Mouse | PO | Acute | 10 g kg ⁻¹ | Well tolerated | |
| | PO | | 10 g kg ⁻¹ | Well tolerated | |
| Nonhuman primate | IV (slow bolus) | 28 days | 10 mL kg ⁻¹ 3× per week | Red/back discoloration of the skin (anogenital region, hind limbs and forelimbs) | 1.01% in sterile water for injection, USP; age 2–3 years; ♂/♀ |
| Rat | PO | 1 month | 250 mg kg ⁻¹ | Well tolerated | |
| | PO | 90 days | 500 g kg ⁻¹ | Diarrhea | |
| | IV (slow bolus) | 6 months | 10 mL kg ⁻¹ 3× per week | None | 1.01% in sterile water for injection, USP; age 7–8 weeks; ♂/♀ |

TABLE E.107 Tween 80™

| Species | Route | Duration | Dose | Adverse Reactions/Toxicity | Notes |
|---------|-------------|-------------|--|---|--|
| Dog | IV, SC, IP | Single dose | Varies | Not tolerated in any significant amount; hypersensitivity | |
| | PO | Single dose | Varies | Poorly tolerated in significant amounts | |
| | PO (gavage) | ADME | 5 mL kg ⁻¹ day ⁻¹ | Well tolerated | 0.1% |
| | PO (gavage) | 90 days | Dose vol. 5 mL kg ⁻¹ , daily dose 10 mg kg ⁻¹ (0.2% w/v) | Well tolerated | Dose concentration 2 mg mL ⁻¹ |
| Mouse | Intranasal | 3 days | 10 µL per nostril | Well tolerated | 0.2% |
| | IP | 1 month | 10 mL kg ⁻¹ | Well tolerated | 2% solution |
| | PO (gavage) | 35 days | 10 mL kg ⁻¹ BID | Distended abdomen, skin cold to the touch, feces few/absent, limb function impaired, head tilt, swollen abdomen, death; vehicle was not tolerated following 35 days of daily administration | 10% in distilled water; age 6 weeks; ♂/♀ |
| | PO (gavage) | 90 days | Dose vol. 10 mL kg ⁻¹ , daily dose 10 mg kg ⁻¹ (0.2% w/v) | Well tolerated | Dose concentration 1 mg mL ⁻¹ |

TABLE E.107 (Continued)

| Species | Route | Duration | Dose | Adverse Reactions/Toxicity | Notes |
|------------------|--------------|----------|---|----------------------------|---|
| Nonhuman primate | PO (gavage) | Efficac. | 5 mL kg ⁻¹ day ⁻¹ | Well tolerated | 0.1% |
| | PO (gavage) | 90 days | Dose vol. 5 mL kg ⁻¹ , daily dose 10 mg kg ⁻¹ (0.2% w/v) | Well tolerated | Dose concentration 2 mg mL ⁻¹ |
| Rat | IP | Efficac. | 10 mL kg ⁻¹ day ⁻¹ | Well tolerated | 0.2% |
| | IV | Acute | 100 mg kg ⁻¹ | Well tolerated | |
| | PO | Acute | 350 mg kg ⁻¹ | Well tolerated | |
| | PO | 7 days | 10 mL kg ⁻¹ | Well tolerated | 1% solution |
| | PO (dietary) | | 2% (1 g kg ⁻¹) | NOAEL | |
| | PO (gavage) | 4 weeks | 5 mL kg ⁻¹ day ⁻¹ | Well tolerated | 0.1% |
| | PO (gavage) | 90 days | Dose vol. 5 mL kg ⁻¹ , daily dose 10 mg kg ⁻¹ (0.2% w/v) | Well tolerated | Dose concentration 2 mg mL ⁻¹ |

TABLE E.108 Vitamin E TPGS

| Species | Route | Duration | Dose | Adverse Reactions/ Toxicity | Notes |
|---------|-------------|--|--------------------------|--------------------------------|---|
| Dog | PO (gavage) | 32 days | 5 mL kg ⁻¹ QD | None | 20% in DI water; age 5–5.5 months; ♂/♀ |
| Rat | PO (gavage) | QD for 4 days, then off for 4 days (×5 cycles) for total of 32 days | 10 mL kg ⁻¹ | None | 20% in DI water; age 6 weeks; ♂/♀ |
| | PO (gavage) | Single dose | 10 mL kg ⁻¹ | None | 10% Vitamin E TPGS in DI water; age 6 weeks; ♂ |

TABLE E.109 Water

| Species | Route | Duration | Dose | Adverse Reactions/Toxicity | Notes |
|------------------|-------------|------------------------------|---|----------------------------|-------------------------------|
| Dog | PO (gavage) | 14 doses | 10 mL kg ⁻¹ dose ⁻¹ | None | Distilled |
| | PO (gavage) | 28 doses | 5 mL kg ⁻¹ day ⁻¹ | None | Distilled |
| | PO (gavage) | 30 doses | 10 mL kg ⁻¹ | None | Distilled |
| | PO | 9 months | 0.052 mL kg ⁻¹ TID | None | GLP; age 5–6 months; 6♂/6♀ |
| Minipig | SC, IA | q7d × 3 weeks for 90 days | 12 mL total: 2 mL per site | None | GLP; age 5–6 months; 4♂/4♀ |
| Mouse | PO (gavage) | 2 doses | 20 mL kg ⁻¹ day ⁻¹ | None | Distilled |
| | PO (gavage) | 28 doses | 10 mL kg ⁻¹ | None | Deionized |
| Nonhuman primate | PO (gavage) | 14 doses | 10 mL kg ⁻¹ | None | Sterile, USP |
| | PO (gavage) | 28 doses | 10 mL kg ⁻¹ | None | Deionized |

(Continued)

TABLE E.109 (Continued)

| Species | Route | Duration | Dose | Adverse Reactions/Toxicity | Notes |
|---------|-----------------|-------------|---------------------------|----------------------------|------------------------------|
| Pig | Dermal | 9 doses | 10 mL per animal per day | None | Deionized |
| Rat | IV | 11 doses | 5 mL kg ⁻¹ | None | Sterile, USP |
| | IV (slow bolus) | 4 doses | 4 mL kg ⁻¹ | None | Sterile, USP |
| | PO (gavage) | Single dose | 5 mL kg ⁻¹ | None | Sterile, USP |
| | PO (gavage) | Single dose | 5 mL kg ⁻¹ | None | Deionized |
| | PO (gavage) | Single dose | 10 mL kg ⁻¹ | None | Distilled |
| | PO (gavage) | Single dose | 10 mL kg ⁻¹ | None | Deionized |
| | PO (gavage) | Single dose | 10 mL kg ⁻¹ | None | Sterile, USP |
| | PO | Single dose | 20 mL kg ⁻¹ | None | Sterile, USP |
| | PO (gavage) | 5 doses | 10 mL kg ⁻¹ | None | Deionized |
| | PO (gavage) | 14 doses | 5 mL kg ⁻¹ | None | Sterile, USP |
| | PO (gavage) | 14 doses | 10 mL kg ⁻¹ | None | Distilled |
| | PO (gavage) | 14 doses | 10 mL kg ⁻¹ | None | Sterile, USP |
| | PO | 26 weeks | 5 mL kg ⁻¹ TID | None | GLP; age 6 weeks; 25♂/25♀ |
| | PO (gavage) | 28 doses | 5 mL kg ⁻¹ | None | Deionized |
| | PO (gavage) | 28 doses | 10 mL kg ⁻¹ | None | Deionized |
| | PO (gavage) | 30 doses | 10 mL kg ⁻¹ | None | Distilled |

TABLE E.110 Xylitol

| Species | Route | Duration | Dose | Adverse Reactions/ Toxicity | Notes |
|------------------|------------|----------|---|--|---------------------|
| Nonhuman primate | Intranasal | 1 month | Control and high dose of 1200 µL day ⁻¹ ; intermediate dose of 400 µL day ⁻¹ ; low dose of 200 µL day ⁻¹ | Well tolerated at 1200 µL day ⁻¹ | 3.3% in water (w/v) |

TABLE E.111 Combination Formulations

| # | Formulation | Species | Route | Duration | Dose | Adverse Reactions/ Toxicity | Notes |
|----|--|------------------|--------------------------|---|--|--------------------------------|--|
| 1 | Acacia (10%)/Antifoam 1510-US™ (0.05%)/water (purified) | Rat | PO (gavage) | 14 days | 10 mL kg ⁻¹ QD | Well tolerated | Age 9–11 weeks; ♂/♀ |
| 2 | Acacia gum (10%)/DMSO (1%) | Nonhuman primate | PO (gavage) | 2 months (ADME) | 5 mL kg ⁻¹ | Well tolerated | |
| 3 | Acacia gum (10%)/DMSO (1%) | Rat | PO (gavage) | 1 week (ADME) | 5 mL kg ⁻¹ | Well tolerated | Sprague–Dawley |
| 4 | Acacia gum (10%)/DMSO (1%) | Rat | PO (gavage) | 1 month (ADME) | 5 mL kg ⁻¹ | Well tolerated | Sprague–Dawley |
| 5 | Acacia gum (10%)/DMSO (1%) | Rat | PO (gavage) | 1 month (ADME) | 5 mL kg ⁻¹ | Well tolerated | Sprague–Dawley |
| 6 | Acacia gum (10%)/Tween 80™ (0.5%) | Rat | PO (gavage) | Prelim/segment II | 5 mL kg ⁻¹ day ⁻¹ | Well tolerated | |
| 7 | Acacia gum (10%)/Tween 80™ (0.5%) | Rat | PO (gavage) | Segment II | 5 mL kg ⁻¹ day ⁻¹ | Well tolerated | |
| 8 | Acetate buffer (pH 5)/benzyl alcohol (1% pH 5.0) (acetate 100 mM, citrate 10 mM) | Nonhuman primate | IV (into saphenous vein) | 2 weeks | 0.92 mL kg ⁻¹ | Well tolerated | |
| 9 | Acetate buffer pH 5 with 1.0% benzyl alcohol (acetate 100 mM, citrate 10 mM), pH 5.0 | Rat | IV (into tail vein) | 28 days | 0.92 mL kg ⁻¹ | Well tolerated | Sprague–Dawley |
| 10 | Acetic acid (0.01 M)/DMA (95/5) | Dog | 30 min infusion | 1 day | 3 mL kg ⁻¹ in single infusion | None | 99% solution; beagles; age 5 months; ♂/♀ |
| 11 | Acetone/cyclohexane (50/50) | Rat | Dermal | 6 h daily, 5 days per week over 14 days | 1.5 mL kg ⁻¹ | None | 100% solution; no sham group; age ~60 days; ♀ |
| 12 | Acetonitrile/acetic acid (99/1) | Dog | PO (dietary) | Ad libitum over 1 year | | None | Age 7–8 months; ♂/♀ |
| 13 | Benzyl alcohol (60.2%)/citric acid (0.1%)/BHT (0.1%) | Cat | Topical | 24 h | 0.3 mL SD | None | Non-GLP; age 1 year; 1 ♂/1 ♀ |
| 14 | BHT/benzyl alcohol/isopropanol | Dog | Topical | q7d × 2 doses over 7 weeks | 1.4 mL kg ⁻¹ divided into 3 doses 60 min apart | None | Non-GLP; age 7 weeks; 4 ♂/4 ♀ |
| 15 | Capmul MCM™/Koliphor EL™ (50/50) | Nonhuman primate | PO (gavage) | Once weekly over 4 weeks | 5 mL kg ⁻¹ | None | Age 3–6 years; ♂/♀ |
| 16 | Capmul MCM NF™/propylene glycol/Koliphor EL™ in a ratio of 1 : 1 : 1 (by weight) | Rat | PO | 13 weeks | 3 mL kg ⁻¹ BD | Well tolerated | Wistar Hans (CRL); age ~7–9 weeks at study initiation; ♂/♀ |
| 17 | Captisol™ (5.4%)/dextrose (2.5%)/water (pH 4) | Dog | IV (infusion) | 3 days | 2 mL kg ⁻¹ h ⁻¹ , 24 h day ⁻¹ | None | Age 7–8 months; ♂/♀ |

(Continued)

TABLE E.111 (Continued)

| # | Formulation | Species | Route | Duration | Dose | Adverse Reactions/ Toxicity | Notes |
|----|---|------------------|-------------|--------------------|---|---|----------------------------|
| 18 | Citrate buffer (50 mM containing 0.5% Methocel E50 Premium LV™ and 0.2% Tween 20™, pH 4±0.1) | Rat | PO (gavage) | 14 days | 10 mL kg ⁻¹ QD | None | Age 8–9 weeks; ♂/♀ |
| 19 | Citrate buffer (53.8–54.9 mM containing 0.5% Methocel E50 Premium LV™ and 0.2% Tween 20™, pH 4±0.1) | Nonhuman primate | PO (gavage) | 1 day | 5 mL kg ⁻¹ SD | None | Age 3–6 years; ♂ |
| 20 | CMC (0.5%)/Tween 80™ (0.1% (w/w)) | Rabbit | PO (gavage) | Prelim/segment III | 5 mL kg ⁻¹ day ⁻¹ | Well tolerated | |
| 21 | CMC (1%)/Tween 80™ (0.5%) | Dog | PO | 28 days | 10 mL kg ⁻¹ | Well tolerated | |
| 22 | CMC (1%)/Tween 80™ (0.5%) | Rat | PO | 28 days | 20 mL kg ⁻¹ | Well tolerated | |
| 23 | CMC (high viscosity, 0.25%)/Tween 80™ (0.2%)/water (sterile for injection, USP) | Mouse | PO (gavage) | 102 weeks | 10 mL kg ⁻¹ QD | None | Age 6 weeks; ♂/♀ |
| 24 | CMC (low viscosity, 1%)/Tween 80™ (0.01%)/water (distilled) | Rat | PO (gavage) | 105 weeks | 10 mL kg ⁻¹ QD | None | Age 6 weeks; ♂/♀ |
| 25 | CMC (low viscosity, 1%)/Tween 80™ (0.01%)/water (distilled) | Nonhuman primate | PO (gavage) | ≤365 days | 10 mL kg ⁻¹ QD | None | Age 4.5–8 years; ♂/♀ |
| 26 | CMC (low viscosity, 1%)/Tween 80™ (0.01%)/water (distilled) | Mouse | PO (gavage) | 91 weeks | 10 mL kg ⁻¹ QD | None | Age 6 weeks; ♂/♀ |
| 27 | CMC (medium viscosity, 0.5% w/v), Tween 80™ (0.1%, v/v)/water (sterile for injection, USP) | Mouse | PO (gavage) | 13 weeks | 10 mL kg ⁻¹ QD | None | Age 5 weeks; ♂/♀ |
| 28 | CMC (0.5%)/Tween 80™ (0.05%)/in PBS pH 7.4 | Dog | IA | 2 days | 0.35 mL SD | 1 day postdosing, 1 animal had moderate hemorrhage of right stifle fat pad and mildly increased synovial fluid volume on necropsy, 2 had lameness 2 days' postinjection, resolved within 24 h | GLP; age 8 months; 6♂/6♀ |
| 29 | CMC (0.5%)/Tween 80™ (0.05%)/in PBS pH 7.4 | Dog | IA | q30d over 3 months | 350 µL | None | GLP; age 5–6 months; 5♂/5♀ |
| 30 | CMC (0.5%)/Tween 80™ (0.05%)/in PBS pH 7.4 | Dog | IA | q30d over 9 months | 350 µL | None | GLP; age 5–6 months; 5♂/5♀ |

| | | | | | | | |
|----|--|------------------|-----------------|-------------|---|---|--|
| 31 | Corn oil/benzyl alcohol (99 : 1) | Rat | PO (gavage) | 1 month | 1 mL kg ⁻¹ day ⁻¹ | Well tolerated | Sprague–Dawley; during the first 3 days of the study, the vehicle was di(ethylene glycol) ethyl ether. From the 4th day it was replaced by corn oil/benzyl alcohol |
| 32 | Corn oil/EtOH (20%) | Rat | PO (gavage) | 1 month | 2.5 mL kg ⁻¹ day ⁻¹ | Well tolerated | Sprague–Dawley |
| 33 | Corn oil/EtOH (20%) | Rat | PO (gavage) | 1 month | 5 mL kg ⁻¹ day ⁻¹ | Well tolerated | Sprague–Dawley |
| 34 | Cyclodextrin/Ora-Plus™ suspension | Rat | PO (gavage) | 28 doses | 15 mL kg ⁻¹ | Fecal changes (soft, watery, or mucoid); kidney lesions | |
| 35 | Cyclodextrin/Ora-Plus™ suspension | Rat | PO (gavage) | Single dose | 20 mL kg ⁻¹ | None | |
| 36 | Cyclodextrin/Ora-Plus™ suspension | Rat | PO (gavage) | Single dose | 20 mL kg ⁻¹ | None | |
| 37 | Cyclodextrin/Ora-Plus™ suspension | Rat | PO (gavage) | 5 doses | | None | |
| 38 | Dehydrated EtOH (200 proof) (86.249%)/hexylene glycol (10.000%)/dimethiconol Blend 20 (2.500%)/hydroxypropyl cellulose (1.250%)/anhydrous citric acid (0.001%) | Minipig | Topical | 28 days | 5 mL QD | None | GLP; age 3–5 months; 5♂/5♀ |
| 39 | Dextrose/trehalose (25%) | Rat | Bolus injection | 28 doses | 10 mL kg ⁻¹ | None | 5% solution |
| 40 | Dextrose for injection (USP)/EtOH (5%) | Rat | PO (gavage) | Single dose | | None | 5% solution |
| 41 | Diethylacetamide/NaCl | Mouse | IV | | MTD: 1.4 g kg ⁻¹ ; LD ₅₀ : 2.3–3.2 g kg ⁻¹ ; NOEL: 468 mg kg ⁻¹ | | CD-1 mice; 30% of a 5 mL kg ⁻¹ dose volume at MTD; (%v/v) in NaCl |
| 42 | Disodium hydrogen phosphate dihydrate (8 mM)/sodium dihydrogen phosphate dihydrate (7 mM)/NaCl (50 mM)/sucrose (146 mM)/Poloxamer 188™ (0.12 mM)/water (for injection, pH 6.9 ± 0.4) | Nonhuman primate | SC | 26 weeks | 2 mL kg ⁻¹ once weekly | None | Age 1–1.5 years; ♂/♀ |
| 43 | DMSO (1.25%)/5% mannitol (5%)/Kolliphor™ (1.25%) | Nonhuman primate | PO (gavage) | ADME | 1 mL kg ⁻¹ day ⁻¹ | Well tolerated | |

(Continued)

TABLE E.111 (Continued)

| # | Formulation | Species | Route | Duration | Dose | Adverse Reactions/ Toxicity | Notes |
|----|---|------------------|-----------------|------------------------|--|--|--|
| 44 | DMSO (2%)/PEG 400 (10%) | Nonhuman primate | IV | 1 month (ADME) | 1.5 mL kg ⁻¹ | Well tolerated | |
| 45 | DMSO (2%)/PEG 400 (10%) | Nonhuman primate | IV | 1 week (ADME) | 1.5 mL kg ⁻¹ | Well tolerated | |
| 46 | DMSO (2%)/PEG 400 (10%) | Nonhuman primate | IV | 2 months (ADME) | 1.5 mL kg ⁻¹ | Well tolerated | |
| 47 | DMSO (3.5%)/PEG 400 | Nonhuman primate | PO (gavage) | Prelim | 5 mL kg ⁻¹ day ⁻¹ | Soft liquid feces | |
| 48 | DMSO (5%)/mannitol (4.5%)/Kolliphor™ (5%) | Nonhuman primate | IV | 2 weeks | 2 mL kg ⁻¹ day ⁻¹ | Well tolerated | |
| 49 | DMSO/NaCl | Mouse | IV | | MTD: 2.2 g kg ⁻¹ ; LD ₅₀ : 3.8–7.6 g kg ⁻¹ ; NOEL: 1.6 g kg ⁻¹ | | CD-1 mice; 40% of a 5 mL kg ⁻¹ dose volume at MTD; (%v/v) in NaCl |
| 50 | DMSO/PEG 400 (99%) | Rat | PO (gavage) | 11 doses | | None | 1% solution |
| 51 | DMSO/PEG 400 (99%) | Rat | PO (gavage) | 11 doses | 2 mL kg ⁻¹ | None | 1% solution |
| 52 | DMSO/PEG 400/Tris buffer | Nonhuman primate | IV (bolus) | ADME | 2 mL kg ⁻¹ | Well tolerated | |
| 53 | DMSO/PEG 4000/water (DI) (15/35/50) | Rat | IP | 28 days | 10 mL kg ⁻¹ | Well tolerated | |
| 54 | DMSO/PEG 4000/water (DI) (15/35/50) | Mouse | IP | 28 days | 10 mL kg ⁻¹ | Well tolerated | |
| 55 | DMSO/Solutol HS15™/water (pH 3–11) (20/5/75) | Rat | IV (bolus) | | 2 mL kg ⁻¹ | Well tolerated | |
| 56 | DMSO/tetraglycol (50/50) | Minipig | IV | 1 day | 0.4 mL kg ⁻¹ SD | Immediate cardiac arrest; study was cancelled after first dose | Age 3 months; 1 ♂ |
| 57 | DMSO + 1% Tween 80™ at 0.12% in water for injection | Mouse | SC | RNA extraction, 7 days | 4/10 mL kg ⁻¹ per injection | Well tolerated | |
| 58 | EDTA (0.2 mg mL ⁻¹)/citric acid anhydrous (0.8 mg mL ⁻¹)/NaCl (4.6 mg mL ⁻¹ , low iron)/water (sterile for injection, USP, pH 3.9–4.0) | Rat | IV (slow bolus) | 13 weeks | 4.5 mL kg ⁻¹ QD | Red/blue skin discoloration and edema of the tail (nonadverse) | Age 6 weeks; ♂/♀ |

| | | | | | | | |
|----|--|------------------|--------------------------|----------|--|---|--|
| 59 | EDTA pH 7.3 (phosphate 10 mM; NaCl 150 mM; EDTA 0.5 mM) | Rat | IV (into tail vein) | 1 month | 0.92 mL kg ⁻¹ | Well tolerated | Sprague–Dawley |
| 60 | EDTA pH 7.3 (phosphate 10 mM; NaCl 150 mM; EDTA 0.5 mM) | Nonhuman primate | IV (into saphenous vein) | 2 weeks | 0.92 mL kg ⁻¹ | Well tolerated | |
| 61 | EtOH (190 proof) (63.37%)/glycerol (3%)/Carbopol Ultrez™ (10: 2.5%)/Tween 20™ (2%)/propylene glycol (2%)/panthenol (0.15%)/salicylic acid (0.15%)/EDTA (0.05%)/water (DI) (26.78%) | Minipig | Topical | 92 days | 1 g QD | None | GLP; age 3–4 months; 5♂/5♀ |
| 62 | EtOH/Kolliphor EL™/water for injection (10/5/85) | Nonhuman primate | IV | ADME | 0.4 mL kg ⁻¹ | Well tolerated | |
| 63 | EtOH/Kolliphor EL™/water for injection (10/5/85) | Rat | IV | ADME | 2 mL kg ⁻¹ | Well tolerated | |
| 64 | EtOH/NaCl | Mouse | IV | | MTD: 986 mg kg ⁻¹ ; LD ₅₀ : 1.6–4.3 g kg ⁻¹ ; NOEL: 197 mg kg ⁻¹ | | CD-1 mice; 25% of a 5 mL kg ⁻¹ dose volume at MTD; (%v/v) in NaCl |
| 65 | EtOH/propylene glycol/water (12.5/15.5/75, v/v/v) | Rat | PO (gavage) | 2 weeks | 5 mL kg ⁻¹ day ⁻¹ | Well tolerated | |
| 66 | EtOH/propylene glycol/water (12.5/15.5/75, v/v/v) | Rat | SC | 2 weeks | 2 mL kg ⁻¹ day ⁻¹ | Well tolerated | |
| 67 | EtOH/propylene glycol/water (30/10/60) | Dog | IV | | 5 mL kg ⁻¹ at a rate of 0.3 mL kg ⁻¹ | Hemolysis <i>in vitro</i> dog blood | |
| 68 | EtOH/propylene glycol/water (30/20/50) | Dog | IV | | 5 mL kg ⁻¹ at a rate of 0.3 mL kg ⁻¹ | Hemolysis <i>in vitro</i> dog blood | |
| 69 | EtOH/propylene glycol/water (40/10/50) | Dog | IV | | 5 mL kg ⁻¹ at a rate of 0.3 mL kg ⁻¹ | Hemolysis <i>in vitro</i> dog blood | |
| 70 | EtOH/Solutol HS15™/water | Nonhuman primate | PO (gavage) | 39 weeks | 1 mL kg ⁻¹ day ⁻¹ | Well tolerated | |
| 71 | EtOH/Solutol HS15™/water | Nonhuman primate | PO (gavage) | 4 weeks | 1 mL kg ⁻¹ day ⁻¹ | Well tolerated | |
| 72 | EtOH/water (30/70) | Rat | IV | | 5 mL kg ⁻¹ at a rate of 0.3 mL kg ⁻¹ | Hematuria | |
| 73 | EtOH (50%)/propylene glycol (50%)/with BHA (0.05%)/BHT (0.05%) | Minipig | Topical | 16 weeks | 1 mL TID | All males had increased AST at end of study; up to 500 IU L ⁻¹ (ref: 19–263 IU L ⁻¹) | GLP; age 3–4 months; 6♂/6♀ |

(Continued)

TABLE E.111 (Continued)

| # | Formulation | Species | Route | Duration | Dose | Adverse Reactions/ Toxicity | Notes |
|----|---|------------------|-----------------|-------------|--|--|--------------------------------|
| 74 | Gelatin (0.5% m/v)/mannitol (5% m/v)/water for injection | Nonhuman primate | PO (gavage) | Prelim | 5 mL kg ⁻¹ day ⁻¹ | Well tolerated | |
| 75 | Gelucire™/PEG 400/NNMP/Transcutol HP™ (50/30/10/10) | Minipig | PO | 4 days | 5 mL kg ⁻¹ QD | None | Non-GLP; age 4–6 months; 3♀ |
| 76 | Gelucire 44/14™/PEG 400/NNMP/Transcutol HP™ (50/30/10/10) | Minipig | PO | 3 days | 5 mL kg ⁻¹ QD | Mild transient diarrhea 24–30 h postfirst dose | Non-GLP; age 7–8 months; 3♀ |
| 77 | Gelucire 44/14™/PEG 400/NNMP/Transcutol HP™ (50/30/10/10) | Mouse | PO | 5 days | 5 mL kg ⁻¹ QD | None | Non-GLP; age 9–14 weeks; 6♂/6♀ |
| 78 | Histidine (20 mM, pH 6.5)/sucrose (8.8%) | Rat | IV (slow bolus) | 14 days | 5 mL kg ⁻¹ QD | None | Age 6 weeks; ♂/♀ |
| 79 | Histidine (20 mM, pH 6.5)/sucrose (8.8%) | Dog | SC | 14 days | 1 mL kg ⁻¹ QD | None | Age 5–6 months; ♂/♀ |
| 80 | Histidine buffered solution (10 mM, pH 6.5)/NaCl (130 mM)/water (sterile) | Rat | IV (bolus) | 1 day | 7.14 mL kg ⁻¹ single dose | None | Age 5 months; ♂ |
| 81 | HPMC (0.5%)/Tween 80™ (0.1%)/water (DI) | Nonhuman primate | PO (gavage) | 28 days | 5 mL kg ⁻¹ QD | Soft/watery feces | Age 2–4 years; ♂/♀ |
| 82 | HPMC (0.5%)/Tween 80™ (5%) | Rat | PO (gavage) | 2 weeks | 5/1 mL kg ⁻¹ day ⁻¹ | Well tolerated | |
| 83 | HPMC (1%)/Poloxamer 188™ (1%) | Dog | PO | 2 weeks | 5 mL kg ⁻¹ day ⁻¹ | Well tolerated | Beagle |
| 84 | HPMC (1%)/Poloxamer 188™ (1%) | Dog | PO | 2 weeks | 10 or 20 mL kg ⁻¹ day ⁻¹ | Well tolerated | Beagle |
| 85 | HPMC (1%)/Poloxamer 188™ (1%) | Dog | PO | 2 months | 10 mL kg ⁻¹ day ⁻¹ | Well tolerated | Beagle |
| 86 | HPMC (1%)/Tween 80™ (0.25%)/water (purified) | Rat | PO (gavage) | 11 days | 10 mL kg ⁻¹ QD | None | Age 8–10 weeks; ♀ |
| 87 | HPMC/sodium lauryl sulfate (2%)/water (distilled) | Rat | PO (gavage) | Single dose | 10 mL kg ⁻¹ | None | 0.5% solution |
| 88 | HPMC acetate succinate (90 mg mL ⁻¹)/methylcellulose (0.5%) | Mouse | PO | 14 days | 5 mL kg ⁻¹ EOD | None | GLP; age 61–63 days; 20♂/20♀ |
| 89 | HPMC acetate succinate (90 mg mL ⁻¹)/methylcellulose (0.5%) | Minipig | PO | 14 days | 6 mL kg ⁻¹ EOD | None | GLP; age 3–5 months; 5♂/5♀ |
| 90 | HPMC (1%)/fumaric acid (0.5%)/Tween 80™ (0.1%)/glycerol (0.1%)/in water | Dog | PO | 14 days | 5 mL kg ⁻¹ QD | None | Non-GLP; age 7–9 months; 1♂/1♀ |

| | | | | | | | |
|-----|--|------------------|-------------------|----------------------------|--|--|-------------------------------|
| 91 | HPβCD (20% solution)/DMSO (99/1) | Dog | IV | ADME | 4 mL kg ⁻¹ day ⁻¹ | Well tolerated | |
| 92 | HPβCD (20%)/sodium acetate (25 mM, pH 4) | Rat | PO (gavage) | 28 days | 5 mL kg ⁻¹ QD | None | Age 6 weeks; ♂/♀ |
| 93 | HPβCD (20%)/sodium acetate (25 mM, pH 4) | Nonhuman primate | PO (gavage) | 28 days | 2.5 mL kg ⁻¹ dose ⁻¹ BID | None | Age 2–3.5 years; ♂/♀ |
| 94 | HPβCD (30%)/DMSO (5%)/water (purified) (acidic solution, w/w) | Nonhuman primate | IV | 4 weeks | 1 mL kg ⁻¹ day ⁻¹ | Well tolerated | |
| 95 | HPβCD/citric acid solution (0.05 M) | Dog | PO (gavage) | 14 doses | 3 mL kg ⁻¹ | None | 20% solution |
| 96 | HPβCD/NaCl solution (0.9%) | Dog | IV | Single dose | 1 mL kg ⁻¹ | None | 10% solution |
| 97 | HPβCD/NaCl solution (0.9%) | Rat | IV | Single dose | 1 mL kg ⁻¹ | None | 10% solution |
| 98 | HPβCD (12%)/EtOH (8%)/propylene glycol (2%)/water for injection | Dog | IV (1 h infusion) | q7d × 3 doses over 16 days | 2 mL kg ⁻¹ | May cause temporary mild diarrhea | Non-GLP; age 1–4 years; 2♂/2♀ |
| 99 | HPβCD (12%)/EtOH (8%)/propylene glycol (2%)/water for injection | Dog | IV (1 h infusion) | q7d × 3 doses over 16 days | 1.5 mL kg ⁻¹ | May cause diarrhea | Non-GLP; age 1–4 years; 2♂/2♀ |
| 100 | Hydroxyethylcellulose (1%)/Tween 80 TM (0.25%)/Antifoam TM (0.05%)/water (purified) | Rat | PO (gavage) | 1 day | 10 mL kg ⁻¹ single dose | None | Age 6 weeks; ♂ |
| 101 | Hydroxyethylcellulose (1%, w/v)/Tween 80 TM (0.25%)/Antifoam 1510-US TM (0.05%)/water (purified) | Rat | PO (gavage) | 4 days | 10 mL kg ⁻¹ QD | None | Age 8–9 weeks; ♀ |
| 102 | Hymetellose (1%)/Poloxamer 188 TM (0.1%) | Rabbit | PO (gavage) | 1 month | 3 mL kg ⁻¹ day ⁻¹ | Well tolerated | |
| 103 | Imwitor 742 TM /Tween 80 TM (1 : 1, w/w) | Rat | PO (gavage) | 7 days | 10 mL kg ⁻¹ QD | None | Age 7 weeks; ♂/♀ |
| 104 | Imwitor 742 TM /Tween 80 TM (1 : 1, w/w) | Nonhuman primate | PO (gavage) | 7 days | 5 mL kg ⁻¹ QD | Soft/watery feces | Age 2–3 years; ♂/♀ |
| 105 | Imwitor 742 TM /Tween 80 TM (1 : 1, w/w) | Hamster | PO (gavage) | 7 days | 5 mL kg ⁻¹ QD | None | Age 5 weeks; ♂/♀ |
| 106 | Kolliphor EL TM /PEG 300/(4 : 1, w/w) | Mouse | PO (gavage) | 104 weeks | 3 mL kg ⁻¹ day ⁻¹ | Well tolerated | |
| 107 | Kolliphor RH40 TM (41.5%)/PEG 400 (20%)/TPGS (20%)/propylene glycol (10%)/Tween 80 TM (8.5%) | Rat | PO (gavage) | 92 days | 10 mL kg ⁻¹ QD | None | Age 10–23 weeks; ♂/♀ |
| 108 | Kolliphor EL TM /10% 190 proof EtOH/80% 5% dextrose in water (D5W) | Rabbit | IV | Single dose | 5 mL kg ⁻¹ | None | 10% solution |
| 109 | Kolliphor EL TM /10% 190 proof EtOH/80% 5% dextrose in water (D5W) | Rabbit | IV | Single dose | 5 mL kg ⁻¹ | Local irritation, very slight to well-defined erythema | 10% solution |

(Continued)

TABLE E.111 (Continued)

| # | Formulation | Species | Route | Duration | Dose | Adverse Reactions/ Toxicity | Notes |
|-----|--|------------------|------------------------|---|---|--|--|
| 110 | Kolliphor EL™/10% 190 proof EtOH/80% 5% dextrose in water (D5W) | Rabbit | Perivascular injection | Single dose | 0.5 mL kg ⁻¹ | None | 10% solution |
| 111 | Kolliphor EL™/10% 190 proof EtOH/80% 5% dextrose in water (D5W) | Rabbit | Perivascular injection | Single dose | 0.5 mL kg ⁻¹ | Erythema | 10% solution |
| 112 | Labrafil™/Tween 80™ (0.1%) | Rat | PO (gavage) | 2 doses | 2 mL kg ⁻¹ dose ⁻¹ | None | |
| 113 | Labrasol™/Kolliphor HS15™/Transcutol HP™ (60/30/10) | Minipig | PO (gavage) | Once weekly over 4 weeks | 3.75 mL kg ⁻¹ | None | Age 6–7 months; ♂/♀ |
| 114 | Labrasol™/Labrafil™/Transcutol™ | Rat | PO | 4 weeks | 0, 5, 10, or 20 mL kg ⁻¹ day ⁻¹ | Tolerated at 5 mL kg ⁻¹ day ⁻¹ | Changes in appearance and behavior at 10 mL kg ⁻¹ day ⁻¹ ; lethality and renal and hepatic effects at 20 mL kg ⁻¹ day ⁻¹ ; Wistar rats |
| 115 | Labrasol™/PEG-400 (60/40) | Rat | PO | 7 days | 5 mL kg ⁻¹ BD | Well tolerated | Sprague–Dawley rats (Harlan); age ~8–10 weeks; ♂/♀ |
| 116 | Lactated Ringer's injection (USP) | Nonhuman primate | IV | 28 days | 2 mL kg ⁻¹ QD | None | |
| 117 | L-Ascorbic acid/isotonic NaCl | Rat | PO | 90 days | 500 g kg ⁻¹ | Hematologic changes, weight loss | |
| 118 | Mannitol (47 mg mL ⁻¹)/succinic acid (1.181 mg mL ⁻¹)/water (sterile, USP) | Dog | SC | 28 days | 0.1 mL kg ⁻¹ QD | None | Age 5 months; ♂/♀ |
| 119 | Mannitol (5%)/acetate buffer pH 4 (4:6) | Minipig | IM | Tolerance | 0.8/1.2 mL kg ⁻¹ day ⁻¹ | Well tolerated | |
| 120 | Mannitol (5%)/gelatin (0.5%)/Tween 80™ (0.2%) (aqueous solution, % m/v) | Nonhuman primate | PO (gavage) | 2 weeks | 2 mL kg ⁻¹ day ⁻¹ | Well tolerated | |
| 121 | Mannitol (120 mg)/Na2HPO4 (12 mg)/NaCl (10.5 mg)/water (4 mL) | Dog | IV | 1 week (placebo), then 1 day (with the test item) | 0.15 mL kg ⁻¹ | Well tolerated | Beagle dogs |
| 122 | Mannitol (250 mM)/sodium succinate (25 mM)/water pH 4.9 | Rat | SC | 28 days | 4 mL kg ⁻¹ | Well tolerated | |

| | | | | | | | |
|-----|---|------------|-------------|--------------------|---|----------------|---------------------------|
| 123 | Mannitol (250 mM)/sodium succinate (25 mM)/water pH 4.9 | Mouse | SC | 28 days | 20 mL kg ⁻¹ | Well tolerated | |
| 124 | Methanesulfonic acid (5%)/EtOH (5%)/water | Rat | PO | 28 days | 10 mL kg ⁻¹ | Well tolerated | |
| 125 | Methocel™ (0.5%)/Tween 80™ (0.1%)/water (reverse osmosis) | Rat | PO (gavage) | 182 days | 10 mL kg ⁻¹ QD | None | Age 6 weeks; ♀ |
| 126 | Methocel™ (A4M Premium, 0.5%)/simethicone (0.1%)/TPGS (10%)/citrate buffer (17 mM, pH 4±0.05) | Rat | PO (gavage) | 28 days | 10 mL kg ⁻¹ QD | None | Age 8 weeks; ♂/♀ |
| 127 | Methyl methacrylate/glycol dimethacrylate crosspolymer, propylene glycol dicaprylate/dicaprate, BHT | Rat | Topical | 90 days | 2.4 mL kg ⁻¹ × 44–46 day, then 0.75 mL kg ⁻¹ QD | None | GLP; age 8 weeks; 15♂/15♀ |
| 128 | Methylparaben, propylparaben | Minipig | Topical | 91 days | 0.05 mL cm ⁻² TID | None | Age 2–3 months; 6♂/6♀ |
| 129 | Methylcellulose (0.5% w/v)/Tween 80™ (0.1%) | Guinea pig | PO | 28 days | 10 mL kg ⁻¹ | Well tolerated | |
| 130 | Methylcellulose (aqueous, 0.5% w/w)/Tween 80™ (0.5% w/w) | Rat | PO (gavage) | Prelim/segment III | 5 mL kg ⁻¹ day ⁻¹ | Well tolerated | |
| 131 | Methylcellulose/PEG 200 (5%) | Rat | PO (gavage) | Single dose | 10 mL kg ⁻¹ | None | 0.5% solution |
| 132 | Methylcellulose/Tween 80™ (0.1% v/v) | Rat | PO (gavage) | Single dose | 7.5 mL kg ⁻¹ | None | 0.5% solution |
| 133 | Methylcellulose/Tween 80™ (0.1% v/v) | Rat | PO (gavage) | Single dose | 5 mL kg ⁻¹ | None | 0.5% solution |
| 134 | Methylcellulose/Tween 80™ (0.1% v/v) | Rat | PO (gavage) | 28 doses | 5 mL kg ⁻¹ dose ⁻¹ | None | 0.5% solution |
| 135 | Methylcellulose/Tween 80™ (0.1% v/v) | Dog | PO (gavage) | 28 doses | 1 mL kg ⁻¹ dose ⁻¹ | None | 0.5% solution |
| 136 | Methylcellulose/Tween 80™ (0.1% v/v) | Dog | PO (gavage) | Single dose | 5 mL kg ⁻¹ dose ⁻¹ | None | 0.5% solution |
| 137 | Methylcellulose/Tween 80™ (0.1%) | Rat | PO (gavage) | Single dose | 10 mL kg ⁻¹ | None | 0.5% solution |
| 138 | Methylcellulose 1500 cps (0.5%)/Tween 80™ (0.1%)/acetate buffer (10 mM)/water (distilled, pH 4.5±0.1) | Mouse | PO (gavage) | 28 days | 10 mL kg ⁻¹ QD | None | Age 6 weeks; ♂/♀ |
| 139 | Methylcellulose 1500 cps (0.5%)/Tween 80™ (0.1%)/acetate buffer (10 mM)/water (distilled, pH 4.5±0.1) | Mouse | PO (gavage) | 13 weeks | 10 mL kg ⁻¹ QD | None | Age 6 weeks; ♂/♀ |

(Continued)

TABLE E.111 (Continued)

| # | Formulation | Species | Route | Duration | Dose | Adverse Reactions/ Toxicity | Notes |
|-----|---|---------------------|-------------|-------------|---|---|-------------------------------|
| 140 | Methylcellulose 1500 cps (0.5%)/ Tween 80™ (0.1%)/acetate buffer (10 mM)/water (distilled, pH 4.5 ± 0.1) | Nonhuman primate | PO (gavage) | 91 days | 5 mL kg ⁻¹ QD | None | Age 2–4 years; ♂/♀ |
| 141 | Methylcellulose 1500 cps (0.5%)/ Tween 80™ (0.1%)/acetate buffer (10 mM)/water (distilled, pH 4.5 ± 0.1) | Rat | PO (gavage) | 91 days | 10 mL kg ⁻¹ QD | None | Age 6 weeks; ♂/♀ |
| 142 | Methylcellulose 400 cps (0.5%)/ sodium lauryl sulfate (0.5%) | Rat | PO (gavage) | 26 weeks | 10 mL kg ⁻¹ QD | None | Age 6 weeks; ♂/♀ |
| 143 | Methylcellulose 400 cps (0.5%)/ sodium lauryl sulfate (0.5%)/ simethicone (0.01%)/water (DI) (% w/v) | Rat | PO (gavage) | 13–26 weeks | 5 mL kg ⁻¹ QD | None | Age 8 weeks; ♂/♀ |
| 144 | Methylcellulose 400 cps (0.5%)/ sodium lauryl sulfate (0.5%)/ water (DI) | Dog | PO (gavage) | 26 weeks | 5 mL kg ⁻¹ QD | Soft/watery feces (nonadverse) | Age 4–5 months; ♂/♀ |
| 145 | Methylcellulose 400 cps (0.5%)/ sodium lauryl sulfate (0.5%)/ water (DI) | Dog | PO (gavage) | 52 weeks | 5 mL kg ⁻¹ QD | Soft/watery feces (nonadverse) | Age 4–5 months; ♂/♀ |
| 146 | Methylcellulose 400 cps (0.5%)/ sodium lauryl sulfate (0.5%)/ water (DI) | Rat | PO (gavage) | 28 days | 10 mL kg ⁻¹ QD | None | Age 8 weeks; ♂/♀ |
| 147 | Methylparaben (0.17%)/propylparaben (0.03%)/acetyl cysteine (0.5%)/in citrate buffer (100 mM, pH 6.5)/ sodium hydroxide (10%) added to pH 6.0 | Minipig | SC | 56 days | 0.138 mL kg ⁻¹ QD | Temporary dose site irritation postinjection, resolved within a few minutes | GLP; age 4–7 months; 3♂/3♀ |
| 148 | NaCl (0.9%)/propylene glycol/EtOH (50/40/10) | Dog | IV | | 5 mL kg ⁻¹ at a rate of 0.3 mL kg ⁻¹ | Hemolysis <i>in vitro</i> dog blood | |
| 149 | NaCl (0.9%)/EtOH (60/40) | Rat | IV | | 5 mL kg ⁻¹ at a rate of 0.3 mL kg ⁻¹ | Hematuria | |
| 150 | NaCl (0.9%)/EtOH (70/30) | Rat | IV | | 5 mL kg ⁻¹ at a rate of 0.3 mL kg ⁻¹ | Hematuria | |
| 151 | NaCl (0.9%)/EtOH/PEG 400 (50/30/20) | Dog | IV | | 5 mL kg ⁻¹ at a rate of 0.3 mL kg ⁻¹ | Partial hemolysis <i>in vitro</i> dog blood, RBC discolored | |

| | | | | | |
|-----|---|---------|-----------------|--|---|
| 152 | NaCl (0.9%)/EtOH/PEG 400 (50/40/10) | Dog | IV | 5 mL kg ⁻¹ at a rate of 0.3 mL kg ⁻¹ | Partial hemolysis <i>in vitro</i> dog blood, RBC discolored |
| 153 | NaCl (0.9%)/EtOH/propylene glycol (60/30/10) | Dog | IV | 5 mL kg ⁻¹ at a rate of 0.3 mL kg ⁻¹ | Hemolysis <i>in vitro</i> dog blood |
| 154 | NaCl (0.9%)/EtOH/propylene glycol (50/30/20) | Dog | IV | 5 mL kg ⁻¹ at a rate of 0.3 mL kg ⁻¹ | Hemolysis <i>in vitro</i> dog blood |
| 155 | NaCl (0.9%)/EtOH/propylene glycol (50/40/10) | Dog | IV | 5 mL kg ⁻¹ at a rate of 0.3 mL kg ⁻¹ | Hemolysis <i>in vitro</i> dog blood |
| 156 | NaCl (0.9%)/propylene glycol/EtOH (50/30/20) | Dog | IV | 5 mL kg ⁻¹ at a rate of 0.3 mL kg ⁻¹ | Hemolysis <i>in vitro</i> dog blood |
| 157 | NaCl/DMSO/tetraglycol (90/5/5) | Rat | IV | 4.3 mL kg ⁻¹ SD | None GLP; age 8 weeks; 5 ♂/5 ♀ |
| 158 | NaCl/DMSO/tetraglycol (90/5/5) | Minipig | IV | 2.6 mL kg ⁻¹ SD | None GLP; age 2–4 months; 1 ♂/1 ♀ |
| 159 | NaCl for injection (USP)/mannitol (20 mg mL ⁻¹)/Tween 80 TM (4 mg mL ⁻¹), sterile filtered | Rat | IV (infusion) | Single dose | None 0.9% solution |
| 160 | NaCl for injection, USP/10% EtOH and 0.9% | Dog | IV (bolus) | Single dose | None 10% solution |
| 161 | NaCl for injection/10% EtOH | Rat | IV | 5 mL kg ⁻¹ | None 10% solution |
| 162 | NaCl for injection/20% PET in 0.9% | Rat | IV (slow bolus) | 7 doses | None 20% solution |
| 163 | NaCl USP (15.00%)/potassium chloride (7.50%)/L-arginine HCl USP (7.50%)/glyceryl stearate SE (7.00%)/cetyl alcohol NF (7.00%)/propylene glycol (5.00%)/squalene NF (4.00%)/Tween 20 TM NF (2.00%)/sodium hydroxide (1.30%)/oleic acid NF (1.00%)/isopropyl myristate (1.00%)/Keltrol RD (0.50%) (xanthan gum)/Keltrol BT TM (0.30%) (xanthan gum)/water (purified) (40.90%) | Minipig | Topical | 7 days | Severe dose site erythema in 2/6 animals GLP; age 3–5 months; 3 ♂/3 ♀ |
| 164 | Neobee 1053 Oil TM /EtOH/BHT (94.95/5/0.05) | Mouse | IV | 10 mL kg ⁻¹ | Well tolerated |
| 165 | Neobee 1053 Oil TM /EtOH/BHT (94.95/5/0.05) | Rat | PO | 5 mL kg ⁻¹ | Well tolerated |

(Continued)

TABLE E.111 (Continued)

| # | Formulation | Species | Route | Duration | Dose | Adverse Reactions/ Toxicity | Notes |
|-----|--|---------------------|------------------|--------------------------------|--|---|---|
| 166 | NMP/NaCl | Mouse | IV | | MTD: 1.3 g kg ⁻¹ ; LD ₅₀ : 54–3600 mg kg ⁻¹ ; NOEL: 257 mg kg ⁻¹ | | CD-1 mice; 25% of a 5 mL kg ⁻¹ dose volume at MTD; (%v/v) in NaCl |
| 167 | Octoxynol-40, vitamin E | Dog | Ocular (topical) | q1h × 8 doses over 14 days | 0.35 µL (one drop)/eye | None | GLP; age 8 months; 4♂/6♀ |
| 168 | Octoxynol-40, vitamin E | Rabbit | Ocular (topical) | q1h × 8 doses over 14 days | 0.35 µL (one drop)/eye | None | GLP; 5–6 months; 5♂/5♀ |
| 169 | Octoxynol-40, vitamin E | Rabbit | Ocular (topical) | q1h × 8 doses over 13 weeks | 0.35 µL (one drop)/eye | None | GLP; age 6 months; 6♂/6♀ |
| 170 | Oleic acid/PEG 400/Kolliphor EL™ (80/10/10, w/w) | Dog | PO (capsule) | 9 months | 0.6 mL kg ⁻¹ QD | None | Age 6 months; ♂/♀ |
| 171 | Oleic acid/PEG 400/Kolliphor EL™ (80/10/10, w/w) | Rat | PO (gavage) | 104 weeks | 2 mL kg ⁻¹ QD | Decreases in body weight gain (nonadverse) | Age 4 weeks; ♂/♀ |
| 172 | Oleic acid/PEG 400/Kolliphor EL™ (80/10/10, w/w) | Mouse | PO (gavage) | 13 weeks | 2 mL kg ⁻¹ QD | None | Age 6 weeks; ♂/♀ |
| 173 | Oleic acid/PEG 400/Kolliphor EL™ (80/10/10, w/w) | Nonhuman primate | PO (gavage) | 14 days | 2 mL kg ⁻¹ dose ⁻¹ BID | Emesis and fecal changes | Age 2–3 years; ♂/♀ |
| 174 | Olive oil (28.25%)/Tween 80™ (11.25–13.5%)/oleyl alcohol NF (10.00%)/lanolin alcohol NF (8.00%)/ cyclomethicone NF (3.00%)/cetyl acetate (1.5–3.75%)/Shea butter (2.00%)/0.50% sorbitan tristearate/ acetylated lanolin alcohol (0.15–0.75%)/methylparaben NF (0.20%)/propylparaben NF (0.05%)/ water (USP purified) (33.00%) | Minipig | Topical | 14 days | 0.32 mL kg ⁻¹ QD | 4/6 experienced persistent mild (Draize score 1/4) dose site erythema or milary (disseminated in surrounding cutis) erythema | GLP; age 3–5 months; 5♂/5♀ |
| 175 | Peanut oil/EtOH 100 (8:1) | Dog | SC | | 0.33 mL kg ⁻¹ | Well tolerated | Beagle |
| 176 | Peceol™ (Gattefossé)/Tween 80™/ PEG 400/vitamin E (TPPE) (50/40/10/0.2) | Rat | PO (gavage) | 91 days | 10 mL kg ⁻¹ dose ⁻¹ BID | None | Age 6 weeks; ♂/♀ |
| 177 | Peceol™ (Gattefossé)/Tween 80™/ PEG 400/vitamin E (TPPE) (50/40/10/0.2) | Dog | PO (gavage) | 91 days | 5 mL kg ⁻¹ dose ⁻¹ BID | None | Age 7–8 months; ♂/♀ |
| 178 | PEG (5%)/methylcellulose (0.5%) | Pig | PO (gavage) | 28 doses | 5 mL kg ⁻¹ | None | 5% solution |

| | | | | | | |
|-----|--|---------|---|---------|---|--|
| 179 | PEG/DAM (70/30, v/v) | Dog | IV | 2 weeks | 0.32 mL kg ⁻¹ for single IV injection | Well tolerated |
| 180 | PEG/DAM (70/30, v/v) | Rat | IV | 3 weeks | Bolus 0.8–1.07 mL kg ⁻¹ infusion, 0.266–0.356 mL kg ⁻¹ intravenous injection (into tail vein) following by an intravenous injection for 6 h | Well tolerated |
| 181 | PEG/DAM (70/30, v/v) | Dog | IV (into cephalic of saphenous vein) following by an IV injection for 6 h | 2 weeks | Bolus 0.24–0.33 mL kg ⁻¹ infusion, 0.08–0.11 mL kg ⁻¹ h ⁻¹ | Well tolerated |
| 182 | PEG/DAM (70/30, v/v) | Rat | IV | 3 weeks | Bolus 0.8–1.07 mL kg ⁻¹ infusion, 0.266–0.356 mL kg ⁻¹ intravenous injection (into tail vein) following by an intravenous injection for 6 h | Well tolerated |
| 183 | PEG 200/95% methylcellulose (0.5%) | Rat | PO | 3 doses | 5 mL kg ⁻¹ of body weight | None 5% solution |
| 184 | PEG 200/EtOH/dextrose (5%) (70/15/15, v/v/v) | Minipig | IV | 2 weeks | 1 mL kg ⁻¹ | Well tolerated |
| 185 | PEG 300 (40%)/Cavisol W7™ (25/75 v/v) | Dog | PO | 28 days | 10 mL kg ⁻¹ | Well tolerated |
| 186 | PEG 300 (40%)/Cavisol W7™ (25/75 v/v) | Rat | PO | 28 days | 10 mL kg ⁻¹ | Well tolerated |
| 187 | PEG 300/DMA (90/10) | Rat | PO (gavage) | 14 days | 2.5 mL kg ⁻¹ QD | None Age 6 weeks; ♂/♀ |
| 188 | PEG 300/DMA (90/10) | Dog | PO (gavage) | 14 days | 2.5 mL kg ⁻¹ QD | Body weight loss (>20%) (adverse) Age 5–6 months; ♂/♀ |
| 189 | PEG 300/NaCl (0.9%) (40/60, v/v) | Rat | PO (gavage) | ADME | 5 mL kg ⁻¹ day ⁻¹ | Well tolerated |
| 190 | PEG 300/propylene glycol/water (DI) (55/25/20) | Rat | PO (gavage) | 7 days | 10 mL kg ⁻¹ QD | None Age 9–10 weeks; ♂/♀ |

(Continued)

TABLE E.111 (Continued)

| # | Formulation | Species | Route | Duration | Dose | Adverse Reactions/ Toxicity | Notes |
|-----|---|---------|---------------------|----------|---|--|---|
| 191 | PEG 400/Captisol™/EtOH/water (pH 3) (45/7/5/43) | Dog | IV (infusion) | | 1 mL kg ⁻¹ | Well tolerated | |
| 192 | PEG 400/DMA (50/50) | Dog | IV (infusion) | | 0.1 mL kg ⁻¹ | Well tolerated | |
| 193 | PEG 400/DMSO (20%) | Mouse | IV (into tail vein) | Acute | 1128 mg kg ⁻¹ (dose vol. 5 mL kg ⁻¹) | NOEL | 20% PEG 400 |
| 194 | PEG 400/DMSO (20%) | Mouse | IV (into tail vein) | Acute | 3948 mg kg ⁻¹ (dose vol. 5 mL kg ⁻¹) | MTD; ventral recumbency, ataxia, tremors, and hypoactivity shortly after dosing. Tremors for up to 4 min; hypoactivity for up to 10 min is typical | 70% PEG 400 |
| 195 | PEG 400/DMSO (95/5) | Rabbit | PO (gavage) | 12 doses | 0.33 mL kg ⁻¹ | None | 5% solution |
| 196 | PEG 400/EtOH (10%)/DMSO (10%) | Mouse | IV (into tail vein) | Acute | 1692 mg kg ⁻¹ (dose vol. 5 mL kg ⁻¹) | NOEL | 30% PEG 400 |
| 197 | PEG 400/EtOH (10%)/DMSO (10%) | Mouse | IV (into tail vein) | Acute | 2820 mg kg ⁻¹ (dose vol. 5 mL kg ⁻¹) | MTD; ventral recumbency, tremors, ataxia, and hypoactivity shortly after dosing; recovery by 10 min is typical | 50% PEG 400 |
| 198 | PEG 400/EtOH (20%) | Mouse | IV (into tail vein) | Acute | 3384 mg kg ⁻¹ (dose vol. 5 mL kg ⁻¹) | MTD; vocalization and struggling at dosing; ventral recumbency, rapid breathing, tremors, and ataxia shortly after dosing; recovery by 5 min is typical | 60% PEG 400 |
| 199 | PEG 400/EtOH (200 proof) (95/5, v/v) | Rat | PO | 14 days | 10 mL kg ⁻¹ BD | Abnormal clinical observations included anogenital or urogenital staining, soft feces/watery diarrhea, stained body surface, apparent dehydration, staining around mouth or nose/nares, and wet body surface | Wistar Han (CRL) rats; age ~8–10 weeks at study initiation; ♂/♀ |

| | | | | | | | |
|-----|---|------------|---------------|---------|---|--|-------------------------------|
| 200 | PEG 400/EtOH/propylene glycol/ water (sterile) (30/20/20/30) | Rat | IP | 7 days | 2 mL kg ⁻¹ QD | None | Non-GLP; age 7–8 weeks; 6♂ |
| 201 | PEG 400/EtOH/water (pH 3–11) (45/5/50) | Rat | IV (bolus) | | 2 mL kg ⁻¹ | Well tolerated | |
| 202 | PEG 400/EtOH/water (pH 3–11) (45/5/50) | Rat | IV (infusion) | | 5 mL kg ⁻¹ | Well tolerated | |
| 203 | PEG 400/EtOH/water (pH 3–11) (45/5/50) | Dog | IV (bolus) | | 1 mL kg ⁻¹ | Well tolerated | |
| 204 | PEG 400/EtOH/water (pH 3–11) (45/5/50) | Dog | IV (infusion) | | 2 mL kg ⁻¹ | Well tolerated | |
| 205 | PEG 400/EtOH/water (sterile) (1 : 1 : 1) | Guinea pig | IV | 28 days | 2 mL kg ⁻¹ | Well tolerated | |
| 206 | PEG 400/Kolliphor RH40™ (70/30) | Rat | PO (gavage) | 91 days | 10 mL kg ⁻¹ dose ⁻¹ | Sporadic incidences of fecal changes (soft and/ or loose/watery); slight brown/orange staining around anus; fluid contents in the cecum; 10% decrease in mean body weights (male rats); lowered food consumption; mildly increase serum urea; minimally decreased serum sodium and chloride values (male rats); minimally increased total serum cholesterol values (females); alterations in urine electrolytes; organ weight changes; minimal, focal, or multifocal coagulative hepatocellular necrosis (in 3 females and 1 male) | 6♂/6♀ |
| 207 | PEG 400/Kolliphor RH40™ (90/10) | Dog | PO (gavage) | 28 days | 2 mL kg ⁻¹ dose ⁻¹ | Emesis; administration associated with minimal lamina propria hemorrhage in gastric glandular mucosa in 1/3 dogs | Beagle dogs; ♂ |

(Continued)

TABLE E.111 (Continued)

| # | Formulation | Species | Route | Duration | Dose | Adverse Reactions/ Toxicity | Notes |
|-----|---|------------|---------------|----------|--|---|--|
| 208 | PEG 400/Kolliphor RH40™ (90/10) | Dog | PO (gavage) | 28 days | 5 mL kg ⁻¹ dose ⁻¹ | Fecal alterations (loose/watery, mucoid or red) present beginning on day 1; emesis; administration associated with minimal lamina propria hemorrhage in gastric glandular mucosa in 2/3 dogs; single/multifocal red areas in stomach; minimal increase in group mean serum urea | Beagle dogs; ♂ |
| 209 | PEG 400/Labrasol™/Kolliphor EL™ (50/30/20) | Mouse | PO (gavage) | 182 days | 10 mL kg ⁻¹ QD | Unkempt appearance (potential effect) | Age 6 weeks; ♂/♀ |
| 210 | PEG 400/NaCl | Mouse | IV | | MTD: 4.5 g kg ⁻¹ ; LD ₅₀ : 8.6–9.7; NOEL: 1.7 g kg ⁻¹ | Over tested range, expect hypoactivity, tremors, mild ataxia with increasing duration with dose | CD-1 mice; 80% of a 5 mL kg ⁻¹ dose volume at MTD; (%v/v) in NaCl |
| 211 | PEG 400/NaCl (0.9%)/EtOH (50/40/10) | Dog | IV | | 5 mL kg ⁻¹ at a rate of 0.3 mL kg ⁻¹ | Partial hemolysis <i>in vitro</i> dog blood, RBC discolored | |
| 212 | PEG 400/propylene glycol/EtOH/water (pH 3) (40/20/5/35) | Dog | IV (infusion) | | 2.5 mL kg ⁻¹ | Well tolerated | |
| 213 | PEG 400/propylene glycol/Tween 80™/water (25/15/6/54) | Rat | IV | 28 days | 10 mL kg ⁻¹ | Well tolerated | |
| 214 | PEG 400/propylene glycol/Tween 80™/water (25/15/6/54) | Guinea pig | IV | 28 days | 10 mL kg ⁻¹ | Well tolerated | |
| 215 | PEG 400/propylene glycol/Tween 80™/water (25/15/6/54) | Mouse | IV | 28 days | 10 mL kg ⁻¹ | Well tolerated | |
| 216 | PEG 400/PVP K30/TPGS (90/5/5) | Dog | PO (gavage) | 28 doses | 2.5 mL kg ⁻¹ day ⁻¹ | None | 90% solution |
| 217 | PEG 400/PVP K30/TPGS (90/5/5) | Rat | PO (gavage) | 28 doses | 5 mL kg ⁻¹ day ⁻¹ | None | 90% solution |

| | | | | | | | |
|-----|-------------------------------|-----|-------------|---------|--|---|-----------------|
| 218 | PEG 400/Solutol HS15™ (70/30) | Rat | PO (gavage) | 91 days | 10 mL kg ⁻¹ | Males: sporadic fecal changes (soft and/or loose/watery); slight brown/orange staining around anus; 6% decrease in mean body weight; lower food consumption; organ weight changes; increased urine volume; alterations in urine electrolytes (both genders); fluid contents in the cecum (both genders) | 6♂/6♀ |
| 219 | PEG 400/Solutol HS15™ (70/30) | Dog | PO (gavage) | 28 days | 2 mL kg ⁻¹ dose ⁻¹ | Intermittent loose/watery feces; sporadic emesis starting on day 1; 1/3 animals had minimal mucus cell hypertrophy of the ileal mucosa (direct effect of vehicle or effect of loose stools?) | Beagle dogs; 3♀ |
| 220 | PEG 400/Solutol HS15™ (70/30) | Dog | PO (gavage) | 28 days | 5 mL kg ⁻¹ dose ⁻¹ | Consistent incidence of loose/watery feces starting on day 1; sporadic emesis starting on day 1; 2/3 animals had minimal mucus cell hypertrophy of the ileal mucosa (direct effect of vehicle or effect of loose stools?); increase in RBC mass; decrease in urine volume | Beagle dogs; 3♀ |

(Continued)

TABLE E.111 (Continued)

| # | Formulation | Species | Route | Duration | Dose | Adverse Reactions/ Toxicity | Notes |
|-----|---|---------|--------------|-------------|---|---|---|
| 221 | PEG 400/Solutol HS15™ (90/10) | Dog | PO (gavage) | 28 days | 2 mL kg ⁻¹ dose ⁻¹ | Intermittent loose/ watery feces; sporadic emesis starting on day 1 | Beagle dogs; 3♀ |
| 222 | PEG 400/Solutol HS15™ (90/10) | Dog | PO (gavage) | 28 days | 5 mL kg ⁻¹ dose ⁻¹ | Consistent incidence of loose/watery feces starting on day 1; sporadic emesis starting on day 1; all had minimal mucus cell hypertrophy of the ileal mucosa (direct effect of vehicle or effect of loose stools?) | Beagle dogs; 3♀ |
| 223 | PEG 400/TPGS/PVP VA 64/EtOH (80/10/5/5) | Rat | PO (gavage) | 182 days | 2 mL kg ⁻¹ QD | White feces (nonadverse) | Age 8 weeks; ♂/♀ |
| 224 | PEG 400/TPGS/PVP VA 64/EtOH (80/10/5/5) | Dog | PO (gavage) | 9 months | 1 mL kg ⁻¹ QD | None | Age 10–11 months; ♂/♀ |
| 225 | PEG 400/Tween 20™/Poloxamer 124™ (70/20/10) | Rat | PO (gavage) | 26 weeks | 2 mL kg ⁻¹ dose ⁻¹ BID | None | Age 6 weeks; ♂/♀ |
| 226 | PEG 400/Tween 20™/Poloxamer 124™ (70/20/10) | Dog | PO (gavage) | 26 weeks | 0.5 mL kg ⁻¹ dose ⁻¹ BID | None | Age 6.5–7.5 months; ♂/♀ |
| 227 | PEG 400/Tween 20™/TPGS/ Poloxamer 124™ (50/20/20/10) | Mouse | PO (gavage) | 91 days | 2 mL kg ⁻¹ QD | None | Age 6 weeks; ♂/♀ |
| 228 | PEG 400/Tween 20™/TPGS/ Poloxamer 124™ (50/20/20/10) | Dog | PO (capsule) | 91 days | ≤5 mL of vehicle QD | Soft/watery feces (nonadverse) | Age 6–7 months; ♂/♀ |
| 229 | PEG 400/Tween 80™ (95/5) | Mouse | PO | 2-year care | 5 mL kg ⁻¹ BD | During the early stages of the study, a number of mice receiving the vehicle developed gastrointestinal atony resulting in severe bloating and sometimes death | CD-1 (Harlan) mice; age ~7 weeks at study initiation; ♂/♀ |

| | | | | | | | |
|-----|--|------------------|-------------|--------------------------|-------------------------------------|---|--|
| 230 | PEG 400/Tween 80™ (95/5) | Rat | PO | 2-year care | 5 mL kg ⁻¹ BD | Well tolerated | Vacuolation of the tubular epithelium in the kidney associated with PEG-linked proteins; Sprague–Dawley (Harlan) rats; age ~6 weeks at study initiation; ♂/♀ |
| 231 | Phosal 53 MCT™/PEG 400/ Poloxamer 124™/Kolliphor RH40™ (40/20/20/20) | Rat | PO (gavage) | 99 weeks | 2 mL kg ⁻¹ QD | None | Age 6–110 weeks; ♂/♀ |
| 232 | Phosphate (50 mM)/NaCl (100 mM)/ Tween 80™ (0.01%) | Nonhuman primate | SC | Prelim/2 weeks | 1 mL kg ⁻¹ per injection | Well tolerated | |
| 233 | Phosphate buffer 0.5 M at pH 7.5/0.4% Mannitol™ | Rat | SC | Segment I | 0.5 mL per animal per injection | Well tolerated | |
| 234 | Polawax™ (4.80%)/alcohol denatured SDA 40-2 (190 proof) (4.25%)/ propylene glycol (4.00%)/isopropyl myristate (2.50%)/sodium hydroxide solution (10.0) (1.20%)/ phenoxyethanol (1.00%)/Carbomer 974p™ (0.55%)/water (purified) (81.7%) | Minipig | Topical | 210 days | 0.4 g kg ⁻¹ QD | None | GLP; age 3 months; 5♂/5♀ |
| 235 | Polawax™ (4.80%)/propylene glycol (4.00%)/EtOH 200 proof (4.00%)/ isopropyl myristate (2.50%)/sodium hydroxide 10% solution in purified water (1.20%)/phenoxyethanol (1.00%)/Carbomer 974p™ (0.55%) | Mouse | Topical | 90 days | 3.4 mL kg ⁻¹ QD | None | GLP; age 7–8 weeks; 18♂/18♀ |
| 236 | Poloxamer 188™ (0.5%)/NaCl for injection (USP, 0.9%) | Rat | SC | Single dose | 2.5 mL kg ⁻¹ | None | Age 6 weeks; ♂/♀ |
| 237 | Poloxamer 188™ (0.5%)/NaCl for injection (USP, 0.9%) | Minipig | SC | | 1 mL kg ⁻¹ | None | Gottingen minipigs; age 5 months; ♂/♀ |
| 238 | Poloxamer 188™ (1% w/v)/citrate buffer (100 mM, pH 3) | Rat | PO (gavage) | 14 days | 10 mL kg ⁻¹ QD | None | Age 6 weeks; ♂/♀ |
| 239 | Propylene glycol (USP)/glycerol (USP)/EtOH (200 proof, USP) (65/25/10, w/w/w) | Rabbit | SC | 2× per week over 3 weeks | 1 mL kg ⁻¹ | None | Age 6.5 months; ♂/♀ |
| 240 | Propylene glycol (USP)/glycerol (USP)/EtOH (200 proof, USP) (65/25/10, w/w/w) | Nonhuman primate | SC | 1× per week over 3 weeks | 0.5 mL kg ⁻¹ | Scratching and red skin discoloration at the dose site (nonadverse) | Age 3–6 years; ♂/♀ |

(Continued)

TABLE E.111 (Continued)

| # | Formulation | Species | Route | Duration | Dose | Adverse Reactions/ Toxicity | Notes |
|-----|--|---------|---------------|----------|---|--|--|
| 241 | Propylene glycol/Capmul PG8™/ EtOH/water (75/12.5/10/2.5) | Dog | PO (gavage) | 7 days | 5 mL kg ⁻¹ QD | None | Age 5 months; ♂/♀ |
| 242 | Propylene glycol/Capmul PG8™/ EtOH/water (75/12.5/10/2.5) | Rat | PO (gavage) | 21 days | 10 mL kg ⁻¹ QD | Aspiration, salivation, material around the mouth/nose, audible breathing, stereotypical behavior (scratching in the cage following dosing), death | Consistency of vehicle considered to have contributed to aspiration risk and related observations in the rat; age 6–9 weeks; ♂/♀ |
| 243 | Propylene glycol/EtOH/water (20/5/75) | Rat | IV (bolus) | | 1 mL kg ⁻¹ | Well tolerated | |
| 244 | Propylene glycol/EtOH/water (30/20/50) | Dog | IV | | 5 mL kg ⁻¹ at a rate of 0.3 mL kg ⁻¹ | Hemolysis <i>in vitro</i> dog blood | |
| 245 | Propylene glycol/EtOH/water (50/10/40) | Dog | IV | 14 days | 4 mL kg ⁻¹ day ⁻¹ at a rate of 6 mL min ⁻¹ | Frank red urine after first dose; this was occasionally observed throughout the 2-week period; decreases in hematocrit, hemoglobin, and erythrocyte count. Urinalyses were positive for occult blood, bilirubin, ketones, and proteins. Swelling at injection site | |
| 246 | Propylene glycol/EtOH/water (60/20/20) | Dog | IV (bolus) | | 0.5 mL kg ⁻¹ | Well tolerated | |
| 247 | Propylene glycol/EtOH/water (pH 3–11) (40/5/55) | Rat | IV (bolus) | | 2 mL kg ⁻¹ | Well tolerated | |
| 248 | Propylene glycol/EtOH/water (pH 3–11) (40/5/55) | Rat | IV (infusion) | | 5 mL kg ⁻¹ | Well tolerated | |
| 249 | Propylene glycol/EtOH/water (pH 3–11) (40/5/55) | Dog | IV (bolus) | | 1 mL kg ⁻¹ | Well tolerated | |

| | | | | | | |
|-----|--|------------------|---------------|--|--|--|
| 250 | Propylene glycol/EtOH/water (pH 3–11) (40/5/55) | Dog | IV (infusion) | 2 mL kg ⁻¹ | Well tolerated | |
| 251 | Propylene glycol/NaCl | Rat | PO (gavage) | 2 mL kg ⁻¹ day ⁻¹ | Well tolerated | |
| 252 | Propylene glycol/NaCl | Mouse | IV | MTD: 1.5 g kg ⁻¹ ; LD ₅₀ : 5.0–8.6 g kg ⁻¹ ; NOEL: 1 g kg ⁻¹ | | CD-1 mice; 30% of a 5 mL kg ⁻¹ dose volume at MTD; (%v/v) in NaCl |
| 253 | Propylene glycol/NaCl (0.9%)/EtOH (50/40/10) | Dog | IV | 5 mL kg ⁻¹ at a rate of 0.3 mL kg ⁻¹ | Hemolysis <i>in vitro</i> dog blood | |
| 254 | Propylene glycol/PEG 400/water/EtOH (40/25/25/10) | Nonhuman primate | PO (gavage) | 2 mL kg ⁻¹ day ⁻¹ | Well tolerated | |
| 255 | Propylene glycol/TPGS/Capmul MCM NF™ (5:5:2 by weight) | Dog | PO | 5 mL kg ⁻¹ BD | Administration of the vehicle was associated with emesis and abnormal fecal quality throughout the 13 weeks of the study | Beagle dogs (marshall); age ~6 months at study initiation; ♂/♀ |
| 256 | PVP K30 (10%)/sodium citrate buffer (50 mM, pH 5) | Rat | PO (gavage) | 10 mL kg ⁻¹ single dose | None | Age 6 weeks; ♂/♀ |
| 257 | PVP K30 (10%)/sodium citrate buffer (50 mM, pH 5) | Nonhuman primate | PO (gavage) | 5 mL kg ⁻¹ once weekly | None | Age 5–7 years; ♂/♀ |
| 258 | PVP K30 (10%)/sodium citrate buffer (50 mM, pH 5) | Nonhuman primate | PO (gavage) | 5 mL kg ⁻¹ QD | None | Age 2–3 years; ♂/♀ |
| 259 | Sesame oil/EtOH (96/4) | Dog | PO | 2 mL kg ⁻¹ | Well tolerated | |
| 260 | Sesame oil/EtOH (96/4) | Rat | PO | 2 mL kg ⁻¹ | Well tolerated | |
| 261 | Sodium acetate (25 mM, USP)/lactose (70 mg mL ⁻¹)/water (sterile for injection, USP, pH 4.5) | Dog | SC | 1 mL kg ⁻¹ | Red discoloration and swelling at injection sites; chronic active inflammation and hemorrhage | |
| 262 | Sodium acetate (25 mM, USP)/lactose (70 mg mL ⁻¹)/water (sterile for injection, USP, pH 4.5) | Rat | SC | 10 mL kg ⁻¹ | Scabbing at injection site | |
| 263 | Sodium acetate (25 mM, USP)/lactose (70 mg mL ⁻¹)/water (sterile for injection, USP, pH 4.5) | Rat | SC | 4 mL kg ⁻¹ | None | |
| 264 | Sodium acetate (USP, 25 mM)/lactose (USP/EP, 70 mg mL ⁻¹)/water (pH 4.5) | Rat | SC | 10 mL kg ⁻¹ | Well tolerated | |

(Continued)

TABLE E.111 (Continued)

| # | Formulation | Species | Route | Duration | Dose | Adverse Reactions/ Toxicity | Notes |
|-----|--|------------------|----------------------|-----------|--|--------------------------------|---------------------------|
| 265 | Sodium acetate in NaCl, 5 mM | Rat | IV | 1 month | 1 mL kg ⁻¹ | Well tolerated | |
| 266 | Sodium acetate trihydrate buffer (50 mM)/Tween 80 TM (1%) | Mouse | PO (gavage) | 26 weeks | 10 mL kg ⁻¹ day ⁻¹ | Well tolerated | |
| 267 | Sodium acetate trihydrate buffer (50 mM)/Tween 80 TM (1%) | Mouse | PO (gavage) | 7 days | 10 mL kg ⁻¹ day ⁻¹ | Well tolerated | |
| 268 | Sodium citrate/NaCl buffer | Rat | IV | 2 weeks | 4 mL kg ⁻¹ bolus, 2/10 min | Well tolerated | |
| 269 | Sodium CMC (0.1%)/methylparaben sodium (0.1%)/propylparaben sodium (0.02%)/water (purified) (w/v) | Mouse | PO (gavage) | 105 weeks | 10 mL kg ⁻¹ dose ⁻¹ BID | None | Age 7 weeks; ♂/♀ |
| 270 | Sodium CMC (0.1%)/methylparaben sodium (0.1%)/propylparaben sodium (0.02%)/water (purified) (w/v) | Rat | PO (gavage) | 104 weeks | 10 mL kg ⁻¹ dose ⁻¹ BID | None | Age 6 weeks; ♂/♀ |
| 271 | Sodium hydroxide (0.1 M)/NaCl for injection, USP (0.9%) | Nonhuman primate | IV (30 min infusion) | 14 days | 10 mL kg ⁻¹ QD | None | Age 2–4.5 years; ♂/♀ |
| 272 | Sodium hydroxide (10% solution) (2%)/phenoxylethanol (1%)/Carbomer 974P NF TM (1%)/in purified water | Mouse | Topical | 28 days | 3.4 mL kg ⁻¹ QD | None | GLP; age 7 weeks; 12♂/12♀ |
| 273 | Sodium phosphate (10 mM)/NaCl (0.8%)/Tween 20 TM (0.05%)/water (sterile for injection, USP, pH 6.0±0.3) | Rat | SC | 26 weeks | 1500 µL kg ⁻¹ twice weekly | None | Age 6 weeks; ♂/♀ |
| 274 | Sodium phosphate (20 mM)/sucrose (1%)/mannitol (4%)/water (for injection) | Rat | SC | 26 weeks | 1.38 mL kg ⁻¹ QD | None | Age 7 weeks; ♂/♀ |
| 275 | Sodium phosphate (20 mM)/sucrose (1%)/mannitol (4%)/water (for injection) | Rabbit | SC | 39 weeks | 0.58 mL kg ⁻¹ QD | None | Age 5–6 months; ♂/♀ |
| 276 | Sodium phosphate buffer (0.3 M)/PEG 400, pH 8 (70:30, w/w) | Nonhuman primate | PO (gavage) | 4 weeks | 2 mL kg ⁻¹ day ⁻¹ (0.4 mL min ⁻¹ kg ⁻¹) | Well tolerated | |
| 277 | Sodium phosphate buffer (20 mM)/dextrose (4%)/sodium hydroxide (pH 7.9–8.1) | Rat | Infusion | 4 days | 2 mL kg ⁻¹ h ⁻¹ | None | Age 8 weeks; ♀ |

| | | | | | | | |
|-----|--|------------------|-------------|------------------|---|---------------------------------|-------------------------------|
| 278 | Sodium succinate (25 mM)/lactose (45 mg mL ⁻¹)/NaCl (0.45%)/water (sterile for injection, USP) | Dog | IV | 7 doses | 1 mL kg ⁻¹ | None | |
| 279 | Sodium succinate (25 mM)/lactose (45 mg mL ⁻¹)/NaCl (0.45%)/water (sterile for injection, USP) | Dog | SC | 14 doses | 1 mL kg ⁻¹ | None | |
| 280 | Sodium succinate (25 mM)/lactose (45 mg mL ⁻¹)/NaCl (0.45%)/water (sterile for injection, USP) | Rat | IV | 7 doses | 1 mL kg ⁻¹ | None | |
| 281 | Sodium succinate (25 mM)/lactose (45 mg mL ⁻¹)/NaCl (0.45%)/water (sterile for injection, USP) | Rat | SC | 14 doses | 1 mL kg ⁻¹ | None | |
| 282 | Sodium succinate (25 mM)/Mannitol™ (250 mM)/water (sterile for injection, pH 4.6) | Rat | SC | Single dose | 4 mL kg ⁻¹ | None | |
| 283 | Solutol HS 15™/EtOH/water (40/10/50, v/v/v) | Dog | PO | q7d over 8 weeks | 1.5 mL kg ⁻¹ | None | Non-GLP; age 1–4 years; 3♂/3♀ |
| 284 | Solutol HS 15™ (15%)/EtOH (5%)/PBS | Dog | PO | 5 days | 20 mL kg ⁻¹ QD | May cause vomiting, loose stool | Non-GLP; age 1–2 years; 2♂ |
| 285 | Solutol™ (10%)/in NaCl | Rat | IV | 14 days | 6 mL kg ⁻¹ SD | None | Non-GLP; age 10 weeks; 3♂/3♀ |
| 286 | Solutol HS15™/EtOH/water (40/10/50) | Nonhuman primate | PO (gavage) | 9 months | 3 mL kg ⁻¹ day ⁻¹ | Well tolerated | |
| 287 | Sorbitol (5%)/histidine (10 mM)/Tween 80™ (0.01%)/water (sterile for injection, USP, pH 5.8) | Dog | IV (bolus) | 5 weeks | 4.98 mL kg ⁻¹ once weekly | None | Age 2–3 years; ♂/♀ |
| 288 | Soybean oil (50.00%)/coconut oil (23.60%)/mineral oil (5.80%)/cyclomethicone (5.00%)/cetostearyl alcohol (3.50%)/stearic acid (3.00%)/myristyl alcohol (2.50%)/hydrogenated castor oil (2.00%)/white wax (beeswax) (2.00%)/stearyl alcohol (1.50%)/docosanol (1.10%) | Minipig | Topical | 3 weeks | 0.25 mL kg ⁻¹ QD | None | GLP; age 3–4 months; 5♂/5♀ |
| 289 | Sucrose (1%)/NaCl (100 mM)/L-arginine hydrochloride (25 mM)/sodium phosphate (25 mM, pH 6.3)/water (for injection, USP) | Nonhuman primate | SC | 4 weeks | 2 mL kg ⁻¹ twice weekly | Soft/watery feces | Age 2.5–3.5 years; ♂/♀ |
| 290 | Sucrose acetate isobutyrate/EtOH/PEG 300 (90/5/5) | Cat | Oral mucosa | 8 h | 0.1 mL SD | None | Non-GLP; age >6 months; 3♀ |

(Continued)

TABLE E.111 (Continued)

| # | Formulation | Species | Route | Duration | Dose | Adverse Reactions/ Toxicity | Notes |
|-----|--|------------------|-------------|-------------------------------------|---|---|-------------------------------|
| 291 | TPGS (2%)/HPMC acetate succinate (1% HF grade)/PVP K30 (0.25%)/water (DI) | Rat | PO (gavage) | 91 days | 10 mL kg ⁻¹ QD | None | Age 6 weeks; ♂/♀ |
| 292 | TPGS (2%)/HPMC acetate succinate (1% HF grade)/PVP K30 (0.25%)/water (DI) | Dog | PO (gavage) | 91 days | 10 mL kg ⁻¹ QD | None | Age 6–7 months; ♂/♀ |
| 293 | TPGS (2%)/HPMC acetate succinate (1% HF grade)/PVP K30 (0.25%)/water (DI) | Rabbit | PO (gavage) | 14 days | 10 mL kg ⁻¹ QD | None | Age 5–8 months; ♂/♀ |
| 294 | TPGS (2%)/HPMC acetate succinate (1.5% HF grade)/PVP VA 64 (1.5%)/sodium citrate (50 mM pH 5)/water (DI) | Rat | PO (gavage) | 28 days | 10 mL kg ⁻¹ QD | None | Age 6–8 weeks; ♂/♀ |
| 295 | TPGS (5%)/methylcellulose 400 cps (0.5%)/water (DI) | Rat | PO (gavage) | 1 day | 10 mL kg ⁻¹ single dose | None | Age 6 weeks; ♂/♀ |
| 296 | Trehalose (9%)/lactic acid (10 mM) | Dog | IA | q28d × 4 doses over 85 days | 0.4 mL | None | GLP; age 9–11 months; 5 ♂/5 ♀ |
| 297 | Trehalose (9%)/lactic acid (10 mM) | Dog | IV | 2 × per week × 8 doses over 28 days | 0.5 mL kg ⁻¹ | None | GLP; age 1 year; 3 ♀ |
| 298 | Trehalose (9%)/lactic acid (10 mM) | Rat | IV | 14 days | 0.5–5 mL kg ⁻¹ SD | None | GLP; age 16–20 weeks; 6 ♂/6 ♀ |
| 299 | Tween 20™ (0.01%)/sodium acetate (10 mM)/sorbitol (5%, pH 5) | Nonhuman primate | IV (bolus) | 12 weeks | 5 mL kg ⁻¹ once weekly | None | Age 2–4 years; ♂/♀ |
| 300 | Tween 20™ (0.01%)/sodium acetate (10 mM)/sorbitol (5%, pH 5) | Rat | IV (bolus) | 12 weeks | 5 mL kg ⁻¹ once weekly | None | Age 7 weeks; ♂/♀ |
| 301 | Tween 80™/CMC/dimethicone (0.01%) (ratio PS80/CMC of 1 : 1, 0.2% de PS80 and CMC) | Rabbit | PO (gavage) | 1 month | 3 mL kg ⁻¹ day ⁻¹ | Well tolerated | |
| 302 | Tween 80™ (10%)/citric acid (10.5 mg mL ⁻¹)/water (sterile solution) | Rat | IV | 2 doses | 3 mL kg ⁻¹ | None | |
| 303 | Vitamin E (20%)/sodium citrate buffer (50 mM, pH 5) | Rat | PO (gavage) | 28 days | 10 mL kg ⁻¹ QD | None | Age 8 weeks; ♂/♀ |
| 304 | Water (sterile for injection, USP)/sodium hydroxide | Rat | SC | Single dose | 1 mL kg ⁻¹ | Necrosis of the subcutaneous muscle panniculus carnosus, inflamed injection | |
| 305 | Xanthan gum NF (aka Xantural 180™) (0.2%, w/v)/Tween 80™ NF (0.255%, w/v)/water (sterile for injection, USP) | Rat | PO (gavage) | 5 days | 10 mL kg ⁻¹ dose ⁻¹ BID | None | Age 8–10 weeks; ♀ |

REFERENCES

- Ahmad SK, Brinch DS, Friis EP, Pedersen PB. (2004) Toxicological studies on lactose oxidase from *microdochium nivale* expressed in *fusarium venenatum*. *Regul. Toxicol. Pharmacol.* 39(3): 256–270.
- Albers E, Muller BW. (1995) Cyclodextrin derivatives in pharmaceuticals. *Crit. Rev. Ther. Drug Carrier Syst.* 12(4):311–337.
- Ali BH. (2001) Dimethyl sulfoxide: recent pharmacological and toxicological research. *Vet. Hum. Toxicol.* 43(4):228–231.
- Allen B, Gentry R, Shipp A, Van Landingham C. (1998) Calculation of benchmark doses for reproductive and developmental toxicity observed after exposure to isopropanol. *Regul. Toxicol. Pharmacol.* 28(1):38–44.
- American Lecithin. (2007) *Phosal® 53 MCT*. Available at: http://www.americanlecithin.com/TDS/TDS_53MCT.PDF (accessed August 7, 2015).
- Anderson DM. (1986) Evidence for the safety of gum arabic (*Acacia senegal* (L.) Willd.) as a food additive—a brief review. *Food Addit. Contam.* 3(3):225–230.
- Anderson DM. (1989) Evidence for the safety of gum tragacanth (*Asiatic Astragalus* spp.) and modern criteria for the evaluation of food additives. *Food Addit. Contam.* 6(1):1–12.
- Anderson RC, Harris PN, Chen KK. (1950) Toxicological studies on synthetic glycerin. *J. Am. Pharm. Assoc. Sci. Ed.* 39(10):583–585.
- Augustine KA, Zhang Q, Welsh MJ. (2000) A study of vehicles for dosing cultured rodent embryos with non-aqueous soluble compounds. *Toxicologist.* 54(1):297.
- Bachmann E, Weber E, Post M, Zbinden G. (1978) Biochemical effects of gum arabic, gum tragacanth, methylcellulose and carboxymethylcellulose-Na in rat heart and liver. *Pharmacology.* 17(1):39–49.
- Baldrick P, Bamford DG. (1997) A toxicological review of lactose to support clinical administration by inhalation. *Food Chem. Toxicol.* 35(7):719–733.
- Bar A, Til HP, Timonen M. (1995) Subchronic oral toxicity study with regular and enzymatically depolymerized sodium carboxymethylcellulose in rats. *Food Chem. Toxicol.* 33(11):909–917.
- Barrie MD, Dahlstrom DL, Goswami E, Kaetzel R. (2012) Sodium chloride. In: *Patty's Toxicology* (Bingham E and Cohns B, Eds.), 6th Ed, Vol. 1. John Wiley and Sons, Inc., Hoboken, NJ, pp. 1073–1074.
- Bartsch W, Sponer G, Dietmann K, Fuchs G. (1976) Acute toxicity of various solvents in the mouse and rat. LD50 of ethanol, diethylacetamide, dimethylformamide, dimethylsulfoxide, glycerine, n-methylpyrrolidone, polyethylene glycol 400, 1,2-propanediol and Tween 20. *Arzneimittelforschung.* 26(8):1581–1583.
- Beckwith-Hall BM, Holmes E, Lindon JC, Gounarides J, Vickers A, Shapiro M, Nicholson JK. (2002) NMR-based metabonomic studies on the biochemical effects of commonly used drug carrier vehicles in the rat. *Chem. Res. Toxicol.* 15(9):1136–1141.
- Beji S, Kaaroud H, Ben Moussa F, Abderrahim E, Zghidi S, Ben Hamida F, Ben Maiz H, Kheder A. (2006) Acute renal failure following mucosal administration of povidone iodine. *Presse Med.* 35(1 Pt 1):61–63.
- Bellantone R, Bossola M, Carriero C, Malerba M, Nucera P, Ratto C, Crucitti P, Pacelli F, Doglietto GB, Crucitti F. (1999) Structured versus long-chain triglycerides: a safety, tolerance and efficacy randomized study in colorectal surgical patients. *J. Parenter. Enteral Nutr.* 23(3):123–127.
- Bendich A, Cohen M. (1990) Ascorbic acid safety: analysis of factors affecting iron absorption. *Toxicol. Lett.* 51(2):189–201.
- Benita S. (1999) Prevention of topical and ocular oxidative stress by positively charged submicron emulsion. *Biomed. Pharmacother.* 53(4):193–206.
- Bevan C. (2012a) Ethanol. In: *Patty's Toxicology* (Bingham E and Cohns B, Eds.), 6th Ed, Vol. 3. John Wiley & Sons, Inc., Hoboken, NJ, pp. 926–934.
- Bevan C. (2012b) Isopropanol. In: *Patty's Toxicology* (Bingham E and Cohns B, Eds.), 6th Ed, Vol. 3. John Wiley & Sons, Inc., Hoboken, NJ, pp. 937–943.
- Bevan C. (2012c) C16 alcohols. In: *Patty's Toxicology* (Bingham E and Cohns B, Eds.), 6th Ed, Vol. 4. John Wiley & Sons, Inc., Hoboken, NJ, p. 21.
- Briggs D, Lok E, Nera EA, Karpinski K, Clayson DB. (1989) Short-term effects of butylated hydroxytoluene on the Wistar rat liver, urinary bladder and thyroid gland. *Cancer Lett.* 46(1):31–36.
- Buard A, Carlton BD, Floch F, Simon GS. (2003) Subchronic toxicity, mutagenicity and allergenicity studies of a cultured dextrose food product. *Food Chem. Toxicol.* 41(5):689–694.
- Budden R, Kühl UG, Buschmann G. (1978) Studies on the pharmacodynamic activity of several drug solvents. 1st communication: diethyleneglycol monoethylether, n,n-diethylacetamide, dimethylsulfoxide (author's transl). *Arzneimittelforschung.* 28(9):1571–1579.
- Burleigh-Flayer H, Gill M, Hurley J, Bevan C, Gardiner T, Kapp R, Tyler T, Wright G. (1998) Motor activity effects in female Fischer 344 rats exposed to isopropanol for 90 days. *J. Appl. Toxicol.* 18(5):373–381.
- Cambridge MedChem Consulting. (2012) *Formulation*. Cambridge. Available at: <http://www.cambridgemedchemconsulting.com/resources/formulation.html>, updated February 17, 2012 (accessed October 3, 2015).
- Campbell RL, Bruce RD. (1981) Comparative dermatotoxicology. I. Direct comparison of rabbit and human primary skin irritation responses to isopropylmyristate. *Toxicol. Appl. Pharmacol.* 59(3):555–563.
- Caraccio TR, McGuigan MA. (2004) Over-the-counter products. In: *Medical Toxicology* (Dart RC, Michael A and McCuigan MA, Eds.), 3rd Ed. Lippincott Williams & Wilkins, Philadelphia, PA, p. 1051.
- Carlton WW, Boitnott JK, Dungworth DL, Ernst H, Hayashi Y, Mohr U, Parodi AL, Pattengale PK, Rittinghausen S, Ward JM. (2001) Assessment of the morphology and significance of the lymph nodal and hepatic lesions produced in rats by the feeding of certain mineral oils and waxes. Proceedings of a pathology workshop held at the Fraunhofer Institute of Toxicology and Aerosol Research Hannover, Germany, May 7–9, 2001. *Exp. Toxicol. Pathol.* 53(4):247–255.
- Carpenter CP, Shaffer CB. (1952) A study of the polyethylene glycols as vehicle for intramuscular and subcutaneous injection. *J. Am. Pharm. Assoc.* 41:27–29.

- Caujolle F, Canh PH, Dat-Xuong N, Azum-Gelade MC. (1970) Toxicological studies upon acetamide and its n-methyl and n-ethyl derivatives. *Arzneimittelforschung*. 20(9):1242–1246.
- Cavender FL. (2012a) Propylene glycol. In: *Patty's Toxicology* (Bingham E and Cofrissen B, Eds.), 6th Ed, Vol. 4. John Wiley & Sons, Inc., Hoboken, NJ, pp. 611–616.
- Cavender FL. (2012b) Polyethylene glycols. In: *Patty's Toxicology* (Bingham E and Cofrissen B, Eds.), 6th Ed, Vol. 4. John Wiley & Sons, Inc., Hoboken, NJ, pp. 608–611.
- Cavender FL. (2012c) Hydroxypropyl cellulose. In: *Patty's Toxicology* (Bingham E and Cofrissen B, Eds.), 6th Ed, Vol. 4. John Wiley & Sons, Inc., Hoboken, NJ, pp. 945–946.
- Cavender FL. (2012d) Sodium carboxymethyl cellulose. In: *Patty's Toxicology* (Bingham E and Cofrissen B, Eds.), 6th Ed, Vol. 4. John Wiley & Sons, Inc., Hoboken, NJ, p. 943.
- Cavender FL. (2012e) Ethylenediaminetetraacetic acid. In: *Patty's Toxicology* (Bingham E and Cofrissen B, Eds.), 6th Ed, Vol. 2. John Wiley & Sons, Inc., Hoboken, NJ, pp. 483–484.
- Challa R, Ahuja A, Ali J, Khar RK. (2005) Cyclodextrins in drug delivery: an updated review. *AAPS PharmSciTech*. 6(2): E329–E357.
- Cho YA, Gwak HS. (2004) Transdermal delivery of ketorolac tromethamine: effects of vehicles and penetration enhancers. *Drug Dev. Ind. Pharm.* 30:557–564.
- Church AS, Witting MD. (1997) Laboratory testing in ethanol, methanol, ethylene glycol, and isopropanol toxicities. *J. Emerg. Med.* 15(5):687–692.
- Clark DP, Hanke CW, Swanson NA. (1989) Dermal implants: safety of products injected for soft tissue augmentation. *J. Am. Acad. Dermatol.* 5(1):992–998.
- Coon JS, Knudson W, Clodfelter K, Lu B, Weinstein RS. (1991) Solutol®HS15, nontoxic polyethylene esters of 12-hydroxystearic acid, reverses multidrug resistance. *Cancer Res.* 51:897–902.
- Cosmetic Ingredient Review. (1999) Final report on the safety assessment of propylene glycol (PG) dicaprylate, PG dicaprylate/dicaprate, PG dicocoate, PG dipelargonate, PG isostearate, PG laurate, PG myristate, PG oleate, PG oleate SE, PG dioleate, PG dicaprate, PG diisostearate, and PG dilaurate. *Int. J. Toxicol.* 18:35–52.
- Cosmetic Ingredient Review. (2001) Final report on the safety assessment of peanut (arachis hypogaea) oil, hydrogenated peanut oil, peanut acid, peanut glycerides, and peanut (arachis hypogaea) flour. *Int. J. Toxicol.* 20(Suppl. 2):65–77.
- Cosmetic Ingredient Review. (2004) Final report on the amended safety assessment of glyceryl laurate, glyceryl laurate SE, glyceryl laurate/oleate etc. *Int. J. Toxicol.* 23(2):55–94.
- Coussement W, Van Cauteren H, Vandenberghe J, et al. (1990) Toxicological profile of hydroxypropyl β -cyclodextrin (HP β -CD) in laboratory animals (Duchene D). Minutes 5th International Symposium on Cyclodextrins. Editions de Sante, Paris, pp. 552–613.
- Curry DJ, Wright DA, Lee RC, Kang UJ, Frim DM. (2004) Poloxamer 188 volumetrically decreases neuronal loss in the rat in a time-dependent manner. *Neurosurgery*. 55(4):943–949.
- Daher CF, Baroody GM, Howland RJ. (2003) Effect of a surfactant, Tween 80, on the formation and secretion of chylomicrons in the rat. *Food Chem. Toxicol.* 41(4):575–582.
- Dalbey WE, Biles RW. (2003) Respiratory toxicology of mineral oils in laboratory animals. *Appl. Occup. Environ. Hyg.* 18(11):921–929.
- David RM, Bachman AN, Butala JH, Piper JT, Shelp CJ. (2012) Methyl benzoate. In: *Patty's Toxicology* (Bingham E and Cofrissen B, Eds.), 6th Ed, Vol. 4. John Wiley & Sons, Inc., Hoboken, NJ, pp. 152–156.
- DeWitt JC, Meyer EB, Henshel DS. (2005) Environmental toxicity studies using chickens as surrogates for wildlife: effects of vehicle volume. *Arch. Environ. Contam. Toxicol.* 48(2):260–269.
- ChemIDPlus. (2015) *Diethylacetamide*. Rockville, MD. Available at: <http://chem.sis.nlm.nih.gov/chemidplus/name/diethylacetamide>. (accessed August 7, 2015).
- Dordunoos SK, Ford JL, Rubinstein MH. (1996) Solidification studies of polyethylene glycols, gelucire 44/14 or their dispersions with triamterene or temazepam. *J. Pharm. Pharmacol.* 48(8):782–789.
- Dow Corning. (2015a) *Antifoam 1510-US*. MSDS. Available at: <http://www.dowcorning.com/DataFiles/09027701828be040.pdf> (accessed August 7, 2015).
- Dow Corning. (2015b) *Dimethiconol Blend 20*. MSDS. Available at: <http://www.dowcorning.com/DataFiles/090277018258ccb9.pdf> (accessed August 7, 2015).
- Dupont J, White PJ, Carpenter MP, Schaefer EJ, Meydani SN, Elson CE, Woods M, Gorbach SL. (1990) Food uses and health effects of corn oil. *J. Am. Coll. Nutr.* 9(5):438–470.
- Dykes MHM, Meier P. (1975) Ascorbic acid and the common cold: evaluation of its efficacy and toxicity. *J. Am. Med. Assoc.* 231(10):1073–1079.
- Earl LK, Baldrick P, Hepburn PA. (2002) A 13-week feeding study in the rat with shea oleine and hardened shea oleine. *Int. J. Toxicol.* 21:13–22.
- Evangelista CM, Antunes LM, Francescato HD, Bianchi ML. (2004) Effects of the olive, extra virgin olive and canola oils on cisplatin-induced clastogenesis in Wistar rats. *Food Chem. Toxicol.* 42(8):1291–1297.
- Farber TM, Ritter DL, Weinberger MA, Bierbower G, Tanner JT, Friedman MH, Carter CJ, Earl FL, van Loon EJ. (1976) The toxicity of brominated sesame oil and brominated soybean oil in miniature swine. *Toxicology*. 5(3):319–336.
- Feitoza AB, Gostout CJ, Burgart LJ, Burkert A, Herman LJ, Rajan E. (2003) Hydroxypropyl methylcellulose: a better submucosal fluid cushion for endoscopic mucosal resection. *Gastrointest. Endosc.* 57(1):41–47.
- Finì A, Moyano JR, Ginés JM, Perez-Martinez JJ, Rabasco AM. (2005) Diclofenac salts, II. Solid dispersions in PEG6000 and Gelucire 50/13. *Eur. J. Pharm. Biopharm.* 60(1):99–111.
- Fisherman EW, Cohen GN. (1974) Tween 80 exacerbated intrinsic rhinitis and asthma in non-aspirin-sensitive patients. *Ann. Allergy*. 32(6):307–320.
- Fort FL, Heyman IA, Kesterson JW. (1984) Hemolysis study of aqueous polyethylene glycol 400, propylene glycol and ethanol combinations in vivo and in vitro. *J. Parenter. Sci. Technol.* 38(2):82–87.
- Freeman C, Weiner ML, Kotkoskie LA, Borzelleca J, Butt M. (2003) Subchronic and developmental toxicity studies in rats with Ac-Di-Sol croscarmellose sodium. *Int. J. Toxicol.* 22(3):149–157.

- Frim DM, Wright DA, Curry DJ, Cromie W, Lee R, Kang UJ. (2004) The surfactant Poloxamer-188 protects against glutamate toxicity in the rat brain. *Neuroreport*. 15(1):171–174.
- Gad SE. (2014) Cyclohexane. In: *Encyclopedia of Toxicology* (Wexler P, Ed.), 3rd Ed, Vol. 1. Elsevier, Oxford, pp. 1106–1108.
- Gattefossé 2004 *Gattefossé Technical Document*. Gelucire 50/13. October 25, 2004.
- Gattefossé 2003a *Gattefossé Technical Document*. Labrafil M1944 CS. February 24, 2003.
- Gattefossé 2003b *Gattefossé Technical Document*. Plurol Oleique CC497. December 16, 2003.
- Geerling G, Daniels JT, Dart JK, Cree IA, Khaw PT. (2001) Toxicity of natural tear substitutes in a fully defined culture model of human corneal epithelial cells. *Invest. Ophthalmol. Vis. Sci.* 42(5):948–956.
- Gelderblom H, Verweij J, Nooter K, Sparreboom A. (2001) Cremophor EL: the drawbacks and advantages for drug formulation. *Eur. J. Cancer*. 37(13):1590–1598.
- Gelperina SE, Khanansky AS, Skidan IN, Smirnova ZS, Bobruskin AI, Severin SE, Turowski B, Zanella FE, Kreuter J. (2002) Toxicological studies of doxorubicin bound to polysorbate 80-coated poly(butyl cyanoacrylate) nanoparticles in healthy rats and rats with intracranial glioblastoma. *Toxicol. Lett.* 126(2):131–141.
- Genovese RF, Newman DB, Gordon KA, Brewer TG. (1999) Acute high dose artemether toxicity in rats. *Neurotoxicology*. 20(5):851–859.
- Gerloczy A, Hoshino T, Pitha J. (1994) Safety of oral cyclodextrins: effects of hydroxypropyl cyclodextrins, cyclodextrins sulfates and cationic cyclodextrins on steroid balance in rats. *J. Pharm. Sci.* 83(2):193–196.
- Gould S, Scott RC. (2005) 2-Hydroxypropyl-beta-cyclodextrin (HP-beta-CD): a toxicology review. *Food Chem. Toxicol.* 43(10):1451–1459.
- Grindel JM, Jaworski T, Piraner O, Emanuele RM, Balasubramanian M. (2002) Distribution, metabolism, and excretion of a novel surface-active agent, purified poloxamer 188, in rats, dogs, and humans. *J. Pharm. Sci.* 91(9):1936–1947.
- Gupta U, Beaulieu J, Chapin Hopper J, Hagler AR, Hills-Perry P. (1996) Teratogenic evaluation of alternative vehicles: PEG 400, cremophor, carboxy-methylcellulose; comparisons with methylcellulose. *Teratology*. 53(2):111.
- Gutiérrez-Cabano CA. (2000) Protection by ingastric polyethylene glycol 400 in rat stomach against ethanol damage involves α -adrenoceptors. *Dig. Dis. Sci.* 45(1):105–109.
- Hagiwara A, Boonyaphiphat P, Kawabe M, Naito H, Shirai T, Ito N. (1992) Lack of carcinogenicity of tragacanth gum in B6C3F1 mice. *Food Chem. Toxicol.* 30(8):673–679.
- Hardin BD. (1983) Reproductive toxicity of the glycol ethers. *Toxicology*. 27(2):91–102.
- Hardin BD, Goad PT, Burg JR. (1984) Developmental toxicity of four glycol ethers applied cutaneously to rats. *Environ. Health Perspect.* 57:69–74.
- Heimbach J, Rieth S, Mohamedshah F, Slesinski R, Samuel-Fernando P, Sheehan T, Dickmann R, Borzelleca J. (2000) Safety assessment of iron EDTA: summary of toxicological, fortification and exposure data. *Food Chem. Toxicol.* 38(1):99–111.
- Hermansky SJ, Neptun DA, Loughran KA, Leung HW. (1995) Effects of polyethylene glycol 400 (PEG 400) following 13 weeks of gavage treatment in Fischer-344 rats. *Food Chem. Toxicol.* 33(2):139–149.
- Hironishi M, Kordek R, Yanagihara R, Garruto RM. (1996) Maltol (3-hydroxy-2-methyl-4-pyrone) toxicity in neuroblastoma cell lines and primary murine fetal hippocampal neuronal cultures. *Neurodegeneration*. 5(4):325–329.
- Horvath IP, Somafi-Relle S, Hegedus L, Jarman M. (1982) Toxicity, antitumor and haematological effects of 1,2-anhydro-6-bromogalactitol and d-mannitol: a comparison with the related dibromo- and dianhydro-derivatives. *Eur. J. Cancer Clin. Oncol.* 18(6):573–577.
- Hu Z, Tawa R, Konishi T, Shibata N, Takada K. (2001) A novel emulsifier, Labrasol, enhances gastrointestinal absorption of gentamicin. *Life Sci.* 69(24):2899–2910.
- Hu Z, Prasad Yv R, Tawa R, Konishi T, Ishida M, Shibata N, Takada K. (2002) Diethyl ether fraction of Labrasol having a stronger absorption enhancing effect on gentamicin than Labrasol itself. *Int. J. Pharm.* 234(1–2):223–235.
- JECFA. (1997) *Alginic Acid*. 49th JECFA FNP 52 Addendum 5. Food & Agriculture Organization Website. Available at: <http://www.fao.org/ag/agn/jecfa-additives/specs/Monograph1/Additive-009.pdf> (accessed August 6, 2015).
- Jefferson R. (2012) Trisodium phosphate. In: *Patty's Toxicology* (Bingham E and Cohns B, Eds.), 6th Ed, Vol. 1. John Wiley & Sons, Inc., Hoboken, NJ, pp. 942–943.
- Jockovich M, Murray T, Excalona-Benz E, Hernandez E, Feuer W. (2006) Anecortave acetate as a single and adjuvant therapy in the treatment of retinal tumors of LDbetaTag mice. *Invest. Ophthalmol. Vis. Sci.* 47:1264–1268.
- Kärber G. (1951) Über sogenanntes Oxydwachs. *Arch. Exp. Pathol. Pharmacol.* 212:509–521.
- Kawakami K, Miyoshi K, Ida Y. (2004) Solubilization behavior of poorly soluble drugs with combined use of Gelucire 44/14 and cosolvent. *J. Pharm. Sci.* 93(6):1471–1479.
- Kawashima H, Toyoda-Ono Y, Suwa Y, Kiso Y. (2009) Subchronic (13-week) oral toxicity study of dihomo- γ -linolenic acid (DGLA) oil in rats. *Food Chem. Toxicol.* 47(6):1280–1286.
- Kennedy GL Jr. (2012) N-methyl-2-pyrrolidinone. In: *Patty's Toxicology* (Bingham E and Cohns B, Eds.), 6th Ed, Vol. 2. John Wiley & Sons, Inc., Hoboken, NJ, pp. 727–731.
- Kim Y, Oksanen DA, Massefski W Jr, Blake JF, Duffy EM, Chrunk B. (1998) Inclusion complexation of ziprasidone mesylate with beta-cyclodextrin sulfobutyl ether. *J. Pharm. Sci.* 87(12):1560–1567.
- Kligman AM. (1983) Lanolin allergy: crisis or comedy. *Contact Dermatitis*. 9(2):99–107.
- Komatsu H, Asaba K, Suzuki M. (1979) Some biochemical effects of isopropyl myristate and squalane on rabbit skin. *J. Soc. Cosmet. Chem.* 30:263–278.
- Kreckmann KH, Baldwin JK, Roberts LG, Staab RJ, Kelly DP, Saik JE. (2000) Inhalation developmental toxicity and reproduction studies with cyclohexane. *Drug Chem. Toxicol.* 23:555–573.
- Lanigan RS, Yamarik TA. (2002a) Final report on the safety assessment of BHT(1). *Int. J. Toxicol.* 21(Suppl. 2):19–94.

- Lanigan RS, Yamarik TA. (2002b) Final report on the safety assessment of sorbitan caprylate, sorbitan cocoate, sorbitan diisostearate, sorbitan dioleate, sorbitan distearate, sorbitan isostearate, sorbitan olivate, sorbitan sesquiisostearate, sorbitan sesquisteate, and sorbitan triisostearate. *Int. J. Toxicol.* 21(1):93–112.
- Lanigan RS, Yamarik TA. (2002c) Final report on the safety assessment of EDTA, calcium disodium EDTA, diammonium EDTA, dipotassium EDTA, disodium EDTA, TEA-EDTA, tetrasodium EDTA, tripotassium EDTA, trisodium EDTA, HEDTA, and trisodium HEDTA. *Int. J. Toxicol.* 2:95–142.
- Leadscope Data Portal. (2008a) LS-711. Modified May 19, 2008.
- Leadscope Data Portal. (2008b) LS-55950. Modified May 20, 2008.
- Lee KP, Chromey NC, Culik R, Barnes JR, Schneider PW. (1987) Toxicity of n-methyl-2-pyrrolidone (NMP): teratogenic, subchronic, and two-year inhalation studies. *Fundam. Appl. Toxicol.* 9(2):222–235.
- Lemieux P, Guerin N, Paradis G, Proulx R, Chistyakova L, Kabanov A, Alakhov V. (2000) A combination of poloxamers increases gene expression of plasmid DNA in skeletal muscle. *Gene Ther.* 7(11):986–991.
- Li P, Ghosh A, Wagner RF, Krill S, Joshi YM, Serajuddin AT. (2005) Effect of combined use of nonionic surfactant on formation of oil-in-water microemulsions. *Int. J. Pharm.* 288:27–34.
- Li B-Q, Dong X, Fang S-H, Gao JY, Yang GQ, Zhao H. (2011) Systemic toxicity and toxicokinetics of a high dose of polyethylene glycol 400 in dogs following intravenous injection. *Drug Chem. Toxicol.* 34(2):208–212.
- Lina BA, Bos-Kuijpers MH, Til HP, Bar A. (1996) Chronic toxicity and carcinogenicity study of erythritol in rats. *Regul. Toxicol. Pharmacol.* 24(2 Pt 2):S264–S279.
- Liu H, Li S, Wang Y, Yao H, Zhang Y. (2006a) Effect of vehicles and enhancers on the topical delivery of cyclosporin A. *Int. J. Pharm.* 311:182–186.
- Liu Z, Li J, Nie S, Guo H, Pan W. (2006b) Effects of Transcutol P on the corneal permeability of drugs and evaluation of its ocular irritation of rabbit eyes. *J. Pharm. Pharmacol.* 58(1):45–50.
- Lorenz W, Reimann HJ, Schmal A, Dormann P, Schwarz B, Neugebauer E, Doenicke A. (1977) Histamine release in dogs by cremophor EL and its derivatives: oxethylated oleic acid is the most effective constituent. *Agents Actions.* 7(1):63–67.
- Maki KC, Davidson MH, Torri S, Ingram KA, O'Mullane J, Daggy BP, Albrecht HH. (2000) High-molecular-weight hydroxypropylmethylcellulose taken with or between meals is hypocholesterolemic in adult men. *J. Nutr.* 130(7):1705–1710.
- Malley LA, Bamberger JR, Stadler JC, Elliott GS, Hansen JF, Chiu T, Grabowski JS, Pavkov KL. (2000) Subchronic toxicity of cyclohexane in rats and mice by inhalation exposure. *Drug Chem. Toxicol.* 23(4):513–537.
- Martin E, Verhoef JC, Merkus FW. (1998) Efficacy, safety and mechanism of cyclodextrins as absorption enhancers in nasal delivery of peptide and protein drugs. *J. Drug Target.* 6(1):17–36.
- McCarthy DM, Haas M, Thuluvath PJ, Elliott GS, Hansen JF, Chiu T, Grabowski JS, Pavkov KL. (2002) Pulmonary embolization of bovine collagen. *Arch. Pathol. Lab. Med.* 127(2):67–69.
- Mehlman MA. (2012a) Carboxymethylcellulose. In: *Patty's Toxicology* (Bingham E and Cohns B, Eds.), 6th Ed, Vol. 3. John Wiley & Sons, Inc., Hoboken, NJ, pp. 644–645.
- Mehlman MA. (2012b) Hydroxypropyl methylcellulose. In: *Patty's Toxicology* (Bingham E and Cohns B, Eds.), 6th Ed, Vol. 3. John Wiley & Sons, Inc., Hoboken, NJ, pp. 643–644.
- Mehlman MA. (2012c) Methylcellulose. In: *Patty's Toxicology* (Bingham E and Cohns B, Eds.), 6th Ed, Vol. 3. John Wiley & Sons, Inc., Hoboken, NJ, pp. 642–643.
- Meneely GR, Ball CO. (1958) Experimental epidemiology of chronic sodium chloride toxicity and the protective effect of potassium chloride. *Am. J. Med.* 25:713–725.
- Meneely GR, Tucker RG, Darby WJ, Auerbach SH. (1953) Chronic sodium chloride toxicity in the albino rat: II. *J. Exp. Med.* 98:71–80.
- Modderman JP. (1993) Safety assessment of hydrogenated starch hydrolysates. *Regul. Toxicol. Pharmacol.* 18:80–114.
- Moore GL, Boswell GW, Ledford ME, Gildengoren VL, Hankins DJ, Mathewson PJ. (1988) Toxicity and clearance of sodium phosphate intravenously injected into rabbits. *Mil. Med.* 153(4):203–206.
- Moorman M, Safrit C, Richardson J, Gwaltney-Brant S. (Speakers) (2001) *Winter Hazards Part II—Antifreeze Toxicity*. National Animal Poison Control Center, Nanette Walker Smith, N. (Moderator). Available at: http://www.vspn.org/library/rounds/vspn_vspn011210.htm (accessed June 16, 2015).
- Murakami K, Ishida K, Watakabe K, Tsubouchi R, Naruse M, Yoshino M. (2006) Maltol/iron-mediated apoptosis of HL60 cells: participation of reactive oxygen species. *Toxicol. Lett.* 161(2):102–107.
- Nair B. (2001) Final report on the safety assessment of benzyl alcohol, benzoic acid, and sodium benzoate. *Int. J. Toxicol.* 20(Suppl 3):23–50.
- Nakagawa Y, Tayama K, Nakao T, Hiraga K. (1984) On the mechanism of butylated hydroxytoluene-induced hepatic toxicity in rats. *Biochem. Pharmacol.* 33(16):2669–2674.
- Nash JF, Gettings SD, Diembeck W, Chudowski M, Kraus AL. (1996) A toxicological review of topical exposure to white mineral oils. *Food Chem. Toxicol.* 34(2):213–225.
- National Toxicology Program. (1992) NTP toxicology and carcinogenesis studies of polysorbate 80 (CAS no. 9005-65-6) in F344/N rats and B6C3F1 mice (feed studies). *Natl. Toxicol. Program Tech. Rep. Ser.* 415:1–225.
- National Toxicology Program. (2001) Toxicology and carcinogenesis studies of coconut oil acid diethanolamine condensate (CAS no. 68603-42-9 in F344/N rats and B6C3F1 mice (dermal studies). *Natl. Toxicol. Program Tech. Rep. Ser.* 479:5–226.
- O'Sullivan SM, Woods JA, O'Brien NM. (2004) Use of Tween 40 and Tween 80 to deliver a mixture of phytochemicals to human colonic adenocarcinoma cell (CaCo-2) monolayers. *Br. J. Nutr.* 91(5):757–764.
- Obara S, Muto H, Kokubo H, Ichikawa N, Kawanabe M, Tanaka O. (1992) Studies on single-dose toxicity of hydrophobically modified hydroxypropyl methylcellulose in rats. *J. Toxicol. Sci.* 17(1):13–19.

- Oz M, Spivak CE, Lupica CR. (2004) The solubilizing detergents, Tween 80 and Triton X-100 non-competitively inhibit alpha 7-nicotinic acetylcholine receptor function in xenopus oocytes. *J. Neurosci. Methods*. 137(2):167–173.
- Paddock Laboratories. (2015) *Ora-Plus Oral Suspending Vehicle*. Available at: <http://medicalfs.com.au/portal/images/stories/Orarange/Ora-Plus-Sell-Sheet.pdf> (accessed August 7, 2015).
- Passerini N, Perissutti B, Moneghini M, Voinovich D, Albertini B, Cavallari C, Rodriguez L. (2002) Characterization of carbamazepine-Gelucire 50/13 microparticles prepared by a spray-congealing process using ultrasounds. *J. Pharm. Sci.* 91(3):699–707.
- Patel M, Patel G, Patel R, Patel J, Bharadia P, Patel MM. (2005) PEGylation: an innovative approach for protein delivery. *Drug Deliv. Technol.* 5(6):48–56.
- Pestel S, Martin HJ, Maier GM, Guth B. (2006) Effect of commonly used vehicles on gastrointestinal, renal and liver function in rats. *J. Pharmacol. Toxicol. Methods*. 54(2):200–214.
- Pfordte K. (1971) Polyäthylenglykol, ein wenig toxisches lösemittel zur intravenösen applikation wasserunlöslicher substanzen. *Zentralbl Pharm.* 110:449–451.
- PharmPK Discussion. (2002a) *Cosolvents/Surfactants in IV Formulations for Pre-clinical Studies*, October 2, 2002. Available at: <http://www.pharmpk.com/PK02/PK2002075.html> (accessed March 31, 2015).
- PharmPK Discussion. (2002b) *Cremophor EL Toxicity*, October 3, 2002. Fiona Stavros Response. Available at: <https://www.pharmpk.com/PK02/PK2002075.html> (accessed May 30, 2016).
- Prasamthi K, Muralidhara PK, Rajini PS. (2005) Fenvalerate-induced oxidative damage in rat tissues and its attenuation by dietary sesame oil. *Food Chem. Toxicol.* 43(2):299–306.
- Quadbeck G. (1950) Äthylenglykol und verwandte verbindungen als pharmazeutische lösungsmittel. *Chem. Z.* 74:91–94.
- Rajewski RA, Taiger G, Bresnahan J, Jaberaboansari P, Stella VJ, Thompson DO. (1995) Preliminary safety evaluation of parenterally administered sulfoalkyl ether beta-cyclodextrin derivatives. *J. Pharm. Sci.* 84(8):927–932.
- Ramadan LA, El-Habit OH, Arafa H, Sayed-Ahmad MM. (2001) Effect of cremophor-EL on cisplatin-induced organ toxicity in normal rat. *J. Egypt. Natl. Canc. Inst.* 13(2):139–145.
- Ratsimbarzafy V, Bourret E, Duclos R, Brossard C. (1999) Rheological behavior of drug suspensions in Gelucire mixtures and proxiphylline release from matrix hard gelatin capsules. *Eur. J. Pharm. Biopharm.* 48:247–252.
- RTECS. (2015) *Registry of Toxic Effects of Chemical Substances (RTECS)—DMSO*.
- Robertson RP, Harmon J, Tran PO, Tanaka Y, Takahashi H. (2003) Glucose toxicity in B-cells: type 2 diabetes, good radicals, one bad, and the glutathione connection. *Diabetes*. 52:581–587.
- Rosen PA, Brooks AMV, Ramsay RJ, Gullifer KP, Wolfe RJ, Gillies WE. (1987) Efficacy and safety of hypromellose in ocular implant surgery. *Aust. N. Z. J. Ophthalmol.* 15:193–199.
- Rowe VK, Wolf MA. (1982) Polyethylene glycols. In: *Patty's Industrial Hygiene and Toxicology*, Vol. 2C (Clayton GD, Clayton FE). John Wiley & Sons, Inc., New York, pp. 3844–3852, 3902, 3904–3905.
- Rowe RC, Sheskey PJ, Cook WG, Fenton ME. (Eds.) (2012) *Handbook of Pharmaceutical Excipients*, 7th Ed. Pharmaceutical Press, London.
- Ruble GR, Giardino OZ, Fosscedo SL, Cosmatos D, Knapp RJ, Barlow NJ. (2006) The effects of commonly used vehicles on canine hematology and clinical chemistry values. *J. Am. Assoc. Lab. Anim. Sci.* 45(1):25–29.
- Ruchatz E. (1998) *Applications of Solutol® HS15—A Potent Solubilizer with a Low Toxicity*. BASF, Ludwigshafen, Germany.
- Schonwald S. (2004a) Irrigating solutions. In: (Dart RC, Michael A, McCuigan MA) *Medical Toxicology*, 3rd Ed. Lippincott Williams & Wilkins, Philadelphia, PA, p 918–920.
- Schonwald S. (2004b) Acidifying and alkalinizing agents. In: (Dart RC, Michael A, McCuigan MA) *Medical Toxicology*, 3rd Ed. Lippincott Williams & Wilkins, Philadelphia, PA, p 911–913.
- Sellers RS, Antman M, Phillips J, Khan KN, Furst SM. (2005) Effects of miglyol 812 on rats after 4 weeks of gavage as compared with methylcellulose/Tween 80. *Drug Chem. Toxicol.* 28:423–432.
- Serbest G, Horwitz J, Barbee K. (2005) The effect of poloxamer-188 on neuronal cell recovery from mechanical injury. *J. Neurotrauma*. 22(1):119–132.
- Shadnia S, Rahimi M, Pajoumand A, Rasouli MH, Abdollahi M. (2005) Successful treatment of acute aluminium phosphide poisoning: possible benefit of coconut oil. *Hum. Exp. Toxicol.* 24(4):215–218.
- Sharma PK, Chaudhari PD, Badagale KD, Kulkarni PA, Barhate NS. (2006) *Current Trends in Solid Dispersions Techniques*. Pharmainfotech. A Company, New Zealand.
- Shertzer HG. (2012) Methanesulfonic acid. In: *Patty's Toxicology* (Bingham E and Cohns B, Eds.), 6th Ed, Vol. 4. John Wiley & Sons, Inc., Hoboken, NJ, p. 1063.
- Shideman FE, Procita L. (1951) Some pharmacological actions of polypropylene glycols of average molecular weight 400, 750, 1200, and 2000. *J. Pharmacol. Exp. Ther.* 103:293–305.
- Sivilotti MLA. (2004) Ethanol, isopropanol, and methanol. In: *Medical Toxicology* (Dart RC, Michael A and McCuigan MA, Eds.), 3rd Ed. Lippincott Williams & Wilkins, Philadelphia, PA, pp. 1215–1216.
- Smyth HF Jr, Carpenter CP, Shaffer CB, Seaton J, Fischer L. (1942) Some pharmacological properties of polyethylene glycols of high molecular weight (“carbawax” compounds). *J. Ind. Hyg. Toxicol.* 24:281–284.
- Smyth HF Jr, Carpenter CP, Shaffer CB. (1945) The subacute toxicity and irritation of polyethylene glycols of approximate molecular weights of 200, 300, 400. *J. Am. Pharm. Assoc.* 34:172–174.
- Smyth HF Jr, Carpenter CP, Shaffer CB. (1947) The toxicity of high molecular weight polyethylene glycols; chronic oral and parenteral administration. *J. Am. Pharm. Assoc.* 36:157–160.
- Smyth HF Jr, Carpenter CP, Weil CS. (1950) The toxicology of polyethylene glycols. *J. Am. Pharm. Assoc.* 39:349–354.
- Smyth HF Jr, Carpenter CP, Weil CS. (1955) The chronic oral toxicity of the polyethylene glycols. *J. Am. Pharm. Assoc.* 44:27–30.

- Smyth HF Jr, Weil CS, Woodside MD, Knaak JB, Sullivan LJ, Carpenter CP. (1970) Experimental toxicity of high molecular weight poly(ethylene oxide). *Toxicol. Appl. Pharmacol.* 16:442–445.
- Sodicoff M, Lamperti A, Ziskin MC. (1990) Transdermal absorption of radioprotectors using permeation-enhancing vehicles. *Radiat. Res.* 121(2):212–219.
- Solomon GM, Morse EP, Garbo MJ, Milton DK. (1996) Stillbirth after occupational exposure to n-methyl-2-pyrrolidone: a case report and review of the literature. *J. Occup. Environ. Med.* 38(7):705–713.
- Sourkes TL, Koppányi T. (1950) Correlation between the acute toxicity and the rate of elimination of tartaric acid and certain of its esters. *J. Am. Pharm. Assoc.* 39(5):275–276.
- Stella VJ, He Q. (2008) Cyclodextrins. *Toxicol. Pathol.* 36(1):30–42.
- Stokes AH, Kemp DC, Faiola B, Jordan HL, Merrill CL, Hailey JR, Brown RE, Bailey DW. (2013) Effects of Solutol (Kolliphor) and cremophor in polyethylene glycol 400 vehicle formulations in Sprague-Dawley rats and beagle dogs. *Int. J. Toxicol.* 32(3):189–197.
- Strickley RG. (2008) Formulation in drug discovery. In: *Annual Reports in Medicinal Chemistry* (Macor JE, Ed.), Vol. 43. Academic Press, London, pp. 419–451.
- Sullivan DW Jr, Gad SC, Julien M. (2014) A review of the nonclinical safety of Transcutol®, a highly purified form of diethylene glycol monoethyl ether (DEGEE) used as a pharmaceutical excipient. *Food Chem. Toxicol.* 72:40–50.
- Susananta W, Craid DQ, Newton JM. (1995) An evaluation of the mechanisms of drug release from glyceride bases. *J. Pharm. Pharmacol.* 47(3):182–187.
- Szilagyí M. (2012a) Acetic acid. In: *Patty's Toxicology* (Bingham E and Cohns B, Eds.), 6th Ed, Vol. 3. John Wiley & Sons, Inc., Hoboken, NJ, pp. 477–480.
- Szilagyí M. (2012b) Citric acid. In: *Patty's Toxicology* (Bingham E and Cohns B, Eds.), 6th Ed, Vol. 3. John Wiley & Sons, Inc., Hoboken, NJ, pp. 519–521.
- Szilagyí M. (2012c) Succinic acid. In: *Patty's Toxicology* (Bingham E and Cohns B, Eds.), 6th Ed, Vol. 3. John Wiley & Sons, Inc., Hoboken, NJ, pp. 513–514.
- Szilagyí M. (2012d) Tartaric acid. In: *Patty's Toxicology* (Bingham E and Cohns B, Eds.), 6th Ed, Vol. 3. John Wiley & Sons, Inc., Hoboken, NJ, pp. 516–517.
- Takahashi KK, Onodera K, Akiba Y. (1999) Effect of dietary Xylitol on growth and inflammatory responses in immune stimulated chickens. *Br. Poult. Sci.* 40:552–554.
- Talsma J. (2004) Anecortave acetate proving helpful for inhibiting CNV growth. *Ophthalmol. Times.* 29(6):34.
- Temple BR. (2004) Vitamin toxicity. In: *Medical Toxicology* (Dart RC, Michael A and McCuigan MA, Eds.), 3rd Ed. Lippincott Williams & Wilkins, Philadelphia, PA, pp. 1021–1023.
- Thackaberry EA. (2014) Solvent-based formulations for intravenous mouse pharmacokinetic studies: tolerability and recommended solvent dose limits. *Xenobiotica.* 44(3):235–241.
- Thackaberry EA, Kopytek S, Sherratt P, Trouba K, McIntyre B. (2009) Comprehensive investigation of hydroxypropyl methylcellulose, propylene glycol, polysorbate 80, and hydroxypropyl-beta-cyclodextrin for use in general toxicology studies. *Toxicol. Sci.* 117(2):485–492.
- TOXNET. (2015a) *Arabic Gum*. Available at: <http://toxnet.nlm.nih.gov/> (accessed August 7, 2015).
- TOXNET. (2015aa) *Cetyl Alcohol*. Available at: <http://toxnet.nlm.nih.gov/> (accessed August 7, 2015).
- TOXNET. (2015ab) *Coconut Oil*. Available at: <http://toxnet.nlm.nih.gov/> (accessed August 7, 2015).
- TOXNET. (2015abs) *Lactose*. Available at: <http://toxnet.nlm.nih.gov/> (accessed August 7, 2015).
- TOXNET. (2015ac) *Corn Oil*. Available at: <http://toxnet.nlm.nih.gov/> (accessed August 7, 2015).
- TOXNET. (2015ad) *Cottonseed Oil*. Available at: <http://toxnet.nlm.nih.gov/> (accessed August 7, 2015).
- TOXNET. (2015ae) *Cyclodextrin*. Available at: <http://toxnet.nlm.nih.gov/> (accessed August 7, 2015).
- TOXNET. (2015af) *Cyclohexane*. Available at: <http://toxnet.nlm.nih.gov/> (accessed August 7, 2015).
- TOXNET. (2015ag) *Cyclomethicone NF*. Available at: <http://toxnet.nlm.nih.gov/> (accessed August 7, 2015).
- TOXNET. (2015ah) *Diacetylmonoxime*. Available at: <http://toxnet.nlm.nih.gov/> (accessed August 7, 2015).
- TOXNET. (2015ai) *Dextrose*. Available at: <http://toxnet.nlm.nih.gov/> (accessed August 7, 2015).
- TOXNET. (2015aj) *Dichlorvos*. Available at: <http://toxnet.nlm.nih.gov/> (accessed August 7, 2015).
- TOXNET. (2015ak) *N,N-Diethylacetamide*. Available at: <http://toxnet.nlm.nih.gov/> (accessed August 7, 2015).
- TOXNET. (2015al) *Diethylene Glycol Monoethyl Ether*. Available at: <http://toxnet.nlm.nih.gov/> (accessed August 7, 2015).
- TOXNET. (2015am) *Dimethicone*. Available at: <http://toxnet.nlm.nih.gov/> (accessed August 7, 2015).
- TOXNET. (2015an) *N,N-Dimethylacetamide*. Available at: <http://toxnet.nlm.nih.gov/> (accessed August 7, 2015).
- TOXNET. (2015ao) *Dimethylsulfoxide*. Available at: <http://toxnet.nlm.nih.gov/> (accessed August 7, 2015).
- TOXNET. (2015ap) *Di-Sodium Hydrogen Phosphate Dehydrate*. Available at: <http://toxnet.nlm.nih.gov/> (accessed August 7, 2015).
- TOXNET. (2015aq) *Docosanol*. Available at: <http://toxnet.nlm.nih.gov/> (accessed August 7, 2015).
- TOXNET. (2015ar) *Edetic Acid*. Available at: <http://toxnet.nlm.nih.gov/> (accessed August 7, 2015).
- TOXNET. (2015as) *Ethyl Alcohol*. Available at: <http://toxnet.nlm.nih.gov/> (accessed August 7, 2015).
- TOXNET. (2015at) *Fumaric Acid*. Available at: <http://toxnet.nlm.nih.gov/> (accessed August 7, 2015).
- TOXNET. (2015au) *Gelatin*. Available at: <http://toxnet.nlm.nih.gov/> (accessed August 7, 2015).
- TOXNET. (2015av) *Gelucire 44/14*. Available at: <http://toxnet.nlm.nih.gov/> (accessed August 7, 2015).
- TOXNET. (2015aw) *Gelucire 50/13*. Available at: <http://toxnet.nlm.nih.gov/> (accessed August 7, 2015).
- TOXNET. (2015ax) *Gluconic Acid*. Available at: <http://toxnet.nlm.nih.gov/> (accessed August 7, 2015).

- TOXNET. (2015ay) *Glycerol*. Available at: <http://toxnet.nlm.nih.gov/> (accessed August 7, 2015).
- TOXNET. (2015az) *Glycofurol*. Available at: <http://toxnet.nlm.nih.gov/> (accessed August 7, 2015).
- TOXNET. (2015b) *Acetate Ion*. Available at: <http://toxnet.nlm.nih.gov/> (accessed August 7, 2015).
- TOXNET. (2015ba) *Tragacanth*. Available at: <http://toxnet.nlm.nih.gov/> (accessed August 7, 2015).
- TOXNET. (2015bb) *Xanthan Gum*. Available at: <http://toxnet.nlm.nih.gov/> (accessed August 7, 2015).
- TOXNET. (2015bc) *Histidine*. Available at: <http://toxnet.nlm.nih.gov/> (accessed August 7, 2015).
- TOXNET. (2015bd) *Hydrochloric Acid*. Available at: <http://toxnet.nlm.nih.gov/> (accessed August 7, 2015).
- TOXNET. (2015be) *Hydrogenated Castor Oil*. Available at: <http://toxnet.nlm.nih.gov/> (accessed August 7, 2015).
- TOXNET. (2015bf) *Hydroxyethylcellulose*. Available at: <http://toxnet.nlm.nih.gov/> (accessed August 7, 2015).
- TOXNET. (2015bg) *Hydroxypropyl Betadex*. Available at: <http://toxnet.nlm.nih.gov/> (accessed August 7, 2015).
- TOXNET. (2015bh) *Hydroxypropyl Cellulose*. Available at: <http://toxnet.nlm.nih.gov/> (accessed August 7, 2015).
- TOXNET. (2015bi) *Hydroxypropyl Methylcellulose*. Available at: <http://toxnet.nlm.nih.gov/> (accessed August 7, 2015).
- TOXNET. (2015bj) *Hypromellose Acetate Succinate*. Available at: <http://toxnet.nlm.nih.gov/> (accessed August 7, 2015).
- TOXNET. (2015bk) *Hymetellose*. Available at: <http://toxnet.nlm.nih.gov/> (accessed August 7, 2015).
- TOXNET. (2015bl) *Isopropyl Alcohol*. Available at: <http://toxnet.nlm.nih.gov/> (accessed August 7, 2015).
- TOXNET. (2015bm) *Isopropyl Myristate*. Available at: <http://toxnet.nlm.nih.gov/> (accessed August 7, 2015).
- TOXNET. (2015bn) *Polyoxyl 35 Castor Oil*. Available at: <http://toxnet.nlm.nih.gov/> (accessed August 7, 2015).
- TOXNET. (2015bo) *Polyoxyl 40 Hydrogenated Castor Oil*. Available at: <http://toxnet.nlm.nih.gov/> (accessed August 7, 2015).
- TOXNET. (2015bp) *Labrafil*. Available at: <http://toxnet.nlm.nih.gov/> (accessed August 7, 2015).
- TOXNET. (2015bq) *Ringer's Lactate Solution*. Available at: <http://toxnet.nlm.nih.gov/> (accessed August 7, 2015).
- TOXNET. (2015br) *Lactic Acid*. Available at: <http://toxnet.nlm.nih.gov/> (accessed August 7, 2015).
- TOXNET. (2015bt) *Lanolin*. Available at: <http://toxnet.nlm.nih.gov/> (accessed August 7, 2015).
- TOXNET. (2015bu) *Lanolin Alcohols*. Available at: <http://toxnet.nlm.nih.gov/> (accessed August 7, 2015).
- TOXNET. (2015bv) *L-Arginine Hydrochloride*. Available at: <http://toxnet.nlm.nih.gov/> (accessed August 7, 2015).
- TOXNET. (2015bw) *Ascorbic Acid*. Available at: <http://toxnet.nlm.nih.gov/> (accessed August 7, 2015).
- TOXNET. (2015bx) *Propylene Glycol*. Available at: <http://toxnet.nlm.nih.gov/> (accessed August 7, 2015).
- TOXNET. (2015by) *Lycasin*. Available at: <http://toxnet.nlm.nih.gov/> (accessed August 7, 2015).
- TOXNET. (2015bz) *Maltol*. Available at: <http://toxnet.nlm.nih.gov/> (accessed August 7, 2015).
- TOXNET. (2015c) *Acetic Acid*. Available at: <http://toxnet.nlm.nih.gov/> (accessed August 7, 2015).
- TOXNET. (2015ca) *d-Mannitol*. Available at: <http://toxnet.nlm.nih.gov/> (accessed August 7, 2015).
- TOXNET. (2015cb) *Methanesulfonic Acid*. Available at: <http://toxnet.nlm.nih.gov/> (accessed August 7, 2015).
- TOXNET. (2015cc) *Methyl Methacrylate*. Available at: <http://toxnet.nlm.nih.gov/> (accessed August 7, 2015).
- TOXNET. (2015cd) *Methylcellulose*. Available at: <http://toxnet.nlm.nih.gov/> (accessed August 7, 2015).
- TOXNET. (2015ce) *Methylparaben*. Available at: <http://toxnet.nlm.nih.gov/> (accessed August 7, 2015).
- TOXNET. (2015cf) *Methylpyrrolidone*. Available at: <http://toxnet.nlm.nih.gov/> (accessed August 7, 2015).
- TOXNET. (2015cg) *Mineral Oil*. Available at: <http://toxnet.nlm.nih.gov/> (accessed August 7, 2015).
- TOXNET. (2015ch) *Myristyl Alcohol*. Available at: <http://toxnet.nlm.nih.gov/> (accessed August 7, 2015).
- TOXNET. (2015ci) *Octoxynol*. Available at: <http://toxnet.nlm.nih.gov/> (accessed August 7, 2015).
- TOXNET. (2015cj) *Oleic Acid*. Available at: <http://toxnet.nlm.nih.gov/> (accessed August 7, 2015).
- TOXNET. (2015ck) *Oleyl Alcohol*. Available at: <http://toxnet.nlm.nih.gov/> (accessed August 7, 2015).
- TOXNET. (2015cl) *Olive Oil*. Available at: <http://toxnet.nlm.nih.gov/> (accessed August 7, 2015).
- TOXNET. (2015cm) *Panthenol*. Available at: <http://toxnet.nlm.nih.gov/> (accessed August 7, 2015).
- TOXNET. (2015cn) *Peanut Oil*. Available at: <http://toxnet.nlm.nih.gov/> (accessed August 7, 2015).
- TOXNET. (2015co) *Glyceryl Monooleate*. Available at: <http://toxnet.nlm.nih.gov/> (accessed August 7, 2015).
- TOXNET. (2015cp) *Macrogol*. Available at: <http://toxnet.nlm.nih.gov/> (accessed August 7, 2015).
- TOXNET. (2015cq) *Petrolatum*. Available at: <http://toxnet.nlm.nih.gov/> (accessed August 7, 2015).
- TOXNET. (2015cr) *Phenoxyethanol*. Available at: <http://toxnet.nlm.nih.gov/> (accessed August 7, 2015).
- TOXNET. (2015cs) *Phosphate Ion*. Available at: <http://toxnet.nlm.nih.gov/> (accessed August 7, 2015).
- TOXNET. (2015ct) *Poloxalene*. Available at: <http://toxnet.nlm.nih.gov/> (accessed August 7, 2015).
- TOXNET. (2015cu) *Decaglyceryl Monooleate*. Available at: <http://toxnet.nlm.nih.gov/> (accessed August 7, 2015).
- TOXNET. (2015cv) *Povidone*. Available at: <http://toxnet.nlm.nih.gov/> (accessed August 7, 2015).
- TOXNET. (2015cw) *Potassium Chloride*. Available at: <http://toxnet.nlm.nih.gov/> (accessed August 7, 2015).
- TOXNET. (2015cx) *Propylene Glycol Dicaprylate/Dicaprate*. Available at: <http://toxnet.nlm.nih.gov/> (accessed August 7, 2015).
- TOXNET. (2015cy) *Propylparaben*. Available at: <http://toxnet.nlm.nih.gov/> (accessed August 7, 2015).

- TOXNET. (2015cz) *Copovidone*. Available at: <http://toxnet.nlm.nih.gov/> (accessed August 7, 2015).
- TOXNET. (2015d) *Acetone*. Available at: <http://toxnet.nlm.nih.gov/> (accessed August 7, 2015).
- TOXNET. (2015da) *Safflower Oil*. Available at: <http://toxnet.nlm.nih.gov/> (accessed August 7, 2015).
- TOXNET. (2015db) *Salicylic Acid*. Available at: <http://toxnet.nlm.nih.gov/> (accessed August 7, 2015).
- TOXNET. (2015dc) *Sesame Oil*. Available at: <http://toxnet.nlm.nih.gov/> (accessed August 7, 2015).
- TOXNET. (2015dd) *Shea Butter*. Available at: <http://toxnet.nlm.nih.gov/> (accessed August 7, 2015).
- TOXNET. (2015de) *Simethicone*. Available at: <http://toxnet.nlm.nih.gov/> (accessed August 7, 2015).
- TOXNET. (2015df) *Sodium Acetate*. Available at: <http://toxnet.nlm.nih.gov/> (accessed August 7, 2015).
- TOXNET. (2015dg) *Sodium Citrate*. Available at: <http://toxnet.nlm.nih.gov/> (accessed August 7, 2015).
- TOXNET. (2015dh) *Sodium Phosphate, Monobasic, Dihydrate*. Available at: <http://toxnet.nlm.nih.gov/> (accessed August 7, 2015).
- TOXNET. (2015di) *Sodium Hydroxide*. Available at: <http://toxnet.nlm.nih.gov/> (accessed August 7, 2015).
- TOXNET. (2015dj) *Sodium Lauryl Sulfate*. Available at: <http://toxnet.nlm.nih.gov/> (accessed August 7, 2015).
- TOXNET. (2015dk) *Sodium Metabisulfite*. Available at: <http://toxnet.nlm.nih.gov/> (accessed August 7, 2015).
- TOXNET. (2015dl) *Methylparaben Sodium*. Available at: <http://toxnet.nlm.nih.gov/> (accessed August 7, 2015).
- TOXNET. (2015dm) *Sodium Phosphate*. Available at: <http://toxnet.nlm.nih.gov/> (accessed August 7, 2015).
- TOXNET. (2015dn) *Propylparaben Sodium*. Available at: <http://toxnet.nlm.nih.gov/> (accessed August 7, 2015).
- TOXNET. (2015do) *Butanedioic Acid, Disodium Salt*. Available at: <http://toxnet.nlm.nih.gov/> (accessed August 7, 2015).
- TOXNET. (2015dp) *Sodium Sulfite*. Available at: <http://toxnet.nlm.nih.gov/> (accessed August 7, 2015).
- TOXNET. (2015dq) *Solutol HS15*. Available at: <http://toxnet.nlm.nih.gov/> (accessed August 7, 2015).
- TOXNET. (2015dr) *Sorbitan Tristearate*. Available at: <http://toxnet.nlm.nih.gov/> (accessed August 7, 2015).
- TOXNET. (2015ds) *Sorbitol*. Available at: <http://toxnet.nlm.nih.gov/> (accessed August 7, 2015).
- TOXNET. (2015dt) *Soybean Oil*. Available at: <http://toxnet.nlm.nih.gov/> (accessed August 7, 2015).
- TOXNET. (2015du) *Squalene*. Available at: <http://toxnet.nlm.nih.gov/> (accessed August 7, 2015).
- TOXNET. (2015dv) *Stearic Acid*. Available at: <http://toxnet.nlm.nih.gov/> (accessed August 7, 2015).
- TOXNET. (2015dw) *Stearyl Alcohol*. Available at: <http://toxnet.nlm.nih.gov/> (accessed August 7, 2015).
- TOXNET. (2015dx) *Sucrose*. Available at: <http://toxnet.nlm.nih.gov/> (accessed August 7, 2015).
- TOXNET. (2015dy) *Sucrose Acetate Isobutyrate*. Available at: <http://toxnet.nlm.nih.gov/> (accessed August 7, 2015).
- TOXNET. (2015dz) *Sulfobutylether Beta Cyclodextrin*. Available at: <http://toxnet.nlm.nih.gov/> (accessed August 7, 2015).
- TOXNET. (2015e) *Acetonitrile*. Available at: <http://toxnet.nlm.nih.gov/> (accessed August 7, 2015).
- TOXNET. (2015ea) *Tartaric Acid*. Available at: <http://toxnet.nlm.nih.gov/> (accessed August 7, 2015).
- TOXNET. (2015eb) *Terbinafine Hydrochloride*. Available at: <http://toxnet.nlm.nih.gov/> (accessed August 7, 2015).
- TOXNET. (2015ec) *Tetraglycol*. Available at: <http://toxnet.nlm.nih.gov/> (accessed August 7, 2015).
- TOXNET. (2015ed) *Trehalose*. Available at: <http://toxnet.nlm.nih.gov/> (accessed August 7, 2015).
- TOXNET. (2015ee) *Tromethamine*. Available at: <http://toxnet.nlm.nih.gov/> (accessed August 7, 2015).
- TOXNET. (2015ef) *Polysorbate 20*. Available at: <http://toxnet.nlm.nih.gov/> (accessed August 7, 2015).
- TOXNET. (2015eg) *Polysorbate 80*. Available at: <http://toxnet.nlm.nih.gov/> (accessed August 7, 2015).
- TOXNET. (2015eh) *Alpha-Tocopherol*. Available at: <http://toxnet.nlm.nih.gov/> (accessed August 7, 2015).
- TOXNET. (2015ei) *Tocophersolan*. Available at: <http://toxnet.nlm.nih.gov/> (accessed August 7, 2015).
- TOXNET. (2015ej) *White Wax*. Available at: <http://toxnet.nlm.nih.gov/> (accessed August 7, 2015).
- TOXNET. (2015ek) *Xylitol*. Available at: <http://toxnet.nlm.nih.gov/> (accessed August 7, 2015).
- TOXNET. (2015f) *Acetylated Lanolin Alcohol*. Available at: <http://toxnet.nlm.nih.gov/> (accessed August 7, 2015).
- TOXNET. (2015g) *N-Methylacetamide*. Available at: <http://toxnet.nlm.nih.gov/> (accessed August 7, 2015).
- TOXNET. (2015h) *Alginate Acid*. Available at: <http://toxnet.nlm.nih.gov/> (accessed August 7, 2015).
- TOXNET. (2015i) *Anecortave Acetate*. Available at: <http://toxnet.nlm.nih.gov/> (accessed August 7, 2015).
- TOXNET. (2015j) *Cellulose, Microcrystalline*. Available at: <http://toxnet.nlm.nih.gov/> (accessed August 7, 2015).
- TOXNET. (2015k) *Benzoic Acid*. Available at: <http://toxnet.nlm.nih.gov/> (accessed August 7, 2015).
- TOXNET. (2015l) *Benzyl Alcohol*. Available at: <http://toxnet.nlm.nih.gov/> (accessed August 7, 2015).
- TOXNET. (2015m) *beta-Cyclodextrin*. Available at: <http://toxnet.nlm.nih.gov/> (accessed August 7, 2015).
- TOXNET. (2015n) *Butylated Hydroxyanisole*. Available at: <http://toxnet.nlm.nih.gov/> (accessed August 7, 2015).
- TOXNET. (2015o) *Butylated Hydroxytoluene*. Available at: <http://toxnet.nlm.nih.gov/> (accessed August 7, 2015).
- TOXNET. (2015p) *Bicarbonate Ion*. Available at: <http://toxnet.nlm.nih.gov/> (accessed August 7, 2015).
- TOXNET. (2015q) *Calcium Chloride*. Available at: <http://toxnet.nlm.nih.gov/> (accessed August 7, 2015).

- TOXNET. (2015r) *Canola Oil*. Available at: <http://toxnet.nlm.nih.gov/> (accessed August 7, 2015).
- TOXNET (2015s) *Beta-Cyclodextrin Sulfobutyl Ether Sodium Salt*. Available at: <http://toxnet.nlm.nih.gov/> (accessed August 7, 2015).
- TOXNET. (2015t) *Carbomer Homopolymer Type B (Allyl Pentaerythritol Crosslinked)*. Available at: <http://toxnet.nlm.nih.gov/> (accessed August 7, 2015).
- TOXNET. (2015u) *Carbopol Ultrez 10*. Available at: <http://toxnet.nlm.nih.gov/> (accessed August 7, 2015).
- TOXNET. (2015v) *Carboxymethylcellulose*. Available at: <http://toxnet.nlm.nih.gov/> (accessed August 7, 2015).
- TOXNET. (2015w) *Carboxymethylcellulose Calcium*. Available at: <http://toxnet.nlm.nih.gov/> (accessed August 7, 2015).
- TOXNET. (2015x) *Carboxymethylcellulose Sodium*. Available at: <http://toxnet.nlm.nih.gov/> (accessed August 7, 2015).
- TOXNET. (2015y) *Cetostearyl Alcohol*. Available at: <http://toxnet.nlm.nih.gov/> (accessed August 7, 2015).
- TOXNET. (2015z) *Cetyl Acetate*. Available at: <http://toxnet.nlm.nih.gov/> (accessed August 7, 2015).
- Toyoda K, Shoda T, Uneyama C, Takada K, Takahashi M. (1997) Carcinogenicity study of beta-cyclodextrin in F344 rats. *Food Chem. Toxicol.* 35(3–4):331–336.
- Traul KA, Driedger A, Ingle DL, Nakhasi D. (2000) Review of the toxicologic properties of medium-chain triglycerides. *Food Chem. Toxicol.* 38(1):79–98.
- Trimmer GW, Freeman JJ, Priston RA, Urbanus J. (2004) Results of chronic dietary toxicity studies of high viscosity (P70H and P100H) white mineral oils in Fischer 344 rats. *Toxicol. Pathol.* 32(4):439–447.
- Tyl RW, Masten LW, Marr MC, Myers CB, Slaughter RW, Gardiner TH, Strother DE, McKee RH, Tyler TR. (1994) Developmental toxicity evaluation of isopropanol by gavage in rats and rabbits. *Fundam. Appl. Toxicol.* 22(1):139–151.
- Ueda H, Ou D, Endo T, Nagase H, Tomono K, Nagai T. (1998) Evaluation of a sulfobutyl ether beta-cyclodextrin as solubilizing agent for several drugs. *Drug Dev. Ind. Pharm.* 24(9):863–867.
- Walker R, El Harith EA. (1978) Nutritional and toxicological properties of some raw and modified starches. *Ann. Nutr. Aliment.* 32(2–3):671–679.
- Waner T, Borelli G, Cadel S, Privman I, Nyska A. (1995) Investigation of potential oncogenetic effects of beta-cyclodextrin in the rat and mouse. *Arch. Toxicol.* 69(9):631–639.
- White CW, Rodriguez J, Marrs GE Jr. (1984) *Acute Oral Toxicity of DMSO (Dimethyl Sulfoxide) Process Stream Samples in Male and Female Rats*. Govt Reports Announcements & Index (GRA&I), Issue 9, National Technical Information Service (NTIS), Springfield, VA.
- Wieland TM, Lin X, Odle J. (1993) Utilization of medium-chain triglycerides by neonatal pigs: effects of emulsification and dose delivered. *J. Anim. Sci.* 71(7):1863–1868.
- Wood DC, Weber FS, Palmquist MA. (1971) Continued studies in the toxicology of dimethyl sulfoxide (DMSO). *J. Pharmacol. Exp. Ther.* 177(3):520–527.
- Working PK, Newman MS, Johnson J, Cornacoff JB. (1997) Safety of poly(ethylene glycol) and poly (ethylene glycol) derivatives. *Am. Chem. Soc. Symp. Ser.* 680(4):45–57.
- Wu B, Iwakiri R, Ootani A, Tsunada S, Fujise T, Sakata Y, Sakata H, Toda S, Fujimoto K. (2004) Dietary corn oil promotes colon cancer by inhibiting mitochondria-dependent apoptosis in azoxymethane-treated rats. *Exp. Biol. Med.* 229:1017–1025.
- Xcess Biosciences, Inc. (n.d.) *Vehicle Guidance for Preclinical Studies*. San Diego, CA. Available at: <http://xcessbiosciences.com/uploads/tmp/in%20vivo%20study%20formulation%20guidance.pdf> (accessed: June 5, 2015).

APPENDIX F

GLOBAL DIRECTORY OF CONTRACT PHARMACEUTICAL TOXICOLOGY LABS

| Lab Name (Affiliated Labs/ Past Names) | | Location(s) | Phone #(s) | Website | Additional Services | Rat/ Mouse | | Dog | Primate | IV | | | Special | | |
|--|--|--|--|--------------------|--|--|---|-----|---------|----|---|---|---------|----|--|
| Absorption systems | | Exton, PA San Diego, CA | (610)280-7300 (858)560-9000 | www.absorption.com | <i>In vitro</i> , metabolism, hERG | X | X | X | X | X | X | X | X | SP | |
| Accelerera | | Viale Pasteur 10—CP11 | +39 0331 1984 444 | www.accelera.org | | X | X | | X | | | | | | |
| ADVINUS | | Bangalore, India | +91-80-28394959 +91-80-28394015 | www.rallis.co.in | Yes | X | X | X | X | X | | X | X | SP | |
| Advinus Therapeutics Ltd | | Plot No. 21 and 22, Phase-II, PO Box No. 5813, Peenya Industrial Area, Bangalore 560 058, India | 91-80-28394959, 28397338 (p) 91-80-28394015, 28396023 (f) | www.advinus.com | Physical–chemical testing | Toxicity studies (acute, skin sensitization, subacute/ subchronic, chronic, reproductive, carcinogenicity, inhalation) | | | | | | | | | |
| | | | | | Mutagenicity studies (chromosomal aberration test <i>in vivo</i> and <i>in vivo</i> , micronucleus assay <i>in vitro</i> and <i>in vivo</i> , Ames test, gene mutation test <i>in vitro</i>) | | | | | | | | | | |
| | | | | | Environmental toxicity on aquatic and terrestrial organisms | | | | | | | | | | |
| | | | | | Studies on behavior in water, soil, and air, bioaccumulation | | | | | | | | | | |
| | | | | | Residue studies | | | | | | | | | | |
| | | | | | Analytical and clinical chemistry | | | | | | | | | | |
| | | | | | Bioanalytical and TK | | | | | | | | | | |
| | | | | | Drug metabolism and PK and tissue distribution | | | | | | | | | | |
| | | | | | Safety pharmacology | | | | | | | | | | |
| | | | | | GLP | | | | | | | | | | |

[illegible]

(Continued)

[illegible]

(Continued)

| Lab Name (Affiliated Labs/ Past Names) | Location(s) | Phone #(s) | Website | Rat/ Mouse | | IV | | | | Special | | | | | |
|---|---|---|--------------------|-----------------------------------|------|------------|---------|----------------|----------|-----------|---------|------------|------------|---------|----|
| | | | | Additional Services | DART | Inhalation | Primate | Pig | Infusion | Genotoxic | Devices | Metabolism | Analytical | Studies | |
| Drik | 865 Research Parkway, Suite 415, Oklahoma City, OK 73104 | (405) 202-0117 | www.atdrik.com | | X | | | | | | | | | | |
| ENVIGO (now includes Huntingdon, Harlan, and RCC Labs) | East Millstone, NJ | (732) 873-2550 | www.huntingdon.com | Discovery support | X | X | X | X | X | X | X | X | X | X | SP |
| | Cambridgeshire, UK | 44 (0) 1480 892000 | | Radioactivity | X | X | | X ^a | X | | | | | | X |
| | Rehovot, Israel | 972 (0) 8-9409451 | No | Bioproducts, breeder | X | X | | X | | | | X | X | X | SP |
| | Zelgliweg 1, CH-4452 Itingen, Switzerland (previously Safety Pharm) | (410) 385-1666—US sales office +41 61 975 1111 (p) lchaney@harlan.com www.harlan.com | www.rcc.ch | In general | X | X | X | X | X | X | X | X | X | X | SP |
| Germany | | | | | | | | | | | | | | | |
| In vitro studies | | | | | | | | | | | | | | | |
| Spain | | 34-937-190361 | | | X | X | X | X | X | X | X | X | X | X | SP |
| Madison, WI | | (317)806-6080 x2922 | | Nonnaïve canines/NHP | | X | | | X | | | X | | | |
| Indianapolis, IN | | (317)806-6080 x2922 | | Discovery support | X | X | | | | X | X | X | X | | |
| Derby, UK | | 00 44 (0) 1332 792896 | | Yes (fish +) | X | | X | X | X | X | X | X | X | X | SP |
| Enviro-Bio-Tech-Ltd | Bernville, PA | (610)488-7664 | No | (Chicken, cattle) (food animals) | | | | | | | | | X | WL | |
| Estendart Limited/Massey University | Palmerston North 4442, New Zealand | +64(0)6 350 0782 (p) +64(0)6 350 0772 (f) | www.estendart.com | (Sheep, cattle, chickens, horses) | X | | X | | | | X | | | | |
| Eurofins | Dayton, NJ | 732-438-5100 ext 270 732-355-3275 (f) Cel 732-744-6579 | www.eurofins.com | | | | | | | | X | X | | | |
| Evans Labs | Missouri, IA | (573)474-8579 | www.abclabs.com | Bioanalysis, methods development | X | | | X | | | | X | X | X | |
| | N. Ireland, UK | 44 (0) 2870 320 639 | | Radioactivity | | | | | | | | | | | |
| | Columbia, MO | (800)538-5227 | | | | | | | | | | | | | |
| Experimur | Chicago, IL | (773)254-2700 | www.experimur.com | PK | X | X | X | X | X | X | X | X | X | X | X |

| Exygen | State College, PA | (800)281-3219 (p) (800)272-1019 (f) | Methods, development | | | | | | | | | | X | X |
|---|--|--|----------------------------|--|---|---|---|---|---|---|---|---|---|----|
| Fraunhofer ITEM | Nikolai-Fuchs Str., D-30625 Hannover, Germany | +49 511 5353 0 (p) +49 511 5353 155 (f) | www.item.fraunhofer.de | X | X | X | X | X | X | X | X | X | X | X |
| Frontier Biosciences | China | US Office— (301)251-0231 | www.frontierbsi.com | X | X | X | X | X | X | X | X | X | X | SP |
| Genelogic | Gaithersburg, MD | (800)436-3564 | www.genelogic.com | Yes | X | X | X | X | X | X | X | X | X | SP |
| Gentronix | Alderley Edge, Cheshire SK10 4TG, UK | +44 (0) 1625 238700 | www.gentronix.co.uk | | | | | | | | | | X | |
| Gibraltar Laboratories Inc. | Fairfield, NJ | (973)227-6882 (p) (973)227-0812 (f) | www.gibraltarlabsinc.com | X | X | | | | | | | X | X | X |
| GLP Compliant Laboratory, Jubilant Biosys Ltd | 96, Industrial Suburb, 2nd Stage, Yeshwantpur, Bangalore 560022, Karnataka | 91-80-66628339 (p) | www.jubilantinnovation.com | Toxicity studies (acute and subacute) Mutagenicity studies (Ames assay, chromosomal aberration test <i>in vitro</i> , micronucleus assay <i>in vivo</i>) Analytical and clinical chemistry testing GLP | | | | | | | | | | |
| Gwathmey | Cambridge, MA | (617)491-0022 (p) (617)492-5545 (f) | www.gwathmey.com | X | X | | | | | | | | X | X |
| Hammer Institute | RTP, NC | (919)558-1200 | www.thehammer.org/ | X | X | X | X | X | X | X | X | X | X | X |
| Haskell Dupont Laboratory | Wilmington, DE | (306)366-5200 | | Yes | X | X | X | X | X | X | X | X | X | X |
| HemoGenix | Colorado Springs, CO | (719)264-6250 | www.hemogenix.com | <i>In vitro</i> tox./PK | | | | | | | | | | |
| Hurley Consulting Associates Ltd | One Main Street, Chatham, NJ 07928 | 1-770-529-1154 (p) 1-770-529-3783 (f) 973-635-9898 (p) 973-635-9881 (f) | www.hurleyconsulting.com | Statistical analysis, study design, preclinical assessments, data management, QA, regulatory affairs, pharmacovigilance, | | | | | | | | | | X |

(Continued)

(Continued)

| Lab Name (Affiliated Labs/ Past Names) | | Location(s) | Phone #(s) | Website | Additional Services | Rat/ Mouse | | Rabbits | Dog | Carcinogenicity | DART | Inhalation | Primate | Pig | IV | | | | Genotoxic | Devices | Metabolism | Analytical | Special Studies |
|--|--|---|--|--|--|---|---|---------|-----|-----------------|------|------------|---------|-----|----|---|---|--|-----------|---------|------------|------------|--------------------|
| Ibex | | Logan, UT | (435)752-4448 | www. ibexpreclinical.com | Cardio telem. | X | X | X | X | X | X | X | X | X | X | | | | | X | X | | |
| ICP Firefly | | Sydney, NSW, Australia | +61-2 9310 3899 | www. icpfirefly.com.au | Yes | X | X | X | X | X | X | X | X | X | X | | | | | X | X | | X |
| IIBAT | | Padappai, India | 04111-274246 | www.iibat.com | ECOTOX | X | X | X | X | X | X | X | X | X | X | | | | | X | | | |
| IITRI | | Chicago, IL McLean, VA | (312)567-4000 (703)918-4480 | www.iitri.org | Yes | X | | X | X | X | X | X | X | X | X | | | | | X | X | | X |
| ILS | | RTP, NC Shardlow, Derbyshire | (919)544-5857 44(0) 1332-793000 | www.ils- limited.co.uk | Data analysis, eco- toxicology | X | | | | X | X | X | X | X | X | | | | | X | | | SP |
| Int. Non- clinical Assessment (INA) | | Nagano- ken, Japan | +81-265-72-6616 (p) +63-49-541-2294 (f) | www.ina-research. co.jp/en/ | Pediatric toxicity | X | X | X | X | X | X | X | X | X | X | | | | | X | | | |
| Innostar Bio-Tech | | Shanghai, China | (+86) 21-50800333 ext 211 | www.innostarsh. com; www. ncdser.com | Biomedicine, gene technology | X | X | X | | X | | | | X | X | X | X | | | X | | | X |
| Indian Institute of Toxi- cology, Pune | | 32 A/1, Hadapsar Industrial Estate, Pune 411 013, Maharash- tra, India | 0091-20-26819962 (p) 0091-20-26819962 (f) | | Toxicity studies (acute studies, oral, parenteral, dermal, inhalation, dermal irritation/ corrosion, eye irritation/corrosion, skin sensitization; subacute studies, chronic studies) | Mutagenicity studies (Ames assay, mammalian erythrocyte micronucleus test, <i>in vivo</i> and <i>in vitro</i> mamma- lian chromosome aberration test) | | | | | | | | | | | | | | | | | |
| Industrial Toxicology Research Center (ITRC) | | Lucknow, India | +91-522-2211547; +91-522 2228227 (f) | www.iitrindia.org | Irritation, behavioral | X | X | X | | X | X | X | X | | | | | | X | | | | X |

(Continued)

[illegible]

| | | | | | | | | | | | |
|--|--|--|---|---|---|---|---|---|---|---|---|
| JSW Research | Austria | www.jswresearch.com | Yes | X | X | X | X | X | X | X | X |
| Korea Institute of Toxicology (KITOX) | Korea | +82-42-610-8105 www.kitox.re.kr/home/main.php | Yes | X | X | X | X | X | X | X | X |
| Krish Biotech Research Pvt Ltd | T-1, QK-17 (Part) WBI-IDC, Kalyani, Phase III, Nadia 741235, West Bengal | 91-33-25824472 (p) 91-33-71081012 (f) www.krishbiotech.com | Physical–chemical testing Toxicity studies (acute) Mutagenicity studies (micronucleus assay <i>in vitro</i> and <i>in vivo</i> , Ames assay <i>in vitro</i> and <i>in vivo</i>) Environmental toxicity studies on aquatic and terrestrial organisms using fish, honey-bees, alga, daphnia, and earthworms Analytical and clinical chemistry testing GLP | | | | | | | | |
| Laboratory Animal Research Services (LARS), Reliance Life Sciences Private Ltd | Dhirubhai Ambani Life Sciences Centre, R282, Thane-Belapur Road, Rabale, Navi Mumbai 400 701, Maharashtra, India | 0091-22-3911 8500-05 (p) 0091-22-3911 18099 (f) www.relbio.com | Toxicity studies (acute studies—oral, dermal, eye irritation, skin irritation, skin sensitization, subacute studies, subchronic studies) Mutagenicity studies (Ames assay) Analytical and clinical chemistry testing GLP | | | | | | | | |
| Liberty Research | Waverly, NY | (607)565-8131 (p) (607)565-7420 (f) | No | X | | | | | | | |
| Litron Laboratories | Rochester, NY | (877)454-8766 | www.litronlabs.com | | | | | | | | |
| Lovelace | Albuquerque, NM | (505)348-9400 | www.lrii.org | | | | | | | | |
| | | | Clinical | | | | | | | | |
| | | | | | | | | | | | |
| | | | | | | | | | | | |
| | | | | | | | | | | | |
| | | | | | | | | | | | |
| | | | | | | | | | | | |
| | | | | | | | | | | | |
| | | | | | | | | | | | |
| | | | | | | | | | | | |
| | | | | | | | | | | | |
| | | | | | | | | | | | |
| | | | | | | | | | | | |
| | | | | | | | | | | | |
| | | | | | | | | | | | |
| | | | | | | | | | | | |
| | | | | | | | | | | | |
| | | | | | | | | | | | |
| | | | | | | | | | | | |
| | | | | | | | | | | | |
| | | | | | | | | | | | |
| | | | | | | | | | | | |
| | | | | | | | | | | | |
| | | | | | | | | | | | |
| | | | | | | | | | | | |
| | | | | | | | | | | | |
| | | | | | | | | | | | |
| | | | | | | | | | | | |
| | | | | | | | | | | | |
| | | | | | | | | | | | |
| | | | | | | | | | | | |
| | | | | | | | | | | | |
| | | | | | | | | | | | |
| | | | | | | | | | | | |
| | | | | | | | | | | | |
| | | | | | | | | | | | |
| | | | | | | | | | | | |
| | | | | | | | | | | | |
| | | | | | | | | | | | |
| | | | | | | | | | | | |
| | | | | | | | | | | | |
| | | | | | | | | | | | |
| | | | | | | | | | | | |
| | | | | | | | | | | | |
| | | | | | | | | | | | |
| | | | | | | | | | | | |
| | | | | | | | | | | | |
| | | | | | | | | | | | |
| | | | | | | | | | | | |
| | | | | | | | | | | | |
| | | | | | | | | | | | |
| | | | | | | | | | | | |
| | | | | | | | | | | | |
| | | | | | | | | | | | |
| | | | | | | | | | | | |
| | | | | | | | | | | | |
| | | | | | | | | | | | |
| | | | | | | | | | | | |
| | | | | | | | | | | | |
| | | | | | | | | | | | |
| | | | | | | | | | | | |
| | | | | | | | | | | | |
| | | | | | | | | | | | |
| | | | | | | | | | | | |
| | | | | | | | | | | | |
| | | | | | | | | | | | |
| | | | | | | | | | | | |
| | | | | | | | | | | | |
| | | | | | | | | | | | |
| | | | | | | | | | | | |
| | | | | | | | | | | | |
| | | | | | | | | | | | |
| | | | | | | | | | | | |
| | | | | | | | | | | | |
| | | | | | | | | | | | |
| | | | | | | | | | | | |
| | | | | | | | | | | | |
| | | | | | | | | | | | |
| | | | | | | | | | | | |
| | | | | | | | | | | | |
| | | | | | | | | | | | |
| | | | | | | | | | | | |
| | | | | | | | | | | | |
| | | | | | | | | | | | |
| | | | | | | | | | | | |
| | | | | | | | | | | | |
| | | | | | | | | | | | |
| | | | | | | | | | | | |
| | | | | | | | | | | | |
| | | | | | | | | | | | |
| | | | | | | | | | | | |
| | | | | | | | | | | | |
| | | | | | | | | | | | |
| | | | | | | | | | | | |
| | | | | | | | | | | | |
| | | | | | | | | | | | |

(Continued)

(Continued)

| Lab Name (Affiliated Labs/ Past Names) | | Location(s) | Phone #(s) | Website | Additional Services | Rat/ Mouse | | Rabbits | Dog | Carcinogenicity | DART | Inhalation | Primate | Pig | Infusion | Genotoxic | Devices | Metabolism | Analytical | Special Studies |
|--|--|---|--|----------------------------|--|---------------|---|---------|-----|-----------------|------|------------|---------|-----|----------|-----------|---------|------------|------------|--------------------|
| IV | | | | | | | | | | | | | | | | | | | | |
| LPT | | Hamburg, GE | 49 40-70-20-20 49 40-70-20-22-99 | www.LPT-pharm-tox.de | | X | X | X | X | X | X | X | X | X | X | X | X | X | X | SP |
| LyChron | | Mountain View, CA | (650)938-3675 | www.lychron.com | Pig device implant studies | | | | | X | | | | X | | | | | | |
| Primetrics | | Singapore | +65 6801 0630 (Phone) | www.primetrics.com | Safety pharm: ortho-pedic models | X | | | | | X | | | | | | | X | | X |
| MB Research | | Spinnerstown, PA | (215)536-4110 | www.mbresearch.com | In vitro and alternative models | X | X | X | X | | | X | | X | | | | X | | |
| Medicilon/MPI | | Shanghai, China | 8621-51-320228 8621-5859-1500 | www.medicilon.com | | X | | X | | | | X | | X | | | | X | | X |
| Micro Test | | Agawam, MA | (800)631-1680 | www.microtestlabs.com | Yes | | | | | | | | | | | | X | | | X |
| Midwest Bio Research | | Evanston, IL Skokie, IL | (866)416-6820 (p) (847)972-2506 (f) | www.midwestbioresearch.com | consultation | | | | | | | | | | | X | | | X | |
| MPI | | Mattawan, MI | (269)668-3336 | www.mpiresearch.com | Safety pharm, trans-genic mouse CA | X | X | X | X | X | X | X | X | X | X | X | X | X | X | SP |
| Radioactivity | | | | | | | | | | | | | | | | | | | | |
| MWRI | | Kansas City, MO | (816)753-7600 | www.mwrf.org/12 | Yes | X | | | | X | | | | | | | | X | | X |
| Nia Life Sciences Inc. | | Libertyville, IL 60048 | (847)573-1852 | www.nialifesciences.com | Pharmacokinetics, GLP and non-GLP, bioanalytical, fungal and tumor model | X | | | | | | | | | | | | | | |
| Yes | | | | | | | | | | | | | | | | | | | | |
| NAMSA | | Toledo, OH Northwood, OH Irvine, CA Kennesaw, GA | (419)666-9455 (866)666-9455 (949)951-3110 (770)427-3101 | www.namsa.com | | | | | | | | | | | | X | X | | | |
| | | | | | | | | | | | | | | | | | | | | |
| NCDSER | | Shanghai, China | | www.ncdser.com | hERG safety pharma | X | X | X | | | | | X | | | | | X | | X |
| Nelson Labs | | Salt Lake City, UT | (800)826-2088 | www.nelsonlabs.com | | X | X | | | | | | | | | X | X | | X | |
| Next Century Inc. | | Newark, DE | (302)453-0571 | www.nextcent.com/ | Yes | X | X | X | X | X | X | X | X | X | X | X | X | | | |
| Northern Biomedical Research | | Muskegon, MI | (231)759-2333 | www.ntba.org | Yes, surgery, large animals | X | X | X | X | | | | X | X | X | | | | X | |

(Continued)

| Lab Name (Affiliated Labs/ Past Names) | Location(s) | Phone #(s) | Website | Additional Services | Rat/ Mouse | Rabbits | Dog | Carcinogenicity | DART | Inhalation | Primate | IV | | | | Special |
|---|---|---|------------------|---|---------------|---------|-----|-----------------|------|------------|---------|----------|-----------|---------|------------|------------|
| | | | | | | | | | | | | Infusion | Genotoxic | Devices | Metabolism | Analytical |
| QPS | Newark, DE | (302)369-5601 | www.qps-usa.com | Yes | X | X | X | X | X | X | X | X | X | X | X | X |
| | Taipei (ROC) | 800-237-1970 | | Radioactivity | | | | | | | | | | | | |
| Rallis Research Center | Bangalore, India | +91-80-28394959 | www.rallis.co.in | yes | X | X | | X | | | | X | | | X | X |
| | Mumbai | +91-80-28394015 | | | | | | | | | | | | | | |
| | | +91-02266652700 | | | | | | | | | | | | | | |
| RCC Labora- tories India Private Ltd | Genome Valley, Turkapally, Shameerpet (Mandal), Ranga | 0091-40-23480422 (p) | ww.recltd.com | Physical-chemical testing (five-batch analysis) | | | | | | | | | | | | |
| | | | | Toxicity studies (acute, subacute, chronic, subchronic, repro- ductive, inhalation, carcinogenicity) | | | | | | | | | | | | |
| | | 0091-40-23480420 (f) | | Mutagenicity studies (Ames assay, chromosome aberration <i>in vitro</i> and <i>in vivo</i> , micro- nucleus assay <i>in vitro</i> and <i>in vivo</i> , mouse lymphoma assay) | | | | | | | | | | | | |
| | Reddy District 500 078, Hyderabad | | | Environmental toxicity studies on aquatic and terres- trial organisms | | | | | | | | | | | | |
| | | | | Analytical and clinical chemistry testing | | | | | | | | | | | | |
| Ricerca | | | | Residue studies (lab analysis only) | | | | | | | | | | | | |
| | Concord, OH | (888)763-4797 | www.ricerca.com | Synthesis, product development | X | X | X | X | X | X | X | X | X | X | X | X |
| | | | | Radioactivity | | | | | | | | | | | | |
| RTC S.p.A. | Via Tito Speri 12, I-40 | +39 06 910 951 (p) +39 06 910 5737 (f) | www rtc.it | Yes, consulting, regulatory | X | X | X | | X | | | X | X | | | WL |
| | RTP, NC | (919)541-6000 | www.rti.org | Yes | X | | | X | X | | | X | | | X | X |

(Continued)

| Lab Name (Affiliated Labs/ Past Names) | Location(s) | Phone #(s) | Website | Additional Services | Rat/ Mouse | | Dog | Rabbits | DART | Carcinogenicity | Primate | IV | | Special |
|---|--|--|----------------------------------|---|---------------|----------|-----|---------|------|-----------------|---------|-----|-----------|---------|
| | | | | | Inhalation | Infusion | | | | | | Pig | Genotoxic | |
| SNBL USA Ltd | 6605 Merrill Creek Park- way, Everett, WA 98203 | (425)407-0121 (p) (425)407-8601 (f) | www.snblusa.com/ | Formulation | X | | | | X | X | X | X | | SP |
| Southern Research | Birmingham, AL | (888)322-1166 (800)967-6774 | www. southernresearch. org | Yes, formulation, development, tumor models | X | | X | X | X | X | X | X | X | SP |
| SoBran | 855 N. Wolfe Street Suite 622 Baltimore, MD 21205 | 301-648-3404 | www. sobranbioscience. com | Regulatory and accreditation assistance, dose range finding, toxicokinetics | X | | | | X | | | | | X |
| SPI-Bio Bertin Pharma | Montigny le Bretonneux, France | +33 139 30 62 60 | www. bertinpharma.com | Clinical | X | | | | | | | | | X |
| Spring Valley Lab. | Woodbine, MD | (800)864-1839 | www.svlab.com | Ferrets | X | | X | X | X | | | | | X |
| SRI | Menlo Park, CA | (650)859-2000 (866)451-5998 | www.srl.com | Formulation, CTM | X | | X | X | X | X | X | X | X | X |
| Stillmeadow Inc. | Sugarland, TX | (281)240-8828 | www. stillmeadow.com | Companion animal studies | X | | X | X | X | | | | | X |
| STS Duotek | Rush, NY | | www.stsduotek.com | Yes, ocular | X | | X | | | | | X | X | X |
| Ethox | Henrietta Road, Rush, NY | (800)836-4850 | | | | | | | | | | | | |
| SYNECOR | RTP, NC (Durham, NC) | (919)541-9977 | www.synecor.com/ | Surgical | | | X | | | | | X | X | Su |
| Syngene International Ltd | Biocon Park, Plot Nos 2 and 3, Bommasandra, IV Phase, Jigani Link Road, Bangalore 560099, Karnataka | 0091-80-28083879 (p) | www. syngeneintl.com | Physical-chemical testing (charac- terization, five- batch analysis, analytical method development and validation, stability and homogeneity testing, dose con- firmation analysis) | | | | | | | | | | X |

Toxicity studies
(acute toxicity,
subchronic toxicity,
chronic toxicity,
local tolerance,
skin irritation,
guinea pig maxi-
mization studies,
reproductive and
development tox-
icity, neurotoxicity
studies)

| | |
|--|---|
| <p>Mutagenicity studies (Ames assay, micronu- cleus test <i>in vivo</i>, chromosome aberration <i>in vitro</i>)</p> | <p>Bioanalytical (TK analysis) for preclinical studies</p> |
|--|---|

PK, *in vivo*
pharmacology,
ADME/T assays

GLP

(Continued)

(Continued)

| Lab Name (Affiliated Labs/ Past Names) | | | Location(s) | Phone #(s) | Website | Additional Services | Rat/ Mouse | Dog | Rabbits | Inhalation | Primate | Pig | IV | Genotoxicity | Devices | Metabolism | Analytical | Special Studies |
|--|---|--|--------------------------|---|---------|---------------------|---------------|-----|---------|------------|---------|-----|----|--------------|---------|------------|------------|--------------------|
| UIC Tox Research Lab | Chicago, IL | (312)996-9185 | www.uic.edu/ labs/tox | Yes | | Radioactivity | X | X | X | X | X | X | X | X | X | X | X | X |
| Valley Biosystems | West Sacra- mento, CA | (916)374-2735 | | | | | | | | | X | | | | | | | |
| Vanta Bio- sciences | Belgium Czechoslovakia | +32-2-709-2920 | www.vantabio.com | Biotechnology, agri- food, nutritional | X | | X | X | X | | X | | | X | X | X | X | X |
| Vanta Bioscience | K2, 11th Cross Street, SPI- COT Indus- tries Estate, Gummid- pundi 601201, Tamil Nadu | 91-44-67910-300, 329, 306, 303 (p) 91-44-4220-2810 (f) | www.vantabio.com | Toxicity studies (acute, subacute, chronic, reproductive, <i>in vitro</i> skin corrosion and irritation using episkin model) | | | | | | | | | | | | | | |
| | | | | Mutagenicity studies (Ames assay, Micronucleus test, chromosome aberration test <i>in vitro</i> and <i>in vivo</i>) | | | | | | | | | | | | | | |
| | | | | Analytical and clinical chemistry testing | | | | | | | | | | | | | | |
| | | | | GLP | | | | | | | | | | | | | | |
| Vimta Labs Ltd | Life Sciences Facility, Plot No. 5, Alexandria Knowledge Park, Genome Valley, Turkapally, Shameerpet 500 078, Hyderabad | 91-40-67404040 (p) 91-40-39847708 (f) | www.vimta.com | Clinical Toxicity studies (acut- , subacute, chronic, reproductive, carci- nogenicity) | X | | X | X | X | X | X | | | X | X | X | X | X |
| | | | | Mutagenicity studies (Ames assay, chromosome aberration test <i>in vivo</i> and <i>in vitro</i> , mouse lymphoma assay, micronucleus test <i>in vitro</i> and <i>in vivo</i>) | | | | | | | | | | | | | | |

| | Analytical and clinical chemistry testing | | | | | | | | | | | | | |
|---|--|-------------------------------|----------------------|---|---|---|---|---|---|---|---|---|---|---|
| | Bioanalytical, pharmacokinetics, safety pharmacology | | | | | | | | | | | | | |
| | GLP | | | | | | | | | | | | | |
| | Toxicity studies (acute, subacute, reproductive) | | | | | | | | | | | | | |
| Vivo Bio-tech Ltd | Mutagenicity studies (Ames assay, micronucleus assay <i>in vivo</i>) | | | | | | | | | | | | | |
| | Analytical and clinical chemistry testing | | | | | | | | | | | | | |
| | GLP | | | | | | | | | | | | | |
| | Immunodeficient mice | | | | | | | | | | | | | |
| vivoPharm Pty Ltd | Sy #: 349/A, 350/A, 351, 356/3A, Pregnapur Village 502311, Gajwel Mandal, Medak District, Telangana | 91-845-4210411 (p) | www.vivobio.com | X | X | X | X | X | X | X | X | X | X | X |
| | | | | | | | | | | | | | | |
| | | | | | | | | | | | | | | |
| RDDT, a vivoPharm Company | South Australia, West Campus, Plenty Rd., Building 201, Level 2, Suite 11, Bundoora, VIC 3083, Australia | 61-3-9988-1800 (p) | www.vivopharm.com.au | X | X | X | X | X | X | X | X | X | X | X |
| | | | | | | | | | | | | | | |
| | | | | | | | | | | | | | | |
| WestChina-Frontier PharmaTech | Suite 29, 240 Plenty Road, Bundoora, VIC 3083, Australia | 61-3-9988-1800 (p) | www.vivopharm.com.au | X | X | X | X | X | X | X | X | X | X | X |
| | | | | | | | | | | | | | | |
| | | | | | | | | | | | | | | |
| WIL Research Labs Inc. (Acquired by Charles River Labs) | 28 Gaopeng Avenue, Chengdu, China | +86 (28) 85154334 ext 502 (p) | www.glpd.com | X | X | X | X | X | X | X | X | X | X | X |
| | | | | | | | | | | | | | | |
| | | | | | | | | | | | | | | |
| WIL Research Labs Inc. (Acquired by Charles River Labs) | 1407 George Rd., Ashland, OH 44805 | (419)289-8700 (p) | www.wilresearch.com | X | X | X | X | X | X | X | X | X | X | X |
| | | | | | | | | | | | | | | |
| | | | | | | | | | | | | | | |
| WuXi AppTec | St. Paul, MN | (888)794-0077 | www.apptec-usa.com | X | X | X | X | X | X | X | X | X | X | X |
| | | | | | | | | | | | | | | |
| | | | | | | | | | | | | | | |

(Continued)

(Continued)

| Lab Name (Affiliated Labs/ Past Names) | Location(s) | Phone #(s) | Website | Additional Services | Rat/ Mouse | Rabbits | Dog | Carcinogenicity | DART | Inhalation | Primate | Pig | Infusion | IV | Genotoxic | Devices | Metabolism | Analytical | Special Studies |
|---|---|-------------------|---|---------------------------------------|---------------|---------|-----|-----------------|------|------------|---------|-----|----------|----|-----------|---------|------------|------------|--------------------|
| WuXi Pharmatech | 288 FuTe ZhongLu, WaiGaoQiao, Shanghai 200131, PR, China | +82(21)-5046-1111 | www. pharmatechs.com www. wuxiapptec.com | Rodent toxicity (oral, IV, and PK) | X | X | X | X | | | X | | X | X | X | X | X | X | X |
| XenoBiotics | Plainsboro, NJ | (609)799-2295 | www.xbl.com | Formulation | X | X | X | X | | | X | | | | | | X | | X |
| Xenotech | Lenexa, KS | (913)438-7450 | www. xenotechllc.com | Radioactivity Yes | X | X | X | X | | | X | | X | | | | X | | |

SP in the special studies column indicates safety pharmacology; WL in this column indicates wildlife/environment testing is available; EN in this column indicates that environmental testing is available.

* A "yes" or "+" listed in the additional services column indicates that there are more services available than just those listed.

** Limited.

^a Hamster carcinogenicity is available.

INDEX

Note: Italicized letters *f*, and *t* following page numbers indicate figures and tables, respectively.

- 3 Rs (replacement, reduction, refinement)
 - see* humane animal use in research
- 15 day report, 639
- 505(b) (New Drug Application)
 - (1) (traditional), 41
 - (2) (small molecules), 41, 269, 711
- 510(k) (premarket notification for a change in a device), 50, 714
- absorption, 81–111, 384–389, 589
 - inhalation, 385, 464
 - lymphatic channels, 88
 - oral, 86, 384
 - parenteral, 84
 - pulmonary system, 386
 - skin, 84, 387, 474
 - weak acids and bases, 89
- absorption, distribution, metabolism, excretion (ADME), 108*t*, 110, 166*t*, 213, 312*f*, 382, 395, 397, 406, 437, 509, 526*t*, 539, 543*f*, 556*t*
- active cutaneous anaphylaxis assay (ACA), 251, 259
- active systemic anaphylaxis assay (ASA), 259
- active transport, 83, 88, 385, 397
- “activity criterion,” 83, 88, 385, 397
- adaptive immunity, 228
- adaptive repair pathway, 172
- adduct formation, 173, 324
- adulteration, 107
- adverse drug event (ADE), 641–647
- adverse drug reaction (ADR), 593, 622–626, 629, 641–647
- adverse effects, 115, 226, 295–304, 526*t*
- adverse event (AE), 242, 452, 508, 586, 592, 643*f*, 646, 740
- age adjustment, 671
- agonal signs, 140
- alanine aminotransferase (ALT), 124*t*, 277*t*, 282, 355*t*, 612*t*, 787*t*
- albumin-to-globulin ratio (A/G), 248, 355
- alkaline phosphatase (ALP), 121*t*, 122*t*, 124*t*, 355*t*, 473, 612*t*, 613
- allergenicity, 235, 259–261, 440, 555*f*
- alternation, 295, 422
- Ames, Bruce, 181
 - mutagenicity data, 214, 220*t*
 - salmonella/plate incorporation method, 177, 181–185
 - test, 9*f*, 183
- analog plot, 75–76
- analysis
 - covariance (*see* Analysis of Covariance (ANCOVA))
 - screening data, 73, 154, 545
 - tumor incidence, 330–342, 654, 667, 673
 - variance (*see* Analysis Of Variance (ANOVA))
- Analysis Of Covariance (ANCOVA), 285–287, 677–705
- Analysis Of Variance (ANOVA), 284–287, 653–705
- anaphylactic reactions (anaphylaxis), 85, 227, 235, 236*t*, 245*t*, 254, 259, 442–446
- aneuploidy, 176–177, 196, 326
- animal husbandry, 193, 330, 374, 375
- animal models, 263, 292*t*, 382, 385, 409*t*, 455, 515, 732–738 *see also names of specific animals*
- Animal Welfare Act (AWA), 161, 270, 519
- antibiotics, 15, 23, 233–238, 292*t*, 715
- antibody-mediated hypersensitivity *see* antibody-mediated; hypersensitivity, type II (cytotoxic)
- antigenicity, 226, 242, 415*t*, 434*t*, 435*t*, 437*t*, 438*t*, 451–455, 542
- antimetabolites, 232, 233*t*, 259, 292*t*
- Anxiety Status Inventory (ASI), 616
- Area Under the Curve (AUC), 399*t*, 405*t*, 493, 587*t*
- Armstrong method, 258
- Arthus reaction *see* hypersensitivity
- aspartate aminotransferase (AST), 124*t*, 277*t*, 282, 355*t*, 612*t*, 787*t*, 788*t*, 814*t*

- assays *see also names of specific assays*
 chromosomal and genomic mutations, 178*t*
 cytotoxicity, 187–190
 dominant lethal, 178*t*
 gene mutations, 49, 178*t*, 542, 574*t*, 842*t*
 host resistance, 245*t*
in vitro cytogenetic, 49, 178*t*, 193–196, 556*t*
 isolated tissue pharmacological, 420*t*
 local lymph node assay (LLNA), 255–258, 530*t*, 567
 neurochemical, 421
 Asthma, 235, 444*t*, 465, 651
 inhaler, 2*t*
 autoimmunity, 225–248, 264
 Avogadro's law, 470
- bacterial mutation tests, 180
 Bartlett's Test of Homogeneity of Variance, 666*f*, 673–675, 686, 705
 base substitution, 173, 174, 182, 185
 Baycol, 630*t*, 635*t*, 649, 650
 Bayesian inference *see* Bayes' Theorem
 Bayes' Theorem, 702–705
 B-Cell lymphoproliferation response, 249
 B cell /T cell ratio, 227–250
 B6C3F1 mice, 253, 328–334
 beagle, 35, 147, 270–273
 Beck Depression Inventory, 613*t*, 616
 behavioral
 rating scales, 613*t*, 619
 tests, 303, 618, 735
 Bender-Gestalt test, 613*t*, 620
 benign neoplasia, 334–335
 Bernard, Claude, 4
 b-estradiol *see* estradiol
 bias, 285, 298, 333, 341, 622, 655
 bidirectional drug effects, 413
 bile, 92, 124*t*, 392–395, 398–400, 612*t*
 bilirubin, 121*t*, 122*t*, 124*t*, 272, 277*t*, 355*t*, 371, 588*t*, 612*t*
 bioavailability and thresholds, 90–91
 biochemical assays, 421
 bioengineered products, 455
 classification, 435*t*
 biological significance *see* significance
 Biologic License Application (BLA), 7*t*, 19, 39*t*, 594*f*
 biologics, 14*t*, 24, 305*f*, 408, 437*t*, 438*t*, 458, 517, 594*f*
 Biologics Control Act *see* Virus Act of 1902
 biomarkers in heart failure, 615*t*
 biotechnology products, 30, 43*t*, 253, 264, 433–459, safety evaluation program 454
 bioprocess technology, 446
 biosimilars, 458*t*
 gene therapy products, 446
- immunogenicity/allergenicity, 440
 monoclonal antibody technology, 441
 preclinical safety assessment, 437
 recombinant DNA technology, 439
 regulation, 436
 vaccines, 449
 biotransformation, 47, 98, 218, 383*t*, 389–408
 BLA *see* Biologic License Application (BLA)
 b-lactam antibiotics, 233–240
 blockbuster biotechnology approvals, 434*t*
 blocking, 661
 blood
 collection, 270, 275, 397
 compatibility, 82, 104, 367, 486*t*, 517*t*
 red cells, 124, 239, 249, 568, 612*t*
 Urea Nitrogen (*see* Blood Urea Nitrogen (BUN))
 Blood Urea Nitrogen (BUN), 121*t*, 122*t*, 520, 612*t*
 body and organ weights, 705
 body water, by species, 389*t*
 body weight, 37, 105, 116–126, 144, 145*t*, 271–276, 299, 309*f*, 705
 bolus, 164*t*, 272
 bolus vs. infusion, 85, 272, 367
 botanical drug products, 38
 Bovine Corneal Opacity (BCOP), 364, 564, 574*t*
 bovine serum albumin (BSA), 253
 Buehler Assay (BA) *see* Buehler test
 Buehler test (Buehler assay (BA)), 244*f*, 254–259
 Bureau of Biologics (now CBER), 30, 436*t*
- Canadian Centre for Occupational Health and Safety (CCINFO), 64
 capsular polysaccharide vaccine, 449
 captopril, 235*t*, 239, 240
 carcinogenesis, 174, 322, 707
 mechanisms and theories of chemical, 327
 carcinogenicity studies, 321–345
 carcinogens
 epigenetic, 325
 genotoxic, 174, 322
 human, 327
 mechanisms and classes, 322
 cardiotoxicity, 216, 279, 467
 cardiovascular system
 measurements, 282
 safety pharmacology evaluations, 414*t*
 toxicity, 279
 Carroll Rating Scale for Depression, 617
 case I, II, III, IV *see* significance
 Case Report Forms (CRF's), 597
 categorical and ranked data, 679
- causality
 adverse drug reactions, 622
 assessment, 622, 647
 CDER flowchart *see* specific immunotoxicity testing
 cell-mediated immunity, 228*t*, 250
 cellular
 components of immune system, 227–260
 therapies, 33–35
 tissue implants, 711
 censoring, 144, 653, 661, 662, 684
 Center for Biologics Evaluation and Research (CBER), FDA, 19–55, 436, 446, 714–715, 738 *see also* Bureau of Biologics
 Center for Devices and Radiological Health (CDRH), 50–54
 Centers for Disease Control and Prevention (CDC), 640
 central nervous system, 415, 419
 central tendency plots, 74
 chelation, 83*t*, 397*t*
 Chernoff's faces, 697
 Children's Behavior Inventory (CBI), 613*t*, 619
 Children's Diagnostic Scale (CDS), 613*t*, 619
 Chinese Hamster Ovary (CHO), 185, 194
 Chinese hamster V79/HGPRT (hypoxanthine-guanine phosphoribosyltransferase) assay, 178, 185, 187, 194
 Chi Square, 666*f*
 chromosomal aberrations (CAs), 170
 chromosome set damage, 170, 177
 chronic studies, 111, 144, 159
 chronic toxicity, 29*t*, 36, 44, 75, 108*t*, 110*t*, 400, 434*y*, 541*t*, 706
 Cisapride (Propulsid), 630*t*
 Class B sources, 280
 classifying clinical studies according to objective, 599*t*
 clearance, 383*t*, 402–410
 hepatic, 383*t*
 inhaled aerosols, 464
 renal, 284*t*
 total, 284*t*
 clinical
 chemistry, 117–126, 242*t*, 248, 355, 612*t*, 613, 706
 observations, 116, 117*t*, 139–146, 272, 282
 pathology, 117, 121*t*, 243*t*, 299, 355, 520
 signs, 96, 117, 118*t*, 139*t*, 140, 155*t*
 Clinical Trial Application (CTA), 10
 Clinical Trial Certificate Exemption (CTX), 22
 clinical trial safety indicators, 609
 Clinical Trials Certificate (CTC), 22

- cloned human potassium channels, 419
- clumping techniques, 699
- cluster analysis, 699
- Cochran t-test, 666*f*, 685
- combination products, 49, 711
- comet assay, 178, 190–192, 556*t*, 724*t*
- Committee for Proprietary Medicinal Products (CPMP), 241*t*, 328, 414*t*, 418, 637
- common data transformations, 678*t*
- comparative metabolic “fingerprinting,” 382
- compassionate use, 35
- complete acute toxicity testing, 142
- computed tomography (CT), 399, 738
- concurrent control, 68, 115, 147, 153, 342, 659
- congressional committees responsible for FDA oversight, 40*t*
- CONSAM data modelling software, 403
- contaminated diphtheria toxins, 13, 436*t*
- contrast agents, 484, 509
- control charts, 73
- convective absorption, 88
- correlation coefficient, 693
- Council for International Organizations of Medical Sciences (CIOMS), 593, 639, 646
- Cox-Stuart test, 670
- Cox-Tarone binary regression, 341
- creatinine phosphokinase (CPK), 121*t*, 122*t*, 371, 612, 613, 706
- Crestor, 2*t*, 650
- curve-stripping, 403
- cyclooxygenase 2 (COX-2), 353, 540*t*, 635*t*
- cyclosporine, 232
- cytochrome P-450 (CYPs) isoenzymes, 9, 276, 314*t*, 382, 389–394
- cytogenetics, 176–193
- cytokines, 248, 736, 740
- cytokine storm, 225, 442
- cytotoxic
 - hypersensitivity (*see* antibody-mediated); hypersensitivity, type II (cytotoxic)
 - T_k (T killer) cells, 228, 230*t*, 251
- cytotoxicity
 - assay, 187–190, 196
 - confounding factor, 191
- cytotoxic T Lymphocyte-mediated assay (CTL), 250
- Data and Safety Monitoring Boards (DSMBs), 596
- data recording, 196, 664
- dechallenge, 624, 642
- decision tree, 151, 212, 311*f*, 665*f*, 675*f*, 706
 - hypothesis testing procedures, 666*f*
 - modeling procedures, 666*f*
 - reduction dimensionality procedures, 667*f*
- Declaration of Helsinki, 591, 598
- degradation products, 177, 724
- delayed contact hypersensitivity, 216, 226, 263, 567
- delayed liver toxicity (DLT), 216, 570, 586
- Delayed-Type Hypersensitivity (DTH) *see* hypersensitivity, type IV (Delayed-Type Hypersensitivity (DTH))
- deposition of inhaled aerosols, 463
- dermal
 - formulations, 100
 - irritation test, 566*t*
 - route, 82–83, 461, 474
 - toxicity, 108–110, 278
- determination of pregnancy, 301
- developmental and reproductive toxicity (DART), 28, 44, 291–316, 506
- Developmental And Reproductive Toxicology (DART), 28, 842–860
- diagnostic radiopharmaceuticals, 484, 487, 516
- diethylene glycol, 11, 16, 96, 604
- diethyl ether (“sweet oil of vitriol”), 4, 5, 469
- diethylstilbestrol (DES), 233*t*, 234, 326
- Dimethylformamide (DMF), 102, 104
- dimethylsulfoxide (DMSO), 102, 104, 182, 758*t*
- distribution-free multiple comparisons, 682
- DNA adducts, 174, 190*t*, 322–323
- DNA/oligonucleotide hybridization, 450
- dog, 37*t* 95*t*, 147–148, 161*t*, 164*t*, 165*t*, 260, 270–273, 280*t*
 - beagle, 147, 273*t*
 - hemodynamics, ECG and respiration studies, 417
 - 6-lead ECG measurements, 417
 - telemetry studies, 417
- dominant lethal assay, 178*t*
- dose-range finders (DRFs), 154, 166, 438*t*, 475*t*, 492*t*, 548
- dose ranging and selection, 184
- dose-response relationship, 68, 76*f*, 139*t*, 161, 261, 301, 330, 353*t*, 415, 468, 597, 660, 667*t*, 691
- dose selection, 131, 330
- dosing calculations, 105
- D-penicillamine, 235*t*, 237–240
- Draize test
 - eye irritation, 362, 531, 564
 - skin irritation, 360, 374
- drug
 - allergies, 226
 - botanical products, 41
 - distribution, 505, 591
 - formulations, 604
 - immunostimulation, 235–243
 - metabolites, studies of, 601
 - new, 41, 60*f*, 242, 458, 461, 538, 543*f*
 - safety withdrawals, 637–638
 - singularity, 624
 - supplies, 596
 - withdrawn, 630*t*
- Drug Master File (DMF), 23, 107–109
- drug metabolism and pharmacokinetics (DMPK), interspecies differences, 394
- Dry Powder Inhalers (DPIs), 461, 465, 472, 728
- DTH *see* hypersensitivity, type IV (Delayed-Type Hypersensitivity (DTH))
- Duncan’s Multiple-Range Test, 666*f*, 687
- Dunnett’s t Test, 666*f*, 688
- earthworms, 557, 562–563
- EC seventh amendment, 525*t*
- Ehrlich, Paul, 4, 79
- electrocardiograms (ECGs), 271*t*, 272, 274, 282, 315*t*, 417
- electroencephalogram (EEG), 421, 610
- elixir of sulfanilamide, 16, 19, 352, 604
- embryo-fetal development, 44, 47*t*, 293, 295–308, 498
- endoplasmic reticulum, 179, 351, 466
- environmental impact, 23, 563
- enzyme induction, 351, 383*t*, 605
- Enzyme-Linked Immunosorbent Assay (ELISA), 248, 253, 407, 454, 741
- epigenetic carcinogenesis, 325–327
- Escherichia coli* (*E. coli*), 172–182, 450
 - HS-4, 440
 - polypeptides, 453
 - proteins, 453
 - tester strains, 183
- estradiol, 233–234
- estrogens, 234, 326, 344*t*
- ethical tenets, IRBs, 602
- European Economic Community (EEC), 526, 542, 563
- European Medicines Agency (EMA), 124, 240, 312, 382, 413, 437, 446, 728, 731, 739
- European Union (EU), 41, 47, 241, 637, 646
- excipients, 79, 96, 97*t*, 106–110, 367, 725
 - drug formulations, 604
 - regulation, 168–170
- excision repair, 171, 182, 190, 321
- excretion, 175*f*, 191, 383*t*, 394, 398, 509, 741*t*
- expired air, 395, 398
- Exploratory Data Analysis (EDA), 67, 284, 544, 673, 678
- expression cloning, 451
- extractables, 711, 721, 727–728
- extrathoracic, 461, 463

- faces *see* Chernoff's faces
- facilitated diffusion, 88
- "false negative" *see* biological significance; statistical
- "false positive" *see* biological significance; statistical
- federal drug law, important dates, 14*t*
- Federal Virus-Serum-Toxin Act (TOPN), 436*t*
- feed consumption, 139*t*, 143*f*, 144, 146*f*, 153*f*, 271*t*, 272, 274, 282
- female reproductive studies, 49
- ferret, 273–275
- fertility, male and female, 294 *see also* reproductive studies; sterility
- fetal
- examinations, 301–305
 - risk summary, 310
- fiducial limit, 143*t*, 655*t*, 656
- First-In-Man (FIM), 8, 10, 124, 442, 599
- Fischer 344 (F344), 253, 328–338, 343*t*
- Fisher's exact test, 666*f*, 679
- fixed-dose procedure, 135, 136*t*
- flip-flop pharmacokinetics, 403
- fluidized-bed dust generator, 472
- Food and Drug Administration
- Modernization Act (FDAMA), 18, 19, 23, 37
- Food Drug and Cosmetic (FD&C) Act, 106, 436*t*, 602
- Kefauver-Harris amendment, 17, 636
- Form 3500/3500A *See* U.S. Food and Drug Administration (FDA)
- formulation, 79–111, 371, 473, 604
- dermal, 100
 - development, 97
 - excipients, 604
 - interactions, 102
 - mucosal damage, 473
 - oral, 103
 - parenteral, 104
 - test materials, 96
- forward mutation tests, 180, 185
- Fourier (time) analysis, 667*f*, 695, 699
- frameshift mutations, 173, 181, 323
- functional observational battery (FOB), 74*f*, 117*t*, 118*t*, 146, 155*t*, 270, 419
- functional reserve capacity, 264
- Gantt chart, 542–545
- gastric emptying rate, 386*t*, 400, 415, 427
- gastric pH changes, 89, 427
- Gelsinger, Jesse, 446
- gene
- regulation, 171
 - therapies, 34–35, 446
 - therapy products, 447
- general case oral drug, 9*f*
- general criteria affecting drug/device determination, 54
- Generally Recognized As Effective (GRAE), 52, 717
- Generally Regarded As Safe (GRAS), 52, 109, 483, 717
- genetoxic vs. nongenotoxic mechanisms of carcinogenesis, 174
- genotoxicity, 32, 169, 198, 215, 486*t*, 493*t*, 574*t*
- tests recommended by ICH, 516
- Geometric Standard Deviation (GSD), 463–472
- geriatric claims, 37
- germ theory, Louis Pasteur, 4
- Gilbert's syndrome, 418, 588*t*
- glucocorticosteroids, 232
- glutamate dehydrogenase (GDH), 124*t*
- Good Clinical Practices (GCP), 592
- Good Laboratory Practices (GLP), 161, 519
- Good Manufacturing Practices (GMP), 24, 31
- goodness-of-fit tests, 676
- granuloma formation, 238
- gravid uterine weights, 299–300
- guinea pig, 374
- Guinea Pig Maximization Test (GPMT), 244*f*, 244*t*, 254–256, 256*f*
- Haber's rule, 468
- Hamilton Anxiety (HAMA) scale, 613*t*
- HAMilton Depression (HAMD) scale, 613*t*
- Harber and Shalita Method, 258
- hazard assessment, 67, 531
- heavy metals, 60, 234
- hematocrit, 121*t*, 612*t*, 613, 706
- hematology, 1, 355*t*, 520
- hemodynamics, 315*t*, 414*t*, 417
- hemolytic potential, 378
- hepatocyte, 382, 391*t*, 572*t*
- hierarchical techniques, 699
- histograms, 75, 674
- histopathological lesion incidences, 706
- histopathology, 125, 351, 520
- tissues, 352*t*
- historical controls, 154, 165, 304, 342, 673, data. 72*f*
- homologue, 409
- homology, 439, 453
- host resistance assays, 245*f*
- human carcinogens *see* carcinogens, human
- humane animal use in research, 544
- Human Equivalent Doses (HED), 446
- human Ether- α -go-go-Related Genes (hERG), 70, 216, 419
- humanizing mice, 457
- human peripheral blood lymphocytes, 194, 567
- Human Repeat Insult Path Test (HRIPT), 258
- humoral
- immune response, 246
 - immunity, 228*t*, 229*t*, 231
- Hydra* system, 316
- hydrolysis, 171, 324, 390*t*
- hydroxybutyric dehydrogenase (HBDH), 122*t*, 706
- hyperpharmacology, 159, 225, 433, 550
- hyperplasia/hyperplastic, 246*t*, 326, 667
- hypersensitivity
- type I (immediate; reaginic), 235, 253
 - type II (cytotoxic; antibody-mediated), 234, 235, 237, 240, 242*t*, 253, 259
 - type III (Arthus reaction), 236*t*, 237, 242*t*, 259
 - type IV (Delayed-Type Hypersensitivity (DTH)), 235*t*, 237, 238, 250, 254, 258
- hypoxanthine-guanine phosphoribosyltransferase (HGPRT), 178
- identification threshold, 215, 722–725
- imaging agents, 483
- immune system, 125, 225, 245
- immunoassay methods, 253, 397, 407
- immunogenicity, 259
- biotech drugs, 261
- immunomodulatory drugs, 263, 417*t*, 586
- immunopathological assessments, 227, 246
- immunostimulation, 235–243
- immunosuppression, 231, 242
- immunotoxicology, 126*t*, 225, 260
- implant counts, 301
- impurities, 215, 721
- inactivated vaccines, 449
- inactive metabolites, 561*t*
- inducing agents, 179
- informed consent, 603
- inhalation, 29*t*, 82, 108, 110, 385, 461
- inhalation exposure techniques, 470
- in-life phase, 549
- innate immunity, 227
- insertional mutagenesis, 448
- Institutional Review Board (IRB), 602
- insulin pump, 711
- interactions between skin, vehicle, and test chemical, 102
- Interagency Coordinating Committee on the Validation of Alternative Methods (ICCVAM), 135, 575
- intercenter consultation, 51
- interferons (IFNs), 230*t*, 264, 408, 433
- interleukins, 264, 433
- International Classification of Diseases
- 9th Revision, Clinical Modification (ICD-9 CM), 645
 - 10th Revision (ICD-10), 645

- International Conference on Harmonisation (ICH) *see also* S1.2, 3, 4, 5, 6, 7, 8, 9, 10, 11
- International Federation of Pharmaceutical Manufacturers Association (IFPMA), 44, 640
- International Pharmaceutical Excipients Council (IPEC), 79, 109
- interspecies differences, 166, 394, 473, 532
drug metabolism and pharmacokinetics (DMPK), 394
- intramuscular route, 86, 272, 281*t*, 367
- intraperitoneal route (IP), 81, 86, 447, 515
- intravenous route (IV), 84
- Investigational Drug Exemption (IDE), 33, 51–54, 716
- Investigational New Drug (IND), 23*t*, 242
enabling, 17, 35–36 (*See also* First-In-Man (FIM))
- in vitro* cytogenetic assays, 193
- Irwin screen, 419
- isolated tissue pharmacological assays, 420*t*
- Japan, 1, 47, 424*t*, 542, 646
- Johns Hopkins, 591
- Kaplan–Meier, 341, 684, 700
- Karnofsky’s Law, 308
- Kefauver–Harris amendment *see* Food Drug and Cosmetic (FD&C) Act
- Kelsey, Francis, 17
- Kendall’s coefficient of rank correlation, 666*f*, 694
- kidneys, 122*t*, 125*t*, 155, 237, 254, 352*t*
slices, 302
- “kill the losers as early as possible,” 8
- “known to be human carcinogens” *see* carcinogens, human
- Kruskal–Wallis nonparametric tests, 183
- lactate dehydrogenase (LHD), 124*t*
- latent period, 174, 278, 520, 623
- leachables, 721–728
- Lethal Dose 50% (LD₅₀), 47, 88, 96, 130–156, 214–219, 560
- lethality, 116*t*, 214, 560
- lethality testing, 130
- levels of models, safety assessment, 557*t*, 794
- life table analysis, 667*f*
- life tables, 339, 683, 700
- limit doses for toxicological studies, 110*t*
- limit tests, 134
- Limulus* Amebocyte Lysate (LAL) test, 377, 573
- linear regression, 183, 653, 666*f*, 691
- lithium, 235, 240
- litter size, effect of, 296, 305–306, 398
- liver toxicity, 124*t*, 153, 467, 570
- local effects, 81, 463
- Local Lymph Node Assay (LLNA) *see* assay, LLNA (local lymph node assay)
- longitudinal analysis, 284
- Lotronex, 540*t*, 631*t*
- lupus, 229*t*, 232, 240, 422*t*, 423*t*
- macrophage function, 231, 251
- “magic bullets,” 5, 6, 80*f*
- Magnusson and Kligman, 255
- Major Histocompatibility Complex (MHC), 228, 231, 237
- malformation, 175, 295, 304–309
- Mann–Whitney U test, 284, 682
- Marketing Authorization Application (MAA), 23
- Mass Median Aerodynamic Diameter (MMAD), 463
- Material Safety Data Sheets (MSDS), 63
- maximum likelihood, 131, 288, 342, 676
- Maximum Nonlethal Dosage (MNLD), 137
- Maximum Recommended Human Dose (MRHD), 331
- Maximum Tolerated Dose (MTD), 331, 475*t*
- measurement of drug activity, early, 600
- mechanisms and theories of chemical carcinogenesis, 327
- median, 635*t*, 638*t*
- Medical Dictionary for Regulatory Activities (MedDRA), 645
- MEDLINE, 64
- MedWatch, 452, 636, 639–642
- Merck, 3, 5
- Merck Index, 62
- meta-analysis, 662, 701
- metabolic activation, 110*t*, 178, 198, 323, 392
- metabolism cages, 272, 275, 398, 401
- metabolites, studies, 416
- metal catalysts and metal reagents, class exposure and concentration limits, 729*t*
- metered-dose inhaler (MDI), 472, 721
- micronuclei, 170*t*, 197, 574*f*
- micronucleus test, 108*t*, 110*t*, 197, 574*t*, 724*t*
- minimal acute toxicity test, 139
- Minimal Lethal Dosage (MLD), 134, 137
- Minimum Active Biological Effective Level (MABEL), 446
- minipig, 275, 417
- missing data, 672
- missing values, 288
- mitochondrion, 115, 351, 389
- mitochondrial mixed functional oxidase (MMFO), 277
- Mixed-Lymphocyte Response (MLR) assay, 250
- modeling, 210, 404, 690
- monoclonal antibodies (mAbs), 228*t*, 408*t*, 442
- Mouse Lymphoma Assay (MLA), L5178YTK^{+/−}, 178, 187, 192
- mouse specific locus test, 193
- mucociliary transport system, 385, 466
- MultiDimensional Scaling (MDS), 695, 697
- multiple comparisons, 183, 285, 656, 682, 688
- multivariate data, 75, 545, 696–699
- multivariate techniques, 285
- muromonab-CD3; Orthoclone OKT3 (OKT3), 442, 444*t*
- muscle irritancy, 371, 566
- mutations, 172, 173
- myelosuppression, 232–234
- nasal administration, 91*f*, 94, 282
- Natalizumab *see* Tysabri
- National Library of Medicine (NLM), 62
- natural killer (NK) cells, 230*t*, 231, 233*t*, 235*t*, 251, 313*t*
- natural products, 5–7, 24*t*, 715*t*
- negative accuracy, 68, 546
- neoplasia, events leading to, 175*f*
- neurochemical assays, 421
- neuromuscular screen, 139*t*, 146, 147*t*, 546
- neurotoxicology, 125, 126*t*, 421
- New Chemical Entities (NCE), 129, 192, 406, 525, 593
- New Drug Application (NDA), 14*t*, 17, 22–25, 29*t*, 36, 41, 242, 486*t*, 517*t*
- new drug, definition, 22
- nitrogen mustards, 234, 323
- N-nitrosodimethylamine (NDMA), 672
- NonHuman Primate (NHP), 147, 269, 279
- NONLIN, 403
- nonlinear regression, 222, 666*f*, 692
- nonmetric scaling, 697
- Nonsteroidal Anti-Inflammatory Drugs (NSAIDs), 417*t*, 635*t*, 638*t*
- No-Observable-Effect Level (NOEL) estimation, 310, 669, 726–727
- normalize the data, 677
- No-Statistical-Significance-Of-Trend (NOSTASOT) dose, 671
- Nuremberg Code, 591
- occupational exposure limits (OELS), 531
- Occupational Safety and Health Administration (OSHA), 63, 215, 525–532
- occupational toxicology, 523
- Occupational Toxicology Roundtable, 531
- ocular irritation testing, 219, 362, 564

- off-target effects, 433
- oncology
 - cytotoxic, 492, 606
 - drugs, 8, 489
 - imaging, 510
 - protein-targeted molecules, 492*t*
- one-tailed comparisons, 338
- optical isomers, 36
- oral
 - absorption (k_a), 400
 - contraceptives, 35, 422*t*
 - elimination (E), 403
 - formulations, 103
 - route, 86
- Organisation for Economic Co-operation and Development (OECD), 30*t*, 178*t*, 560*t*, 574*t*
- organ weights, 117*t*, 124, 155*t*, 166*t*, 241*t*, 246*t*, 354, 389, 520
- Orphan Drugs, 14*t*, 38
- osmoreceptor, 386*t*
- oxidation, 390*t*
- Paracelsus, 467
- parenteral
 - formulations, 104, 166, 371
 - irritation, 367–369
 - route, 367
- partition coefficients, 83, 87, 210, 385, 397
- passive absorption, 87, 392
- Passive Cutaneous Anaphylaxis (PCA)
 - assay, 251, 415
- Pasteur, Louis, 4
- “patent medicines,” 5
- patients with renal and hepatic
 - dysfunction, 601
- pediatric claims, 19
- Pediatric Use Labeling Rule, 37
- periocular route, 94
- permissible dosing volumes for nonhuman
 - primates, 281*t*
- Permitted Daily Exposure (PDE), 726–729
- Peto Analysis, 341
- Pharmaceutical Research and Manufacturers of America (PhRMA), 37, 526, 538, 583
- pharmacodynamics, 304, 406, 437*t*, 464, 586, 589, 599*t*, 600, 712, 754
- pharmacogenetics, 556, 589
- pharmacokinetics, 47, 164, 381, 395, 464, 589
- pharmacovigilance (PV), 414, 597, 637
- phase 0 (safety) studies, 129, 456
- phase I (metabolism) studies, 170*t*, 390*t*
- phase II (drug efficacy) studies, 70*t*
- phase III (clinical, therapeutic confirmatory) studies, 328
- phase III/IV (therapeutic use) studies, 599*t*
- phase IV (postmarketing) studies, 586
- phase Ib/I (therapeutic exploratory) studies, 314*t*
- phocomelia, 636
- photosensitization, 238, 258, 372, 555*f*, 566–568
- phototoxicity, 108*t*, 238, 278, 371–376
- Physiologically Based Pharmacokinetic (PBPK) modelling, 404–405, 531*t*
- pie charts, 696–697
- pig, 275, 417
- pilot toxicity studies, 154
- pinocytosis, 88, 408
- placebo control, 606
- Plaque-Forming Cell (PFC) assay, 242, 249, 253
- plasma protein binding, 83*t*, 382, 388
- plasmids, 172, 173
- platelets, 123*t*, 124, 234, 244*t*
- platinum salts, 729
- p53^{+/−} mouse model, 336
- polypharmacy, 497, 625
- pooling, 668, 801
- positive accuracy, 68, 546
- post hoc* tests, 679, 687
- postnatal development, 37, 47, 295–315
- potential new drugs in U.S. clinical trials, 7
- Precision, Accuracy, Sensitivity, Selectivity (PASS) validation, 154
- preclinical male fertility studies, 48
- pregnancy *see* developmental and reproductive toxicity (DART)
- preincubation tests, 183
- pre-IND
 - meeting, 9*f*, 38
 - nonclinical safety, 493
- preliminary cytotoxicity testing, 186
- preliminary studies, 297
- premarket approval applications (PMA), 594*f*
- prenatal development, 47, 295–315
- Prescription Drug User Fee Act (PDUFA), 18, 434*t*, 648
- Primary Dermal Irritation (PDI) test, 360
- primary mode of action (PMOA), 711–719
- probit/log transforms, 691
- probit method, 132, 149
- product class review responsibilities, 24*t*, 715*t*
- product withdrawals, safety reasons, 649
- Program Evaluation and Review Technique (PERT), 542, 543*f*
- project management, 542
- Prospulsid *see* Cisapride
- protein binding, 381, 384*t*, 388
- pulmonary
 - sensitization, 108*t*, 466
 - system, 313*t*, 386
- Pure Food and Drugs Act of 1906, 13, 436*t*
- Purkinje fibers, 418
- p value, 338–342, 655–656
- “pyramiding” studies, 133–138, 147
- pyrogenicity, 376, 452, 566, 573
- Q3A, ICH (impurities in new drug substances), 721–725
- Q3B(R2), ICH (impurities in new drug products), 725
- Q3C, ICH (impurities, guideline for residual solvents), 24, 725, 732
- QT/QTc, 42*t*
- qualification threshold, 215, 722–725
- Quantitative Structure-Activity Relationship (QSAR) models, 69, 209–222, 321, 531, 546, 574*t*, 722
- rabbit, 28, 47*t*, 110*t*
 - acute intramuscular irritation, 369
 - phototoxicity test, 373
 - vaginal irritation study, 365
- Radioactive Drug Research Committee (RDRC), 483
- radiochemical methods, 396
- Radio-Immune Assays (RIA), 254, 407
- Radithor, 15
- randomization, 116, 163, 365, 604, 661, 677
- range-finding study, 130–152, 297, 548, 708
- reaginic *see* hypersensitivity, type I (immediate; reaginic)
- receptors slowing gastric emptying, 386*t*
- rechallenge, 609, 624, 642
- Recombinant DNA Advisory Committee (RAC), 446, 738
- recombinant DNA technology, 7, 253, 433, 439
- rectal
 - administration, 91*f*, 94, 386*f*
 - routes, 81, 384
- red blood cells (RBCs), 123*t*, 198, 239, 355*t*, 568, 612*t*
- reduction of dimensionality, 573, 655, 658, 694–699
- regulatory pyramid, 649
- relevance to humans, 342
- renal
 - cells, 355, 490*t*
 - function, 312*t*, 413, 428, 601, 613
- repeat-dose studies, 30*t*, 48, 110, 160*t*, 269, 273, 365, 705, 741
- repeated-dose vaginal irritation, 365
- repeated measures, 284–288
- replacement, reduction, refinement *see* humane animal use in research
- replication, 170–177, 198, 231, 324, 659, 732–740
- reporting threshold, 722–725

- reproducibility, 68, 329, 384, 546, 657.
 reproductive
 effects, 116*t*, 176
 studies, female, 49
 Request For Designation (RFD), 711
 residual
 metals, 728
 solvents, 487, 516, 725
 residuals, 678, 693
 respiratory system, 120*t*, 414, 422–425, 464, 571*t*
 Resusci-Dog, 554
 Reye's syndrome, 353, 496*t*
 Rezulin (troglitazone), 635*t*, 648
 rheumatoid arthritis (RA), 229*t*, 232–240, 514
 ribosomal DNA (rDNA), 5, 433, 439–455
 rising dose tolerance, 152
 Risk Evaluation and Mitigation Strategies (REMS), FDA, 648–651 *see also* tier 1, 2, 3, 4, 5
 “rolling” acute test, 547*f*
 routes of administration, 81, 91–93, 139*t*, 272, 282, 332, 483, 515, 524, 586, 601, 726
 rules for form design and preparation, 664*t*

 Safe Medical Devices Act (SMDA) of 1990, 713
 safety studies, 163, 253, 279, 425, 712, 722, 732
 sample size, 31, 68, 284, 332, 654*t*, 662
 sampling, 397, 659
 sampling interval, 399, 661
 Sandoz Clinical Assessment-Geriatric (SCAG) scale, 618
 scattergram, 673
 Scheffe's Multiple Comparisons, 688
 screens, 67, 150, 544, 614
 secondary organ system, 415, 427
 selection of dosages, 437
 semiquartile distance, 658
 sensitivity, 68, 73
 sequential sampling, 401
 Serum Glutamic Oxaloacetic Transaminase, a.k.a. AST (SGOT), 121*t*, 122*t*, 612*t*, 613
 setting doses, 163
 Severely Toxic Dose (STD10), 36, 492*t*
 S1, ICH (carcinogenicity studies), 44, 45*t*, 328, 345
 S2, ICH (genotoxicity studies), 45*t*, 199, 475*t*
 S3, ICH (toxicokinetics and pharmacokinetics), 45*t*
 S4, ICH (toxicity testing), 45*t*
 S5, ICH (reproductive toxicology), 44, 45*t*, 293*t*
 S6, ICH (biotechnological products), 24, 32, 45*t*, 437, 456
 S7, ICH (pharmacology studies), 45*t*, 49, 413, 414*t*, 418, 419, 486*t*, 517*t*
 S8, ICH (immunotoxicology studies), 45*t*, 126, 225, 240, 241*t*, 242, 245, 261, 345
 S9, ICH (nonclinical evaluation for anticancer pharmaceuticals), 46*t*, 169
 S10, ICH (photosafety evaluation), 46*t*
 S11, ICH (nonclinical safety testing), 46*t*
 significance
 biological, 160, 191, 249, 260, 654, 705
 case I, II, III, IV, 654
 “false negative” (case III), 654
 “false positive” (case II), 654
 statistical, 154, 160, 196, 211, 341, 544, 654
 SIMUSOLV, 403
 singularity of drug, 624
 Sister Chromatid Exchange (SCE) assay, 176, 197
 six-lead ECG measurement, dogs, 417
 skeletal fetal examination, 302
 sorbitol dehydrogenase (SDH), 122*t*
 special classes of studies, 586
 special patient groups, 601
 special populations, 37, 601
 species selection, 162, 260, 439*t*, 455, 519, 732
 species selection for protein therapeutic, 439*t*
 specific immunotoxicity testing, CDER flowchart, 244*t*
 specificity, 68*t*, 214, 220, 406, 546, 597, 644, 702
 specific toxicity screening, 153
 Sprague Dawley rats, 106*t*, 305, 328, 343*t*
 Standard Error of the Mean (SEM), 658
 “statin,” 6
 statistical
 analysis, 304, 338, 655, 665–707
 graphics, by function, 696
 randomization, 163
 significance (*see* significance)
 status of nonanimal methods that have regulatory standards, 574*t*
 sterility, 367
 strategies for development, 8
 stratification, 585, 661, 668
 Structure-Activity Relationship (SAR), 210, 562, 804
 Student's *t* Test, 666*f*, 685
 studies of drug metabolites, 601
 subchronic study, 117, 131, 144, 161, 660
 subcutaneous (SC) route, 81, 85, 163, 281*t*, 367
 subunit vaccines, 449
 sulfonamide, 7*t*, 235*t*, 237, 344*t*, 422*t*
 Summary Basis of Approval (SBA), 23
 supplemented acute studies, 146
 suppressor mutations, 173
 survival, 332
 synthetic chemistry, 5, 7*t*
 Systemic Lupus Erythematosus (SLE), 232, 235*t*, 239, 354, 422*t*, 423*t*
 targeted therapeutics, 433
 target organ toxicity biomarkers, 331, 355
 Tarone's trend test, 669, 694
 T-Cell-Dependent Antibody Response (TDAR), 253
 T-Cell Lymphoproliferation Response, 250
 test material, 96, 682
 Tg.AC (v-Ha-ras) Transgenic Mouse Model, 335
 TGN1412, 11, 225, 433, 442
 Tg.rasH2 Mouse Model, 336
 thalidomide, 629, 636
The Jungle, (Upton Sinclair), 1, 436
 therapeutic drug delivery by dermal route, 474
 therapeutic index, 137, 381, 437, 537, 590, 608, 638*t*
 therapeutic products withdrawn from marketplace due to safety reasons, 630*t*
 Threshold Limit Value (TLV), 473, 523
 thresholds for action on impurities in drug product, 722*t*
 tier 1: mandatory studies, FDA REMS, 649
 tier 2: labeling and assessment, FDA REMS, 649
 tier 3: enhanced communication, FDA REMS, 650
 tier 4: safe-use restriction, defined by provider, FDA REMS, 650
 tier 5: safe-use restriction, defined by patient, FDA REMS, 651
 time course of effect, 591
 timing of studies, 297
 tissues for histopathology *see* histopathology
 T lymphocytes, 231
 tolerance, 544
 top 20 selling pharmaceuticals (2013), 2
 top 25 drug companies by sales (2014), 3
 TOPN (Federal Virus-Serum-Toxin Act), now Viruses, Serums, Toxins, and Anallygous Products (2015), 436
torsades de pointes (TdP), 418, 636
 toxicokinetics, 297, 381
 Toxic Release Inventory (TRI), 63
 Toxic Substances Control Act (TSCA), 215, 525, 560*t*
 transcription, 171
 transformations, 667*f*, 677
 transgenic mouse models, 261, 335
 translation, 171, 451
 trend analysis, 666*f*, 669, 694

- trend tests, 340, 669
- Troglitazone *see* Rezulin
- TSCA *see* Toxic Substances Control Act
- Tumor Necrosis Factor (TNF), 230*t*, 423*t*, 455, 615*t*
- type I error (false positives), 68, 154, 546, 654, 655*t*, 687–689
- type I hypersensitivity *see* hypersensitivity, type I (immediate; reaginic)
- type II error (false negatives), 68, 546, 654, 655*t*
- type II hypersensitivity *see* antibody-mediated; hypersensitivity, type II (cytotoxic)
- type III hypersensitivity *see* hypersensitivity; type III (Arthus reaction)
- type I immunotoxicity test, 126*t*
- type IV hypersensitivity *see* hypersensitivity, type IV (Delayed-Type Hypersensitivity (DTH))
- types of hypersensitivity responses, 236*t*
- types of screens, 71
- Tysabri (Natalizumab), 444*t*, 540*t*
- unconjugated bilirubin (UBili), 124*t*
- univariate
 - parametric tests, 684
 - repeated-measures, 284
- unscheduled DNA synthesis (UDS) assay, 178*t*, 190*t*, 574*t*, 724
- up/down method, 131, 152
- urinalysis, 117*t*, 121*t*, 355*t*, 428*t*, 614
- urine, 612*t*, 614, 706
- U.S. Code of Federal Regulations (CFR)
 - biotechnology products, 436
 - clinical trial subjects, 592, 593
 - 21 CFR (drugs and medical devices, testing, manufacture, sale), 21–24, 51
 - excipients, 106–109, 107*t*
 - FDA submissions, 716
 - imaging agents, 483
 - IRBs, 602–603
 - radiopharmaceuticals, 517
 - regulatory requirements, 639–640
- U.S. Food and Drug Administration (FDA), 13–19
 - Form 3500/3500A, 593, 640, 643*f*
 - Redbook II, 249, 293*t*, 331, 333, 382
 - Risk Evaluation and Mitigation Strategies (REMS), 648
- U.S. Pharmacopeia (USP), 25, 60, 106, 728
- Vaccine Adverse Event Reporting System (VAERS), 452, 636
- vaccines approved since 1986, 452
- vaginal
 - administration, 94
 - irritation, female rabbit, 364
- variant, 295
- vectors, 34, 447
- vehicle controls, 182, 530*t*
- Virus Act of 1902 (Virus-Toxin Law; Biologics Control Act), 13, 14*t*, 30
- Virus-Toxin Law *see* Virus Act of 1902
- V79 lines/system *see* Chinese hamster
- volume of distribution (V_D), 384*t*, 389, 561
- water solubility, 88, 93, 98, 389
- Wechsler Intelligence Scale for Children (WISC), 613*t*, 620
- whole-body autoradiography, 401
- Wilcoxon Rank-Sum Test, 681
- Wiley, Harvey, 13
- Williams' *t*-test, 689
- women of childbearing potential, 44, 298, 309, 438, 452
- world marketplace for drugs, 1
- World Medical Association (WMA), 598
- Wright Dust-Feed (WDF), 472
- xenobiotic metabolism, minipig, 277
- XPA^{-/-} mouse model, 337
- zebra fish, 563