Safety Pharmacology in Pharmaceutical Development and Approval

Safety Pharmacology

in Pharmaceutical Development and Approval

Shayne C. Gad

CRC PRESS

Boca Raton London New York Washington, D.C.

Library of Congress Cataloging-in-Publication Data

Gad, Shayne C., 1948-
Safety pharmacology in pharmaceutical development and approval / Shayne C. Gad.
p. ; cm.
Includes bibliographical references and index.
ISBN 0-8493-1380-5
1. Drugs—Toxicology. 2. Drugs—Safety measures. I. Title.
[DNLM: 1. Drug Evaluation, Preclinical. 2. Drug Approval. 3. Toxicity Tests. QV 771 G123s 2003]
RA1238.G333 2003
615′.704—dc21 2003051468

International Standard Book Number 0-8493-1380-5
Library of Congress Card Number 2003051468
Printed in the United States of America 1 2 3 4 5 6 7 8 9 0
Printed on acid-free paper

Dedication

To my beloved Joyce

Contents

Preface

Safety pharmacology is the evaluation and study of the pharmacological effects of a potential drug that are 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 a drug's safety, these evaluations are usually conducted at doses close to the intended clinical dose. Recent International Conference on Harmonization (ICH) guidelines, followed by versions by national regulatory authorities including those of the U.S. Food and Drug Administration (FDA), have made such evaluations mandatory before a potential drug is introduced into humans — while also failing to provide clear guidance as to how the requirements are to be met.

General/safety pharmacology has been an emerging discipline within the pharmaceutical industry in which unanticipated effects of new drug candidates on major organ function (i.e., secondary pharmacological effects) 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 U.S. was (and remains) in a position behind Japan and the European Union (EU) in establishing formal requirements and in implementing industrial programs. Most companies have traditionally conducted evaluations of cardiovascular and central nervous system (CNS) functions, while few have evaluated respiratory, gastrointestinal, and renal functions; a few conduct a ligand-binding/activity panel as part of their pharmacological profiling. Resources needed to complete a company's standard safety pharmacology program are approximately one to four full-time persons per compound. One-third of companies use a maximum tolerated dose (MTD) for safety pharmacology studies; two-thirds use multiples of pharmacological or therapeutic doses. Until 2000, only half conducted safety pharmacology studies with Good Laboratory Practices (GLPs), using the 1992 Japanese guidelines as a guide or outline. Company clinicians were most often cited as the primary customers for whom safety pharmacology studies were done, followed by research and development scientists, and then

* Chapter 1, References 19 and 20.

regulatory authorities. These results suggest that most companies primarily conducted safety pharmacology for its contribution to risk assessment and critical care management.

It is important that the tests employed detect bidirectional drug effects and that the tests performed be validated in both directions with appropriate reference (control) substances. 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 overly inclusive, but should be performed to the most exacting standards, including GLP compliance. This is, of course, backward; such human tolerance is properly an extension (and expression) of the nonclinical safety pharmacology.

The other point of view 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 biologicals 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. In large animal species, such as dogs and nonhuman primates, smaller numbers 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 biologicals 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 endpoints in the context of toxicity studies (Martin et al. 1997; Matsuzawa et al. 1997).*

In Europe, the numbers of registered drugs and drug expenditure are increasing rapidly. Within the EU, there are no longer any regulations requiring that new drugs have to be better than old ones. At the same time, pharmacoepidemiology studies in Europe and the U.S. show that adverse drug reactions now may account for up to 10% of the admissions of patients to hospitals at a cost of hundreds of millions of U.S. dollars annually (Sjouist, 2000)**. Compared to 20 years ago, this represents a considerable increase.

* Chapter 1, References 37 and 38, respectively.
** Chapter 1, Reference 39.

The many shortcomings of clinical trials and their relevance to health care provide a partial explanation. Adverse drug reactions are often poorly studied and documented in these trials and 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 because up to 50% of adverse drug reactions (ADRs) are dose dependent and thus preventable. Hopefully, the rapid progress in molecular and clinical pharmacogenetics will provide new tools to enable clinicians to choose and dose drugs according to the needs of individual patients. A good starting point for those not well versed in pharmacology and the range of potential mechanisms of action and interaction can be found in *Goodman & Gilman's The Pharmacological Basis of Therapeutics* (Hardman and Limbird, 1996).*

In this essential and rapidly changing field, it is hoped that this first volume on the subject will answer many questions and add clarity to existing requirements.

* Chapter 1, Reference 40.

The Author

Dr. Shayne C. Gad, Ph.D. (Texas, 1977), **DABT, ATS**, is Principal of Gad Consulting Services. His experience includes safety assessment and product development in the pharmaceutical, biotechnology, medical device, and chemical industries. He has published 26 books and more than 300 chapters, articles, and abstracts in the fields of toxicology, statistics, pharmacology. and safety assessment, and is on the editorial boards of *Toxicology Methods, Journal of Applied Toxicology, Journal of Fire Science, Journal of Acute Toxicology,* and the *International Journal of Toxicology.* He has served on the NIH, NIEHS, Canadian government, and nongovernmental organization grant review boards. He has written and filed 58 INDs plus numerous BLAs, PLAs, 510(k)s, IDEs, NDAs, PMAs, and CTDs.

He is a past president of the American College of Toxicology, the Roundtable of Toxicology Consultants, and three specialty sections of the Society of Toxicology. He is also a member of the Regulatory Affairs Professional Society, Society of Toxicology, Teratology Society, Biometrics Society, Society of Toxicologic Pathologists, American Statistical Association, and Drug Information Association, and is a Diplomate of the American Board of Toxicology and a Fellow of the Academy of Toxicologic Sciences. He is also a retired line Captain (USNR) with extensive overseas service.

chapter 1

Safety pharmacology: background, history, issues, and concerns

"The adverse drug reactions, which the standard toxicological test procedures do not aspire to recognize, include most of the functional side effects. Clinical experience indicates, however, that these are much more frequent than the toxic reactions due to morphological and biochemical lesions … ".[1]

Safety pharmacology is the evaluation and study of the pharmacologic effects of a potential drug (or excipient) that are 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 a drug's safety, these evaluations are usually conducted at doses close to the intended clinical dose.

Such pharmacologically based adverse effects have not figured in phase I (first in man) deaths of volunteers for some years. Although such volunteer deaths are very rare, severe adverse events are not (Table 1.1).

In the U.S. from 1954 until 1980, 7,000,000 healthy volunteers participated in studies with only one death (in 1980, not due to adherence to study participation).[2] Indeed, in the U.S., there was one volunteer death in 2001[3] and none in 2002.

Similarly, in a 12 month period, there were no deaths in 8163 healthy volunteers with the 3 reported severe adverse events being a severe dermatitis, an anaphylactic shock, and perforation of an ulcer.[4] In a Phase I center in France in 1015 healthy young volunteers, there were 34 severe adverse events, but no deaths.[5] None of the drugs later withdrawn from the market for QT interval prolongation were detected in such studies.

This suggests that screening in normal, healthy volunteers is unlikely to serve to detect functional (pharmacological) effects such as those that have served to remove drugs from the market. When conducted properly, such studies are in the wrong population (a healthy and, therefore, fairly insensitive one) and at too low a dose (one safe to the volunteers) to detect a signal of concern.

Table 1.1 Risks to Healthy Volunteers in Clinical Pharmacology Studies

Year	Country	Number of volunteers	Moderately severe AE	Potentially life-threatening AE	Deaths
1965–77	U.S.	29,162	58 (0.2%)	—	0
1980	U.S.	—	—	—	1
1983	Ireland	—	—	—	1
1984	U.K.	—	—	—	1
1986–87	U.K.	8,162	45 (0.55%)	3 (0.04%)	0
1986–95	France	1,015	43 (3%)	0	0
2001	U.S.	—	—	—	1

Note: AE = adverse effect.

Source: Modified from Darragh, A. et al., *Lancet*, 1:93–94, 1985.

General vs. safety pharmacology

A wide variety of general pharmacology screens have been available since the 1940s and 1950s.[7,8] The value of many of these screens has been demonstrated in a number of instances where potentially important adverse effects of a novel therapeutic were identified prior to use in clinical trials.[9–13] Such general pharmacology tests evaluate affinity for a pharmacologic target (receptor), pharmacodynamic activity, and interactions in pharmacodynamic processes.[14–18] However, with the extensive advances in biomedical research in the last two decades — especially in molecular biology, immunology, and neurobiology — many novel types of therapeutic compounds are being developed.

At the same time, pharmaceutical and biotechnology companies are being pressured by wary investors, competition, and advocates for the terminally ill to deliver safe therapeutics more rapidly. In this atmosphere, it is incumbent upon the research scientist to carefully consider each new therapeutic under development and determine the optimal methods for measuring the degree of safety of a compound as much as possible prior to its use in humans. No simple formula or set group of safety pharmacology screens is ideal for all kinds of therapeutic agents. Knowledge of the pharmacology of a compound and any knowledge gained from traditional toxicology studies, or the incorporation of some safety pharmacology endpoints within a toxicology study, can help to better determine and assess the safety of new therapeutics. Working with regulatory agencies, clinical pharmacologists, and basic scientists, researchers who are developing compounds in industry can establish the most appropriate schedules of studies and screens to safely advance their products into the clinic. Objectives of safety pharmacology studies include:

- Identifying undesirable pharmacodynamic properties
- Evaluating adverse pharmacodynamic and pathophysiological effects
- Investigating the mechanism of the adverse pharmacodynamic effects observed or suspected

General/safety pharmacology has been an emerging discipline within the pharmaceutical industry in which unanticipated effects of new drug candidates on major organ function (i.e., secondary pharmacological effects) 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. Adverse functional effects were overlooked in importance by all but a few[19–22] for many years.The Japanese clearly became the leaders in developing and requiring such information,[23–24] while the U.S. was in a position behind Japan and the European Union (EU) in having formal requirements and in implementing industrial programs. Although major companies were aware and largely addressing the need by the mid-1990s,[25–28] initial Food and Drug Administration (FDA) guidelines were not proposed and promulgated until the middle of 2001, and the Committee for Proprietary Medicinal Products (CPMP) only 2 years before this.[29–31] The operational issue of what is to be done, how it is to be done, and how the resulting data will be used is still being worked out in the third quarter of 2003.[32] Table 1.2 presents examples of known adverse functional events.

History

The increased interest and eventual regulatory requirements for functional safety (safety pharmacology) testing have resulted due to the occurrence of a number of cases having to be withdrawn due to unacceptable levels of unanticipated deaths occurring after drugs have entered the market place, mandating the drugs' withdrawal.

Table 1.3 presents a summary of drugs withdrawn from the market, since 1990, due to safety reasons.[33] It should be noted that only 4 (denoted by an asterisk) of the 14 withdrawals were due to adverse safety pharmacology; 3 of the 4 were due to cardiovascular incidents.

Adverse effects can be considered as one of five types (Table 1.4), of which only one is likely to be detected by conventional safety pharmacology testing.

Most companies conduct evaluations of cardiovascular and central nervous system (CNS) functions; fewer evaluate respiratory, gastrointestinal, and renal functions; a few conduct a ligand binding/activity panel as part of their pharmacological profiling. The resources needed to complete a company's standard safety pharmacology program would include approximately one to four full-time persons per compound. One-third of large companies use a maximum-tolerated dose (MTD) for safety pharmacology studies; two-thirds use multiples of pharmacological or therapeutic doses. Company clinicians are most often cited as the primary customer for whom safety pharmacology studies are done, followed by research and development scientists, and then regulatory authorities. These results suggest that most companies primarily conduct safety pharmacology for its contribution to risk assessment and critical care management.[34]

It is important that the tests employed detect bidirectional drug effects and that the tests performed be validated in both directions with appropriate

Table 1.2 Examples of Functional Toxicity

Extensions of primary pharmacological activity	
Compound/compound class	**Adverse effect**
Sympathomimetic bronchodilators	Tachycardia, palpitations
Digitalis	Hypotension arrhythmias, ventricular tachycardia
Antihistamines	Sedation
Antihypertensives	Postural hypotension
Suxamethonium	Neuromuscular blockade (esp. prolonged apnea)
Corticosteroids	Adrenal hypofunction
Insulin	Tachycardia
Antiarrhythmics	Arrhythmia
Nitroglycerin	Hypotension
Anticoagulants	Hemorrhage
Unrelated to primary pharmacological action	
Compound/compound class	**Adverse effect**
β-lactam antibiotics	Convulsions
Aminoglycoside antibiotics	Neuromuscular blockade
Trycyclic antidepressants	Anticholinergic effects (dry mouth, blurred vision, hypertension), arrhythmias
Emetine	Hypotension, tachycardia, EKG abnormalities
Digitalis	Gynecomastia
Antihypertensives	Sedation
Streptozolocin	Hyperglycemia
Antihistamines	Anticholinergic effects
Sulfonamides	Diarrhea, nausea, emesis
Vancomycin	Hypotension

Note: EKG = electrocardiogram.

reference (control) substances. This requirement is less appropriate for multi-parameter 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 over-inclusive, but should be performed to the most exacting standards, including Good Laboratory Practice (GLP) compliance.[35] General pharmacology screens are generally not conducted in accordance with GLPs, and thus the quality and reproductivity can be suspect.[36] Tier I safety pharmacology data must be available during the planning stage for Phase I studies, but such is only just now recognized to be the case. 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,

Table 1.3 Post-Approval Adverse Side Effects and Related Drug Withdrawals after 1990

Year	Drug	Indication/class	Causative effects
1991	Enkaid (4 years on market)	Antiarrhythmic	Cardiovasuclar (sudden cardiac death)[a]
1992	Temafloxacin	Antibiotic	Blood and kidney damage
1997	Fenfluramine[a]/Dexafluamine (combo used since 1984) (24 years on market)[a]	Diet pill	Heart valve abnormalities
1998	Posicor (Mibefradil) (1 year on market)	Ca^{++} channel blocker	Lethal drug interactions (inhibited liver enzymes)
	Duract (bronfenac sodium) (early preapproval warnings of liver enzymes)	Pain relief	Liver damage
1999	Tronan (use severely restricted)	Antibiotic	Liver/kidney damage
	Raxar	Quinolone antibiotic	QT internal prolongation/ventricular arrhythmias (deaths)[a]
	Hismanal	Antihistamine	Drug–drug interactions
	Rotashield	Rotavirus vaccine	Bowel obstruction
2000	Renzulin (approved Dec. 1996)	Type II diabetes	Liver damage
	Propulsid	Heartburn	Cardiovascular irregularities/deaths[a]
	Lotonex	Irritable bowel syndrome	Ischemic colitis/death[a]
2001	Phenylpropanolamine (PPA)	OTC ingredient	Hemorrhagic stroke
	Baychlor	Cholesterol reducing (satin)	Rhabdomyolysis (muscle-weakening) (deaths)

Note: 51% of approval drugs had serious postapproval identified side effects.

[a] Functional (pharmacologic) effects.

Source: From Gad, S.C. *Drug Safety Evaluation*, John Wiley & Sons, New York, 2002.

Table 1.4 Classification of Adverse Drug Reactions (ADRs) in Humans

Type A	Dose-dependent; predictable from primary, secondary, and safety pharmacology	Main cause of ADRs (~75%), rarely lethal
Type B	Idiosyncratic response, not predictable, not dose-related	Responsible for ~25% of ADRs, but majority of lethal ones
Type C	Long-term adaptive changes	Commonly occurs with some classes of drug
Type D	Delayed effects (e.g., carcinogenicity, teratogencity)	Low incidence
Type E	Rebound effects following discontinuation of therapy	Commonly occurs with some classes of drug

Note: Conventional safety pharmacology studies can only reasonably be expected to predict Type A ADRs. Functional toxicology measurements may predict Type C ADRs. Conventional toxicology studies address Type D ADRs. Prediction of Type B responses requires a more extensive preclinical and clinical evaluation, often only addressing risk factors for the idiosyncratic reponses (e.g., QT prolongation for torsades de pointes). Type E ADRs are rarely investigated preclinically using functional measurements unless there is cause for concern.

Source: Modified from Redfern, W.S. et al., *Fundamental Clinical Pharmacol.,* 16:161–173, 2002.

of course, backward; such human tolerance is properly an extension and expression of the nonclinical safety pharmacology.

Reasons for poor predictive performance

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 biologicals 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.

In large animal species, such as dogs and nonhuman primates, smaller numbers 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 biologicals 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 endpoints in the context of toxicity studies.[37,38]

For a variety of reasons, preclinical safety pharmacology tests may not accurately predict human adverse effects:

1. Species differences in the presence or functionality of the molecular target mediating the adverse effect.
2. Differences in ADME between test species and humans.
3. Sensitivity of the test system — observations of a qualitative nature should be followed up with specific, quantitative assessment.
4. Poor optimization of test conditions — the baseline level has to be set correctly to detect drug-induced changes.
5. Study designs that are underpowered statistically.
6. Inappropriate timing of functional measurements in relation to T_{max}.
7. Delayed effects — safety pharmacology studies generally involve a single administration with time points for up to 24 h postdose.
8. Difficulty of detection in animals — adverse effects such as headache, disorientation, and hallucinations are quite a challenge to safety pharmacologists.

In Europe, the numbers of registered drugs and drug expenditures are increasing rapidly. Within the EU, there are no longer any regulations requiring that new drugs have to be better than old ones. At the same time, pharmacoepidemiology studies in Europe and the U.S. show that adverse drug reactions now may account for up to 10% of the admissions of patients to hospitals at a cost of hundreds of millions of U.S. dollars annually.[39] This represents a considerable increase compared to 20 years ago. Shortcomings of clinical trials and their relevance for health care provide a partial explanation. ADRs are often poorly studied and documented in these studies and 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 use to dose drugs according to the needs of individual patients. A good starting point for those not well versed in pharmacology and the range of potential mechanisms of action and interaction can be found in Hardman and Limbird.[40]

Why tiers?

One somewhat confusing aspect of the promulgated guidelines is the division of organ systems to be evaluated into Core Tier I – which must be performed prior to human studies and includes evaluations of central nervous system,

Table 1.5 Differences between Safety Pharmacology and General Pharmacology

	Safety pharmacology	General pharmacology
Dose range	Efficacious to effect level or dose producing toxicity	Efficacious to multiple of efficacious dose
Organizational Location	Development	Discovery
GLP	Yes	No
Objectives		
Identify ancillary actions/ side effects	Yes	Yes
Establish effect/no effect levels	Yes	No

cardiovascular, and respiratory/pulmonary system and Optional Tier II – which must be considered on a case-by-case basis prior to drug registration and includes the other organ systems.

The rationale for this tier approach has to do with the judged relative hazard presented by effects on the different organ systems. Acute adverse function effects on core organ systems can be fatal. Adverse effects on these Tier II systems in the optional tier, while of concern and potentially fatal, are generally unlikely to be acutely fatal. The exception to this is the immune system, but it is felt that the existent separate guideline for immuno-toxicity evaluation calls for adequate steps to guard against acutely lethal effects on this system.

Study designs and principles

As a starting place and unlike older pharmacology studies (Table 1.5), safety pharmacology studies are normally conducted as GLP studies. At the same time, unlike other safety assessment studies, these do not need to vastly exceed intended therapeutic doses in order to reveal signs of toxicity. In this sense, they are closer to hazard tests. General considerations in the selection and design of studies include:

- 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

Table 1.6 presents a summary of guidance for the key design features for safety pharmacology studies.

Table 1.6 Summary of Key Study Design Features of the Safety Pharmacology Core Evaluation

Attribute	Recommendation	Comment
Animal model	Conscious, unrestrained, telemeterized, and trained	Should approximate conditions of Phase I clinical model
Test species	Rodent or nonrodent species used in toxicology studies, or most appropriate species based upon scientific considerations	Generally, the toxicology species are the most appropriate and specific synergies are available when these species are used
Statistical design	Random blocked study (each animal receives all treatments and serves as its own control)	Blocked study designs reduce animal used by up to 75% without loss of statistical power
Group size	Sufficient to rule out a significant effect of the drug	Based upon laboratory experience with the model and a statistical power analysis
Controls	1. Negative (vehicle or placebo) 2. Positive	1. Always 2. Only when deemed necessary to ensure appropriate functionality of the test system
Route of administration	Best approximation of the clinical route	Should approximate conditions of Phase I clinical model
Test article formulation	Best approximation of the clinical formulation, given constraints posed by specific species	Formulations should be those used in toxicology or DMPK studies, so that bioavailability data may be used to assist in safety pharmacology study design and data interpretation
Dose range	Three doses including a maximum tolerated dose	Dose-response requirements (S7A)
High dose	1. Maximum tolerated dose 2. Maximum feasible dose 3. Limit dose	1. Overlap toxicological dose range 2. Limited by formulation or dose volume constraints 3. 1–2 g/kg

(*continued*)

Table 1.6 (continued) Summary of Key Study Design Features of the Safety Pharmacology Core Evaluation

Attribute	Recommendation	Comment
Pharmacokinetics	Systemic exposure should be documented within the study or referenced to another relevant study	If SP study design is coordinated with toxicology program, toxicokinetic data may be used to establish safety pharmacology exposures
Endpoints measured	Arterial blood pressure, heart rate, EKG	ICH S7A requirements for cardiovascular core study
Duration	Generally performed as a single dose study Sufficient to cover:	
	1. Initial distribution	1. Acute, transient (anaphylactoid) effects
	2. Cmax	2. Acute concentration-dependent effects
	3. AUC phases (generally 24 h)	3. Delayed effects
Timing with respect to clinical development	Prior to first administration in humans	ICH S6: S7A Guidelines
Applicability of GLPs	Core battery studies should be performed to GLP standards to the greatest extent possible	Aspects of the study not conducted to GLPs should be identified and explained in the study protocol

Note: AUC = area under the curve.

From International Conference on Harmonization [ICH] S7A.

Selection of methodology and species

The following selection criteria are worth considering:

1. It is preferable to use the same species for *in vivo* tests as those used in drug metabolism and pharmacokinetics (DMPK) and toxicology — generally rats and dogs (an exception would be when the primary pharmacological target in those species is different from the human form).
2. The methods should be well established (i.e., not a test invented internally that has not been subject to external evaluation).
3. The methods/tests should be in common use in research and not in major decline.
4. The methods/tests should be validated in-house with at least one reference substance with known effect in humans.
5. The methods/tests should give reliable, reproducible results every time.
6. The level of technical difficulty should be compatible with routine use in a pharmaceutical/CRO environment.

General guidance for dose (or concentration) section for such studies is:

1. *In vivo* studies:
   - Should be designed to define the dose response curve of the adverse effects.
   - Doses should include and exceed primary pharmacodynamic or therapeutic ranges.
   - In the absence of safety pharmacology parameters, the highest doses should equal or exceed some adverse effects (toxic range).
2. *In vitro* studies:
   - Generally should be designed to establish and effect concentration relationship (range of concentrations).

Consideration in the selection and design of specific studies is straightforward.

1. 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
2. A hierarchy of organ systems can be developed:
   - Importance with respect to life-supporting functions:
     - Cardiovascular

    - Respiratory
    - Central nervous system
- Functions that can be transiently disrupted without causing irreversible harm:
    - Renal/urinary system
    - Autonomic nervous system
    - Gastrointestinal system
    - Other organ systems
- No testing is usually considered necessary for:
    - 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 (refer to toxicology studies)
    - New salts having similar pharmacokinetics and pharmacodynamics

Issues

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.[41] 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 biologies arises from both their physicochemical properties and their similarity to endogenous substances that are regulated in a carefully controlled manner. To overcome the 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 necessary only if kinetics/dynamics are substantially altered in comparison to the active substance previously tested.

Special considerations also exist with respect to how to statistically evaluate specific aspects of these studies. Specifically, analysis of time to event becomes very important.[42]

Integral vs. separate

Two general approaches to the *in vivo* preclinical evaluation of safety pharmacology are available. Such evaluation can either be performed as (separate) free-standing studies focused solely on the pharmacologic endpoints of concern or as integral evaluations conducted on animals that are part of a modified safety assessment study.[38,43]

The arguments for conducting free-standing studies are: (1) toxicology studies are generally conducted at a higher dose level than is the case of safety pharmacology studies, and (2) special manipulations, such as implanted sensors, which are potentially desirable for safety pharmacology evaluations, may not be practical in toxicity studies.

The issues are subject to considerable disagreement. Certainly the first argument is weak at best — while traditional toxicology studies do not include a group dosed at the projected clinical dose, the lowest dose is usually a modest multiple of such a dose (5 or 10×) and within the range of doses desirable to be covered in a safety pharmacology study. The second point presupposes two points: (1) special manipulations are required to adequately evaluate potential hazard, and (2) required manipulations compromise the inclusion of involved animals in the regular toxicology evaluations.

The first appears to be a weak point at best and likely not valid. Certainly in the case of the pre-investigational new drug (IND) CNS safety pharmacology evaluation, the proposed free-standing Irwin screen evaluation was the basis for the functional observational battery (FOB). FOB is standardly inducted in the rodent GLP toxicity studies, and was developed specifically for this purpose.[44] Similarly, the EKG evaluation performed in the standard pivotal nonrodent (particularly dog) toxicology study — which supports an IND — should serve to provide a sensitive (certainly adequate) detector of any effects on heart rate or cardiac electrophysiological events. Only the regulatorily required evaluation of respiratory effects does not appear to be addressed currently.

The second is a stronger argument, though again not in the case of the CNS pharmacology evaluation. As Table 1.7 summarizes, the coverage of desired endpoints in an evaluation of cardiovascular and respiratory endpoints in current "standard" toxicology designs might address such shortcomings with alterations in current study designs, but are not currently so.

The arguments for integration into existing study designs may also be viewed as a mixed bag. These are generally seen as: (1) the desirability of avoiding both the use of additional animals and the cost of such additional studies, and (2) the potential greater value and ease of interpretation of results when (potentially) isolated effects with one endpoint can be viewed in the context of the more thorough array of information captured in a well-designed and executed GLP toxicity study.

Table 1.7 Integration of Parameters into Safety Studies

Parameter	Safety pharmacology	Toxicology	
		Rodent	Non-rodent
Functional observational battery	++	++	+
Cardiovascular system	++	-	++
EKG			
Blood pressure	+	-	+
Systolic	+	+	-
Diastolic	+	-	-
LVP	+	-	-
Respiratory system	+++	++	—
Inspiration time	+	+	
Expiration time	+	Plethysmography in	
Peak inspiratory flow	+	conscious	
Peak expiratory flow		unrestrained	
Respiratory rate	+	animals	
Tidal volume	+	+	
Resistance	+	-	
Compliance	+	-	
Hyperreactivity	+	-	

+++, already integrated; ++, integration after adaptation easy; +, integration after adaptation possible; -, integration not feasible.

The first argument against unwarranted use of animals is powerful and must be addressed. It should be remembered, however, that most safety pharmacology studies are not terminal — the animals employed are most commonly "washed out" and then reused in subsequent evaluations. The cost issue, while important, is less compelling when weighed against concerns of potential safety.

The second argument, integrated data evaluation, must also be carefully considered. It is, in a sense, the other side of the argument which says that separate studies can best address the full range of functional endpoints of concern (as presented in Table 1.7). Is one side a necessary and essential trade-off for the other?

Summary

The regulatorily mandated safety pharmacology (functional safety) studies clearly should serve to decrease potential and undue hazards identified postmarket in potentially large populations. However, the actual implementation of requirements and the use of the resulting data in risk/benefit decisions are still in a state of flux and will require time to be fully worked out and understood.

References

1. Zbinden, G., *Pharmacological Methods in Toxicology,* Elmsford, NY: Pergamon, 613, 1979.
2. Kolata, G.B., The death of a research subject, *The Hastings Center Report,* 10:5–6, 1980.
3. Marshall, E., Volunteer's death prompts review, *Science,* 292:2226–2227, 2001.
4. Orme, M. et al., Healthy volunteer studies in Great Britain: the results of a survey into 12 months activity in this field, *Br. J. Clin. Pharmacol.*, 27:125–133, 1989.
5. Sibille, M. et al., Adverse events in Phase I studies: a report in 1015 healthy volunteers, *Eur. J. Clin. Pharmacol,* 54:13–20, 1998.
6. Darragh, A. et al., Sudden death of a volunteer, *Lancet*, 1:93–94, 1985.
7. Irwin, S., Drug screening and evaluation of new compounds in animals, *Animal and Clinical Pharmacologic Techniques in Drug Evaluations,* Nodine, H. and Siegler, P.E., Eds., Philadelphia: Year Book Medical, 1964.
8. Turner, R.A., *Screening Methods in Pharmacology,* Vols. I and II, New York: Academic, 42–47, 60–68, 27–128, 1965.
9. Bramm, E., Binderup, L., and Arrigoni-Martelli, E., An unusual profile of activity of a new basic antiinflammatory drug, Timegadine, *Agents Actions*, 11:402–409, 1981.
10. Graf, E. et al., Animal experiments on the safety pharmacology of Lofexidine, *Arzneim.-Forsch./Drug Res.,* 32(II)(8a):931–940, 1982.
11. Lumley C.E., General pharmacology, the international regulatory environment, and harmonization of guidelines, *Drug Dev. Res.*, 32:223–232, 1994.
12. Proakis, A.G., Regulatory considerations on the role of General Pharmacology Studies in the development of therapeutic agents, *Drug Dev. Res.*, 32:233–36, 1994.
13. Igarashi, T., Nakane, S., and Kitagawa, T., Predictability of clinical adverse reactions of drugs by general pharmacology studies, *J. Toxicol. Sci.,* 20:77–92, 1995.
14. Hite, M., Safety pharmacology approaches, *Int. J. Toxicol.,* 16:23–31, 1995.
15. Folke, S., Drug safety in relation to efficacy: the view of a clinical pharmacologist, *Pharmacol. Toxicol.,* 86:30–32, 2000.
16. Fujimori, K., 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, 1999.
17. Thompson, E.B., Drug bioscreening, in *Drug Evaluation Techniques in Pharmacology,* New York: VCH Publishers, 366, 1990.
18. Sills, M., In vitro screens and functional assays to assess receptor pharmacology, *Drug Dev. Res.,* 32:260–268, 1994.
19. Zbinden, G., The significance of pharmacologic screening tests in the preclinical safety evaluation of new drugs, *J. New Drugs*, 6:1–7,1966.
20. Zbinden, G., Neglect of function and obsession with structure in toxicity testing, *Proc. 9th Int. Cong. Pharmacol.,* Vol. 1, New York: Macmillan, 43–49, 1984.
21. Williams, P.D., The role of pharmacological profiling in safety assessment, *Reg. Toxicol. Pharmacol.,* 12:238–252, 1990.

22. Williams, P.D., Proposal for a core battery-package for Phase I, Sundwall A. et al., *Workshop In the Use of Pharmacology Studies in Drug Safety Assessment — Present Situation and Future Perspectives*, Tryckgruppen: Stockholm, 133–138, 26–27, September 1994.
23. Anon., *Guidelines for the Safety Pharmacology Study Required for Application for Approval of Manufacturing (Import) of New Drugs: Notification No. 4*, Director of New Drug Division, Pharmaceutical Affairs Bureau, Takushin-Yaku: Japan Ministry of Health and Welfare, 29 January, 1991.
24. Anon., Guidelines for general pharmacology, *Drug Approval and Licensing Procedures in Japan,* Tokyo: Yakugyo Jiho, 137–140, 1992.
25. Kinter, L.B., Gossett, K.A., and Kerns, W.D., Status of safety pharmacology in the pharmaceutical industry, 1993, *Drug Dev. Res.,* 32:208–216, 1994.
26. Sullivan, A.T. and Kinter, L.B., Status of safety pharmacology in the pharmaceutical industry, *Drug Dev. Res.,* 12:238–252, 1995.
27. Kurata, M., Kanai, K., and Mizuguchi, K., Trends in safety pharmacology in the U.S. and Europe, *J. Toxicol. Sci.*, 22: 237–248,1997.
28. Olejiniczak, K., Development of a safety pharmacology guideline, *Human Exper. Toxicol.,* 18:502, 1999.
29. CPMP, Note for Guidance on Safety Pharmacology Studies in Medicinal Product Development, 1998.
30. ICH, Safety Pharmacology Studies for Human Pharmaceuticals, 2000.
31. ICH, Safety Pharmacology Studies for Assessing the Potential for Delayed Ventricular Repolarization (QT Interval Prolongation), by Human Pharmaceuticals, 2000.
32. Claude, J.R., Safety pharmacology in the nonclinical assessment of new medicinal products: definition, place, interest and difficulties, *Fund. Clin. Pharmacol.,* 16:75–78, 2002.
33. Gad, S.C., *Drug Safety Evaluation*, New York: John Wiley & Sons, 2002.
34. Redfern, W.S. et al., Safety pharmacology — a progressive approach, *Fundamental Clinical Pharmacol.,* 16:161–173, 2002.
35. Anon., Applicability of good laboratory practices, Committee on Proprietary Medicinal Products III (3824/92): Rev. 1, item 9, 1992b.
36. Spindler, P. and Seiler, J.P., The quality management of pharmacology and safety pharmacology studies, *Fund. Clin. Pharmacol.,* 16:83–90, 2002.
37. Martin, L.I., Horvath, C.J., and Wyand, M.S., Safety pharmacology screening: practical problems in drug development, *Int. J. Toxicol.*, 16:41–65, 1997.
38. Matsuzawa, T. et al., Current status of conducting function tests in repeated dose toxicity studies in Japan, *J. Toxicol. Sci.,* 22:374–382, 1997.
39. Sjouist, F., Drug-related hospital admissions, *Ann. Pharmacol.,* 34:832–839, 2000.
40. Hardman, J.C. and Limbird, L.E., *Goodman & Gilman's The Pharmacological Basis of Therapeutics,* 9*th* ed., New York: McGraw-Hill, 1996.
41. Green, M.D., Problems associated with the absence of activity in standard models of safety pharmacology used to assess biological products, *Int. J. Toxicol.,* 16:33–40, 1997.
42. Anderson, H. et al., Statistical analysis of time to event data from preclinical safety pharmacology studies, *Tox. Methods,* 10:111–125, 2000.
43. Luft, J. and Bode, G., Integration of safety pharmacology endpoints into toxicology studies, *Fund. Clin. Pharmacol.,* 16:91–103, 2002.
44. Gad, S.C., A neuromuscular screen for use in industrial toxicology, *J. Toxicol. Environ. Health,* 9:691–704, 1982.

chapter 2

Regulatory requirements of ICH, U.S. FDA, EMEA,* and Japan MHW

Regulatory requirements

Although the International Conference on Harmonization (ICH) guidelines promulgated in November 2000 (implemented in Europe and the U.S. in June 2001) are the announced international standards for regulation, the actual situation in different countries remains mixed.

Japan continues to operate in conformance with the Ministry of Health and Welfare (MHW) draft (Guidelines for Safety Pharmacology Studies).[1] In 1999, this was revised. The basic principle of the revision is to harmonize the guideline with the international concepts. The working group decided to change the title of General Pharmacology to Safety Pharmacology because the objective of this guideline is to assess the safety of a test substance in humans by examining the pharmacodynamic properties of the substance (note that Japan has separate guidelines for general pharmacology).[2] The proposed guideline includes essential studies on vital functions that should be performed prior to human exposure. Studies are also required when predictable or unexpected observed effects are concerned. The working group recommends a case-by-case approach to select the necessary test items in consideration of the variable information available.

In the European Union (EU), the Committee on Proprietary Medicinal Products (CPMP) issued a draft *Note for Guidance on Safety Pharmacology Studies in Medicinal Product Development* in 1997,[3] but it was not finalized or put in force until the middle of 2001. The U.S. Food and Drug Administration (FDA) promulgated equivalent guidance at the same time, but the exact details of compliance and implementation, as will be seen in this volume, are still being worked out.[4]

* European Agency for the Evaluation of Medicinal Products

Table 2.1 Regulatory Guidelines for Safety Pharmacology

	ICH	U.S. FDA	EMEA/CPMP	JAPAN/MHW
Safety assessment of pharmaceuticals	M3 Nonclinical Safety Studies for the Conduct of Human Clinical Trials for Pharmaceuticals[6]	Guidance for Industry: Nonclinical Safety Studies for the Conduct of Human Clinical trials for Pharmaceuticals[7]	—	New Drugs Division Notification No. 9/99: Guidelines for Toxicity Studies of Drugs
Safety assessment of biotech therapeutics	S6 Preclinical Safety Evaluation of Biotechnology-Derived Pharmaceuticals[8]	Guidance for Industry: Providing Clinical Evidence of Effectiveness for Human Drug and Biological Products	Note for guidance on comparability of medicinal products containing biotechnology-derived proteins as a drug substance	—
Safety pharmacology	Guidance for Industry: S7A Safety Pharmacology Studies for Human Pharmaceuticals[5]	Guidance for Industry: S7A Safety Pharmacology Studies for Human Pharmaceuticals[4]	CPMP: Note for Guidance on Safety Pharmacology Studies in Medicinal Product Development[3]	Notification No. 4 – Guidelines for General Pharmacology[1]
QT interval	Safety Pharmacology Studies for Assessing the Potential for Delayed Ventricular Repolarization (QT Interval Prolongation) by Human Pharmaceuticals[9]	—	CPMP Points to consider: The assessment of the potential for QT interval prolongation by noncardiovascular medicinal products	—

Table 2.2 Cardiovascular System Safety Pharmacology Evaluations

Core
• Hemodynamics (blood pressure, heart rate)
• Autonomic function (cardiovascular challenge)
• Electrophysiology (EKG in dog)
QT prolongation (non-core)
An additional guideline, ICH S7B, is in preparation and 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*
• EKG (QT measurements) in a cardiovascular study that would be covered in the core battery
• HERG channel interactions (HERG expressed in HEK 293 cells)

Table 2.3 Respiratory System Safety Pharmacology Evaluation

Respiratory function	Measurement of rate and relative tidal volume, in conscious animals
Pulmonary function	Measurement of rate, tidal volume, and lung resistance and compliance in anaesthetized animals

The actual requirements of the final June 2001 ICH guidelines[5] are broadly sketched. They call for conducting studies in a core battery to assess effects on the cardiovascular (Table 2.2), respiratory (Table 2.3), central nervous system (CNS) (Table 2.4), and secondary organ systems (Table 2.5). Follow-up studies for the core battery are also required on a case-by-case basis for the three main organ systems, the functional compromise of which is considered immediately life threatening:

- CNS
  - Behavioral pharmacology, learning and memory, specific ligand binding, neurochemistry, visual, auditory, and electrophysiology examinations
- Cardiovascular system
  - Hemodynamics (blood pressure and heart rate), electrophysiology (EKG in dog and other such screens for QTc prolongation as appropriate), and autonomic function in response to a pharmacologic challenge
- Respiratory system
  - Tidal volume, bronchial resistance, compliance, pulmonary arterial pressure, blood gases

Table 2.4 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
Anaesthetic interactions
Test for central interaction with barbiturates
Anti/pro-convulsant 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

Note: Tier II or secondary organ system requirements are less precisely presented, as Table 2.5 makes clear.

[a] Usually a functional observational battery (FOB) is integrated into a rodent (rat) repeat dose toxicity studies to meet this requirement.

Table 2.5 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
Gastrointestinal (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

Studies are not necessary under certain conditions:

- 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 (refer to toxicology studies)
- New salts having similar pharmacokinetics and pharmacodynamics

The FDA has, meanwhile, published draft guidance[10] calling for safety pharmacology evaluation of all new pharmaceutical excipients.

References

1. Anon., Guidelines for the Safety Pharmacology Study Required for Application for Approval of Manufacturing (Import) of New Drugs: Notification No. 4, Takushin-Yaku: Japan Ministry of Health and Welfare, 29 January 1991.
2. Anon., Guidelines for general pharmacology, in: *Drug Approval and Licensing Procedures in Japan*, 137–140, Takushin-Yaku: Japanese Ministry of Health and Welfare, 1992.
3. CPMP, *Note for Guidance on Safety Pharmacology Studies in Medicinal Product Development*, 1998.
4. FDA, ICH S7A Guidance for Industry: Safety Pharmacology Studies for Human Pharmaceuticals. www.fda.gov/cder/guidance/4461fnl.pdf, 2001.
5. ICH, ICH S7A (Step 4) ICH Harmonized Tripartite Guidelines on Safety Pharmacology Studies for Human Pharmaceuticals. www.ich.org/pdfICH/S7step4.pdf, 2000.
6. ICH, ICH Harmonized Tripartite Guidelines (M3): Timing of Nonclinical Safety Studies for the Conduct of Human Clinical Trials for Pharmaceuticals. www.ich.org/pdfICH/m3mstep4.pdf, 1997a.
7. FDA, Guidance for Industry (M3): Nonclinical Safety Studies for the Conduct of Human Clinical Trials for Pharmaceuticals, www.fda.gov/cder/guidance/1855fnl.pdf, 1997.
8. ICH, ICH Harmonized Tripartite Guidelines (S6): Preclinical Safety Evaluation of Biotechnology-Derived Pharmaceuticals. www.fda.gov/cder/guidance/1859fnl.pdf, 1997b.
9. ICH, ICH S7B (Step 3) Safety Pharmacology Studies for Assessing the Potential for Delayed Ventricular Repolarization (QT Internal Prolongation), by Human Pharmaceuticals. www.ich.org/pdfICH/S7Bstep2.pdf, 2000.
10. FDA, Guidance for Industry: Nonclinical Studies for Development of Pharmaceutical Excipients. www.fda.gov/cder/guidance/3812dft.pdf, 2002.

Web Sources

FDA:	www.fda.gov
CPMP:	www.eudra.org
MHW:	www.mhw.go.jp/english
ICH:	www.ifpma.org

chapter 3

Principles of screening and study design

Introduction

In biological research, screens are tests designed and performed to identify agents or organisms that have 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 (and safety pharmacology, in particular), our use of screens is usually negative (i.e., no activity is found). Agents or objects possessing other than desired pharmacological or biochemical activities are not considered to present enough of a hazard due to these secondary pharmacologic effects, therefore, they are studied no further or developed as potential therapeutic agents without compelling reasons (e.g., cases of extreme benefit such as life-saving qualities). For safety pharmacology in particular, screens are required and intended.

In the broadest terms, everything done in preclinical and in Phase I clinical studies can be considered a form of screening.[1] This is certainly true of safety pharmacology studies. What varies is the degree of effectiveness or our confidence in each of the tests used. As a general rule, even though we think of the expensive and labor-intensive pivotal studies required to support regulatory requirements (4-week to 1-year toxicity studies, carcinogenicity, and segment I-III studies, etc.) as definitive, in fact, they are generally effective but not necessarily efficient screens.

Though toxicologists and pharmacologists 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 risks of harming humans does not exceed the potential benefits. In most cases, this means that if a test or screen identifies a level of risk in which we have confidence (our activity criterion, i.e., our basis for and degree of confidence in the outcome), 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. We are interested in minimizing the number of false negatives in safety assessment. Anderson and Hauck[2] should be consulted for statistical methods to minimize false-negative results.

Figure 3.1 illustrates how decisions are currently more likely to be made on a multidimensional basis, which creates a need for balance between degree of benefit, confidence that there *is* a benefit (efficacy is being evaluated in models or screens at the same time safety is), type of risk (with, e.g., muscle irritation, mutagenicity, acute lethality, and carcinogenicity having various degrees of concern attached to them), and confidence in, and degree of, risk.

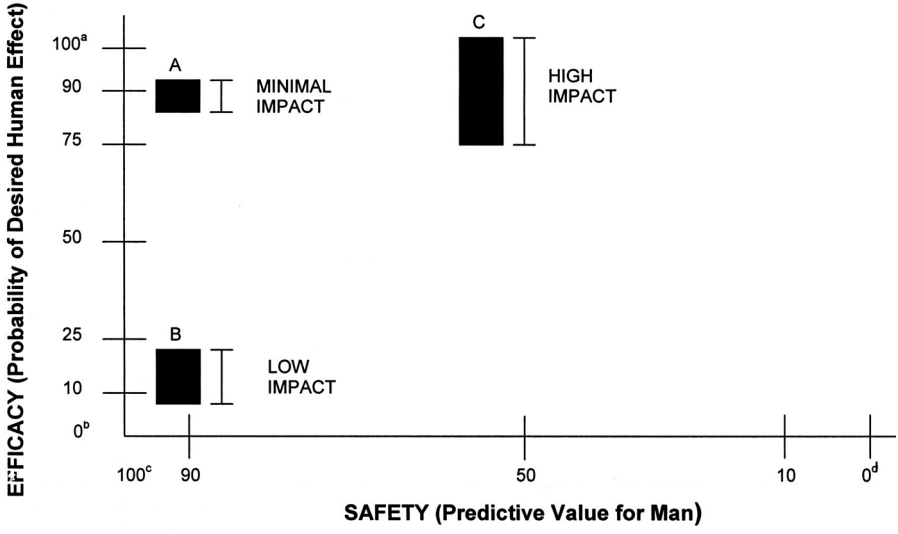


Figure 3.1 Decision making for pharmaceutical candidates based on outcome of screening tests. (a) A 100% probability of efficacy means that every compound that has the observed performance in the model(s) used has the desired activity in man. (b) A 0% probability of efficacy means that every compound 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 probablility of efficacy but also a high probability of having some adverse effect in man. But if that adverse effect is of low impact — for instance, transitory muscle irritation for a lifesaving antibiotic — "A" should go forward. Likewise, "B," which has low probability of efficacy and 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 50% probability of this finding (for instance, neurotoxicity or carcinogenicity) being predictive in man, the adverse effect is not an acceptable one. Here, a more definitive test is required or the compound should be dropped.

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. 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 endpoint. Rather, they confirm that if interest in a compound is sufficient, a more definitive test — a confirmatory test — is required, which frequently provides a basis for selecting between multiple candidate compounds.

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:

- *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)

Later, we will contrast some of these with key elements of study design.

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[3–5] that serve to determine power and define competing error rates:

α, the probability of our committing a type I error (a false positive)
β, the probability of our 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, because 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 much as to have greater statistical significance than all the other compounds.[6] 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 3.1 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 requires close collaboration between the toxicologist, chemist, and statistician. Screening, however, forms only part of a much larger research and development context. Screens may be considered the biological equivalent of exploratory data analysis (EDA). EDA methods, in fact, provide a number of useful possibilities for less rigid and yet utilitarian

Table 3.1 Discovery[a] and Discrimination[b] of Toxicants

	Actual activity of agent tested	
Screen outcome	**Positive**	**Negative**
Positive	a	b
Negative	c	d

[a]Discovery (sensitivity) = a/(a + c), where a = all toxicants found positive; a + c = all toxicants tested.

[b]Discrimination (specificity) = d/(b + d), where d = all nontoxicants found negative; b + d = all nontoxicants tested.

approaches to the statistical analysis of the data from screens, and are one of the alternative approaches presented and evaluated here.[7–10] 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.[11,12]

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 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 to 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 endpoints 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. The analysis techniques used here can then be used to ensure the quality or to modify the 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 (i.e., 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 about what they are seeking), 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.[13]

Uses of screens

The use of screens that first occurs to most pharmaceutical scientists is in pharmacology.[14] Early experiences with the biological effects of a new molecule are almost always in some form of efficacy or pharmacology screen. The earliest 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, central nervous system [CNS], etc.), can frequently provide information of use in safety assessment also — even if only to narrow the limits of doses to be evaluated. 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 QT interval. Although 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.[15,16]

Safety assessment screens are performed in three major settings: discovery support, development (generally considered the real job of safety assessment), and occupational health/environmental assessment testing. Discovery support is the most natural area for the employment of screens. This is 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 the 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 a number of compounds.

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

The use of screens in environmental assessment and occupational health is fairly straightforward. On the occupational side, the concerns discussed in Chapter 11 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, 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 endpoint 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 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 definitive study for carcinogenicity, but, in fact, it was originally designed and functions as a screen for strong carcinogens.[17] 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 that it is unlikely to be potent.

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

Types of screens

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

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* (negative) or equal to or greater than *X* (positive). As more data are accumulated, the criterion should be reassessed.

Sequential

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

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, 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.

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 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, inactive compounds 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. For example, all compounds with observed activities significantly greater than the control at the 5% level could be tagged. However, for early screens, given the power of the assay, a statistical criterion may be too strict 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 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 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. Many forms of graphical analysis methods are available, including some newer forms that are particularly well suited to multivariate data — the type 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.[18] Decreasing group sizes in 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 are acceptable noise levels in a screening system.

Criteria can be simple (presence or not of a pupil reflex) or as complex (alteration of an EKG) as required. The first step in establishing them should be an evaluation of the performance of test systems that have not been treated (negative controls). There will be some innate variability in the population, and understanding this variability is essential to selling some threshold for activity that has an acceptably low level of occurrence in a control population. Figure 3.2 illustrates this approach.

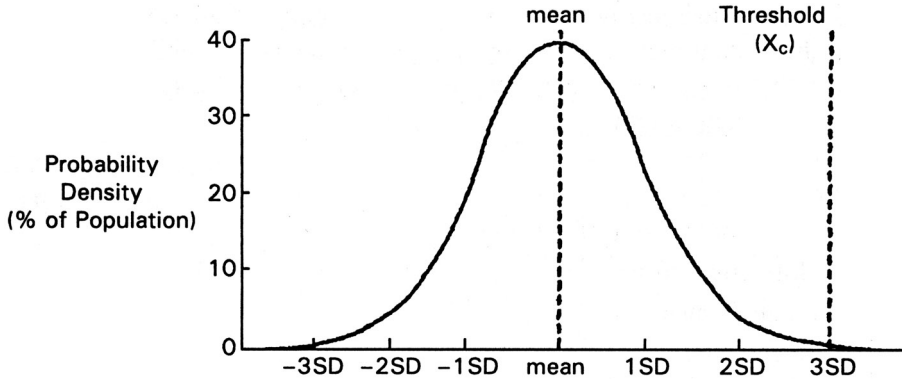


Figure 3.2 Setting thresholds using historical control data. This figure shows a Gaussian ("normal") distribution of screen parameters; 99.7% of the observations in the population are within three standard deviations (SD) of the historic mean. Here, the threshold (i.e., the point at which a datum is outside of "normal") was set at X_c = mean + 3 SD. Note that such a screen is one-sided.

Those endpoints 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 endpoint that is being predicted in man, such as gene mutations predicting 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 is not established. There is also, 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.

Analysis of screening data

Screening data presents a special case that, due to its inherent characteristics, is not well served by traditional approaches.[11,12,19,20]

First, consider which factors influence the power of a statistical test. Gad[11] established the basic factors that influence the statistical performance of any bioassay in terms of its sensitivity and error rates. Recently, Healy[21] presented a review of the factors that influence the power of a study (i.e., the ability to detect a dose-related effect when it actually exists). In brief, the power of a study depends on seven 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)

The following table lists 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.

Univariate data

Control charts

The control chart approach,[22] 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 the 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 control agent's performance (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 $\overline{X}$ of 400 g and a mean range $\overline{R}$ of 90 g. These values were used for the control chart (Figure 3.3). The subgroups are size five. The action limits for the mean and range charts were calculated as follows:

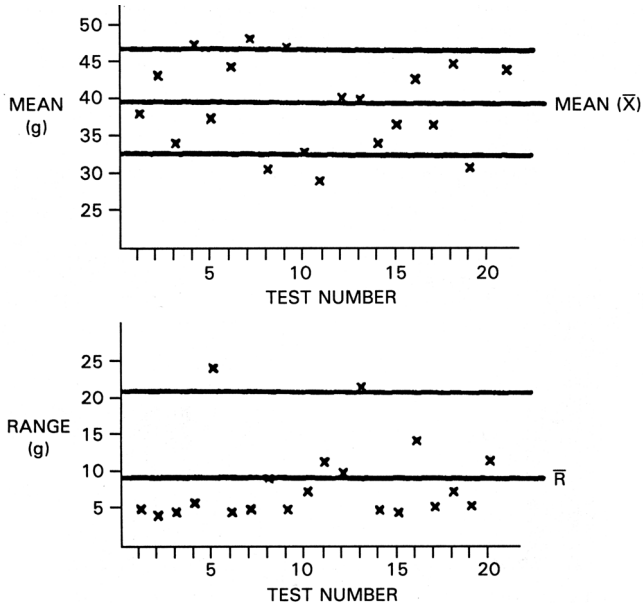


Figure 3.3 Example of a control chart used to "prescreen" data (actually, explore and identify influential variables) from a portion of a functional observational battery. See text discussion for explanation.

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

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 use them as rapid and efficient tools for detecting effects in groups being assessed for that same activity endpoint.

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).[23,24] 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 — for instance, 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 depicted in Figure 4.4. The data for the figure 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 3.4) 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), which allows 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 two or more occurred only 5% of the time by chance. So, we can simply perform an extreme value screen on our collapsed data

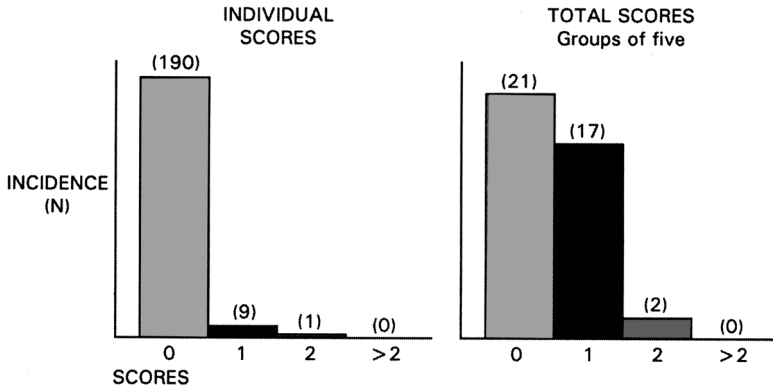


Figure 3.4 Plotting central tendency. Possible individual scores for righting reflexes may range from 0 to 8 (Gad, 1982a). Group total scores would thus range from 0 to 40. (Depicted is the number of groups that contain individual scores in the individual categories.)

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 that will not influence the outcome of the analysis are then identified and dropped from analysis. Group scores may then be simply evaluated against 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.

Multivariate data

The traditional acute, subchronic, and chronic toxicity studies performed in rodents and other species also can be considered to constitute multiple endpoint 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.[25]

The analog plot

The human eye is extremely good at comparing the size, shape, and color of pictorial symbols.[26–30] 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.[31]

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 to 84%) of a percentage variable. For example, for numbers x, symbol radium r, and plotting scaling factor k, a log mapping will give:

$x = 1$	$r = k$
$x = 10$	$r = 2k$
$x = 100$	$r = 3k$

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 or their significance.[32,33] 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 3.5 illustrates relationships for five measures on a time vs. 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 × 2 sexes × 5 groups × 53 times × 50 animals. Means alone constituted some 1600 4-digit numbers.

Body weight gains from the period immediately preceding each consumption measurement were used because these were less correlated. For each variable and at each time, the sums of squares for group differences were divided into four meaningful contrasts:

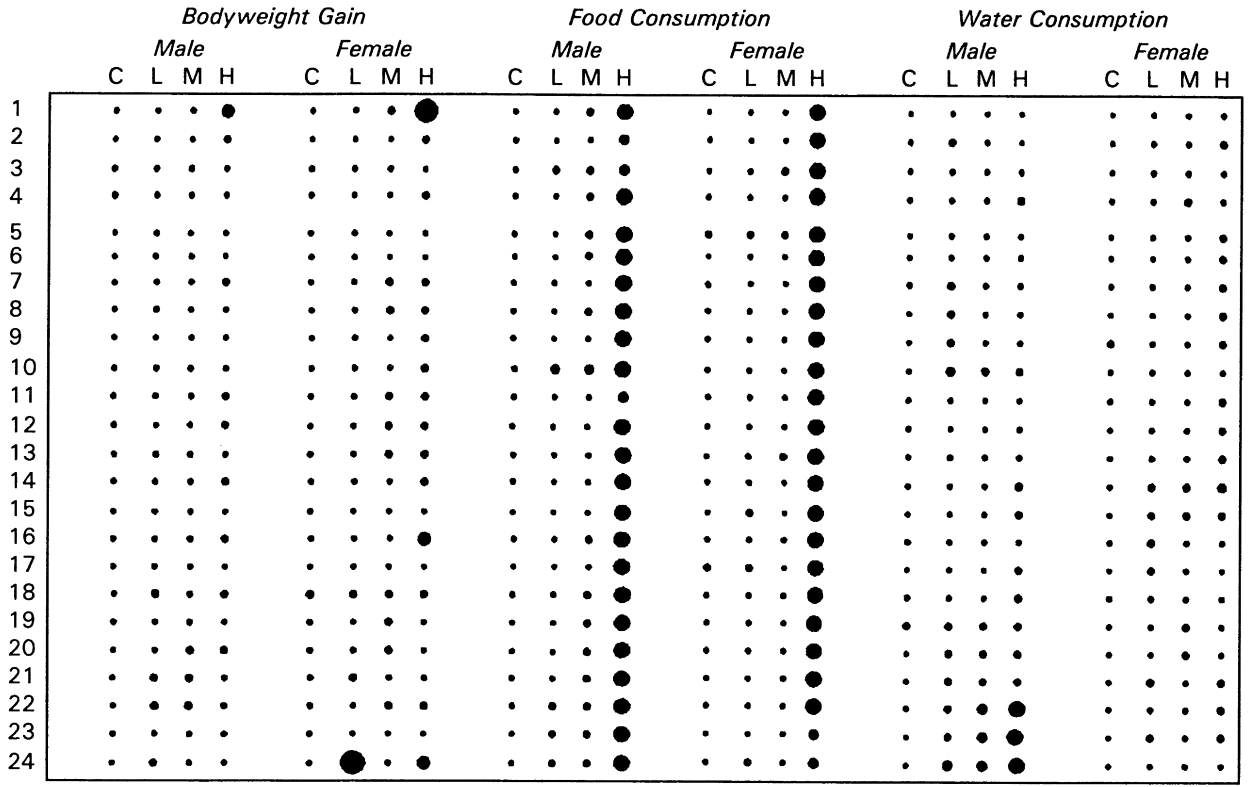
Control A vs. control B
Control A + B vs. low
Control A + B + low vs. medium
Control A + B + low + medium vs. high

To make the variables comparable, the sums of squares 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 appears 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.

Numerous approaches are available to solve the problem of capturing all the information in a set of multi-endpoint data. When the data are continuous in nature, approaches such as the analog plot can be used.[34,35] 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, other forms of analysis must be used. Just as the control chart can be adapted to analyze attribute data, an analog plot can be adapted. Other methods are also available.

Study design

Table 1.6 in Chapter 1 introduced the principal components of study design for which some guidance was given in International Conference on



Note: Square root relationship

Key: C = Control A vs B, L = low vs A + B, M = medium vs A + B + low, H = high vs A + B + low + medium.

Figure 3.5 Analog plot for dose-response contrasts. One of many possible approaches to graphically presenting multidimensional data. 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.

Harmonization (ICH) S7A. Some practical and more detailed guidance is offered next.

Animal models

Although the S7A suggests the use of conscious, unrestrained, telemeterized, or trained animals, this may not be practical in all instances. Some cardiovascular and respiratory measures may require or be best served by using animals that are restrained but acclimated.

Two equally important points are not addressed in the guidelines. The first is whether the animals employed in *in vivo* studies need to be naïve (implying that any animals be utilized for evaluating only one compound, and therefore significantly increasing animal usage and costs) or whether they could be used repeatedly for evaluations of different compounds (each in a range of doses covering a multiple of intended therapeutic dose levels, but not toxicological levels, with a suitable washout period between compounds). The latter is preferable unless there is reason to believe (generally based on knowledge of mechanisms or of actions of the class of compounds, as well as observations during a study) that the compound tested irreparably alters the animals employed in its evaluation.

The second question is whether sensitive, functionally compromised, or disease model animals should be employed for some of the safety pharmacology applications, particularly those of the respiratory and renal systems where the presence of significant functional reserves can serve to reduce the ability of test methods to detect changes. There are at least two good arguments for using sensitive models.[36] The first is that healthy animals and humans possess significant reserve capacities, which would preclude modest but significant effect being detected at doses in and reasonably near the therapeutic range as evidenced by the ability of individuals to lead quite normal lives with only single lungs or kidneys. The second is that drugs are actually used by sick individuals, and certainly the organ systems involved in the targeted disease claim for a drug will be compromised in those individuals using the drug in the market place. Therefore, animal models with similarly compromised functions would best serve to identify potential problems in those individuals using the drug therapeutically.

The major argument against such models is a conservative one: There is inadequate baseline information and experience with such models to be able to reliably evaluate results. This argument is particularly suspect when one considers that the animal models used to identify and verify therapeutic efficacy before going into man are just such disease or compromised models.

Group size

Experience should rule here as any formal statistical power analysis would require assumptions as to what a meaningful change would be and as to

what the variability of a group of animals — fairly well controlled in normal rats, but much less so in dogs or telemeterized and or trained animals — on study will be.

Generally, groups of four rodents or four nonrodents per data point should be adequate, except in the case of a free-standing Irwin screen or functional observatory battery (FOB) where many endpoints are assessed in the same animal with high variability associated with individual endpoints. Here, the animals of one sex would be advisable.

Statistical design

A randomized block design should be the standard here, ensuring adequate washout times are allowed between treatments.

Dose levels/test concentrations

- Should define the dose- (concentration-) response curve.
- Time course should be investigated when feasible.
- Doses should include and exceed the primary pharmacodynamic or therapeutic range. In the absence of adverse effects on safety pharmacology parameters, the highest tested dose should produce moderate adverse effects in this or other studies.
- Some effects in the toxic range (e.g., tremors during EKG recording) may confound the interpretation of safety pharmacology effects and may also limit dose levels.

References

1. Zbinden, G., Elsner, J., and Boelsterli, U.A., Toxicological Screening, *Regul. Toxicol. Pharmacol.,* 4:275–286, 1984.
2. Anderson, S. and Hauck, W., A new procedure for testing equivalence in comparative bioavailability and other clinical trials, *Commun. Statist. Thero. Meth.*, 1983, 12:2663–2692, 1983.
3. Gad, S.C., A neuromuscular screen for use in industrial toxicology, *J. Toxicol. Environ. Health*, *9*:691–704, 1982a.
4. Gad, S.C., Statistical analysis of behavioral toxicology data and studies, *Arch. Toxicol.* (Suppl.), 5:256–266, 1982b.
5. Gad, S.C., *Statistics and Experimental Design for Toxicologists*, 3rd ed., Boca Raton, FL: CRC Press, 1999.
6. Bergman, S.W. and Gittins, J.C., Screening procedures for discovering active compounds, in *Statistical Methods for Pharmaceutical Research Planning,* Peace, K., Ed., New York: Marcel Dekker, 1985.
7. Tukey, J.W., *Exploratory Data Analysis,* Reading, PA, Addison-Wesley: 1977.
8. Redman, C., Screening compounds for clinically active drugs, *Statistics in the Pharmaceutical Industry* in Buncher, C.R. and Tsya, J., Eds., New York: Marcel Dekker, 19–42, 1981.

9. Hoaglin, D.C., Mosteller, F., and Tukey, J.W., *Understanding Robust and Exploratory Data Analysis,* New York: John Wiley & Sons, 1983.
10. Hoaglin, D.C., Mosteller, FD., and Tukey, J.W., *Exploring Data Tables, Trends, and Shapes,* New York: John Wiley & Sons, 1985.
11. Gad, S.C., An approach to the design and analysis of screening studies in toxicology, *J. Am. Coll. Toxicol.,* 8:127–138, 1988.
12. Gad, S.C., Principles of screening in toxicology with special emphasis on applications to neurotoxicology, *J. Am. Coll. Toxicol.,* 8:21–27, 1989a.
13. Garrett, R.A. and London, J.P., *Fundamentals of Naval Operations Analysis,* Annapolis: U.S. Naval Institute, 1970.
14. Martin, Y.C., Kutter, E., and Austel, V., *Modern Drug Research,* New York: Marcel Dekker, 31–34, 155, 265–269, 314–318, 1988.
15. Pennie, W.D., Use of cDNA microassays to probe and understand the toxicological consequences of altered gene expression, *Toxicol. Lett.*, 112–113:473–477, 2000.
16. Nuwaysir, E.F. et al., Microassays and toxicology: the advent of toxicogenomics, *Mol. Carcinog.,* 24:153–159, 1999.
17. Page, N.P., Concepts of a bioassay program in environmental carcinogenesis, in *Environmental Cancer,* Kraybill, H.F. and Mehlman, M.A. Eds., New York: Hemisphere Publishing, 87–171, 1977.
18. Bickis, M.G., Experimental design, in *Handbook of In Vivo Toxicity Testing,* Arnold, D. L., Grice, H.C., and Krewski, D.R., Eds., San Diego: Academic Press, 128–134, 1990.
19. Gad, S.C., Screens in neurotoxicity: objectives, design and analysis, with the observational battery as a case example, *J. Am. Coll. Toxicol.,* 8:1–18, 1989b.
20. Gad, S.C., Statistical analysis of screening studies in toxicology with special emphasis on neurotoxicology, *J. Am. Coll. Toxicol.,* 8:171–183, 1989c.
21. Healy, G.F., Power calculations in toxicology, *A.T.L.A.,* 15:132–139, 1987.
22. Montgomery, D.C., *Introduction to Statistical Quality Control,* New York: John Wiley & Sons, 1985.
23. Velleman, P.F. and Hoaglin, D.C., *Applications, Basics and Computing of Exploratory Data Analysis*, Boston: Duxbury Press, 1981.
24. Tufte, E.R., *The Visual Display of Quantitative Information*, Cheshire, CT: Graphic Press, 1983.
25. Young, F.W., Multidimensional scaling, in *Encyclopedia of Statistical Sciences,* Vol. 5, Katz, S. and Johnson, N.L., Eds., New York: John Wiley & Sons, 649–659, 1985.
26. Anderson, E., A semigraphical method for the analysis of complex problems, *Technometrics,* 2:387–391, 1960.
27. Andrews, D.F., Plots of high-dimensional data, *Biometrics,* 28:125–136, 1972.
28. Davison, M.L., *Multidimensional Scaling,* New York: John Wiley & Sons, 1983.
29. Schmid, C.F., *Statistical Graphics*, New York: John Wiley & Sons, 1983.
30. Cleveland, W.S. and McGill, R., Graphical perception and graphical methods for analyzing scientific data, *Science,* 229:828–833, 1985.
31. Wilk, M.B., and Gnanadesikan, R., Probability plotting methods for the analysis of data, *Biometrics,* 55:1–17, 1986.
32. Kruskal, J.B., Multidimensional scaling by optimizing goodness of fit to a nonmetric hypothesis, *Psychometrika,* 29:1–27, 1964.
33. Kass, G.V., An exploratory technique for investigating large quantities of categorical data. *Appl. Stat.,* 29:119–127, 1980.

34. Chernoff, H., The use of faces to represent points in K-dimensional space graphically, *J. Am. Stat. Assoc.*, 68:361–368, 1973.
35. Chambers, J.M. et al. *Graphical Methods for Data Analysis*, Boston: Duxbury Press, 1983.
36. Milano, S., Introducing disease animal models in safety pharmacology: a way to reevaluate the risk, *MDS Safety Pharmacology Symp.*, Lyon, France: MDA Pharma, 5–6 December 2002.

chapter 4

Cardiovascular system

Introduction

The cardiovascular system is a primary vital function that must 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. 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.[1–3] In the basic program, a detailed hemodynamic evaluation is carried out in the anaesthetized 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.

History

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 known as *torsades de pointes* (TdP), and death.[4–8] Many of the drugs now known to alter ventricular repolarization were developed as antiarrhythmics, such as dofetelide and sotalol, but others, such as cisapride and terfenadine, were developed without the expectation of any effect upon electrically excitable membranes. This has lead to International Conference on Harmonization (ICH) promulgating ICH S7B[9] with specific guidance for evaluation.

Prolongation of the cardiac action potential duration (APD), reflected in the electrocardiogram (EKG) as QT interval prolongation (QT prolongation), is associated with potentially lethal polymorphic ventricular tachyarrhythmia, or TdP.[10–12] QT prolongation may occur spontaneously in

Table 4.1 Drugs Causing QT Prolongation and Torsades de Points

Agent	Drugs
Class IA anitarrhythmic agents	Quinidine, disopyramide, procainamide
Class III antiarrhythmic agents	Amiodarone, dofetilide, D-sotalol, dibutilide
Calcium antagonists	Bepridil, terodiline
Antihypertensives	Ketanserin (adrenocepto antagonist)
Antidepressant agents	Amitryptiline, citalopram, clomipramine, desipramine, doxepin, imipramine, maproptiline, nortriptiline, zimelidine
Anitfungal agents	Fluconazole, itraconazole, ketoconazole, micoconazole
Antihistamine agents	Astemizole, dephenhydramine, hydroxizine, terfenadine
Anticancer agents	Amsacrine, doxorubicine, zorubicine
Antimicrobial agents	Amantidine, amphotericin, clarithromycin, clindamycin, cotrimoxazole, erythromycin, grepafloxacin, pentamidine, sparfloxacin, spiramycin, trimethoprin-sulphamethoxazole, troleandomycin
Antimalarial agents	Chloroquine, halofantrine, quinine
Antipsychotic agents	Chlorpromazine, haloperidol, flufenazine, lithium, mesoridazine, pimozide, prochlorperazine, roperidol, sultopride, sertindole, risperidone, thioridazine, trifluoperazine
Miscellaneous agents	Cisapride, probucol, indapamide, K^+ wasting diuretics (e.g., furesemide)

Source: From Altman, E.M., *Cardiovascular Therapeutics*, 2nd ed., Philadelphia: W.B. Saunders Company, 2002; and Hardman, J.C. and Limbird, L.E., *Goodman & Gilman's The Pharmacological Basis of Therapeutics*, 9th ed., New York:, McGraw-Hill, 1996.

subjects having mutations in genes encoding for potassium or sodium channels regulating normal cardiac repolarization (congenital long QT syndromes, or LQTS).[13,14] By the mid-1990s, QT prolongation was also associated with a growing number of drugs, antiarrhythmic as well as non-antiarrhythmic, and drug combinations (acquired LQTS). Many, if not most, of these drugs prolong QT by inhibiting the rapid component of the delayed rectifying potassium current, I_{Kr}; the α-subunit through which this current flows is encoded by the human ether-a-go-go gene (HERG).[15,16] The likelihood of developing TdP is considered to be related to the magnitude of prolongation of ventricular repolarization.[17] Nonantiarrhythmic drugs that affect QT at therapeutic doses typically show less than 10 to 15 msec mean QT prolongation (Table 4.1). The frequency with which these drugs evoke TdP is very low, estimated in the range of 1 in 2000 to 20,000+ patients, dependent on other risk factors in the treated population; hence, treatment-evoked TdP is unlikely to be detected within the context of a

Table 4.2 Cardiovascular System Safety Pharmacology Evaluations

Core

- Hemodynamics (blood pressure, heart rate)
- Autonomic function (cardiovascular challenge)
- Electrophysiology (EKG in dog)

QT prolongation (non-core)

An additional guideline, ICH S7B, is in preparation and 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*
- EKG (QT measurements) in a cardiovascular study that would be covered in the core battery
- HERG channel interactions (HERG expressed in HEK 293 cells)

traditional clinical development program. Thus, there is an imperative to develop techniques and strategies to detect potential for TdP hazard early in development, ideally, prior to first exposure in man (FIM).

Regulatory developments

The first regulatory body to issue comments on specific concerns related to the potential for drugs to induce QT interval prolongation was the Committee for Proprietary Medicinal Products (CPMP) (See Table 4.2 for the assessment of the potential for QT interval prolongation by noncardiovascular medicinal products in 1997). CPMP identified specific methodologies for both preclinical and clinical studies, including stimulation frequencies for *in vitro* preparations, numbers of EKG leads and chart speed, and QT interval correction (QTc) strategies for heart rate changes.

During the last few years, the discussion has evolved as other regulatory bodies have provided guidance (Table 4.3). This debate occurred in parallel with the ICH process to rationalize safety pharmacology studies for pharmaceuticals. The S7A and S7B guides describe in detail study design consideration for these evaluations, including a strong preference for *in vivo* EKG studies conducted in conscious, unstressed animal preparations. Arguably, the QT prolongation issue made safety pharmacology a regulatory subdiscipline during 2000 and 2001.

The objective of preclinical cardiovascular safety testing is to identify either EKG signals of TdP hazard or a change in heart rate (pulse) and blood pressure indicators for purposes of evaluating the potential risk to the first humans to be exposed to a new drug, and for stipulating additional precautions and safeguards to protect those individuals while the effects of a new drug on human ventricular repolarization are being established. The preclinical testing scheme is well developed in ICH S7B, with general aspects of study design referenced to ICH S7A. The assessment of risks with the

Table 4.3 Current International Guidelines and Draft Documents on Preclinical Assessment of TdP Hazard

Document	Date	Comment
Notes on Applications for Approval to Manufacture (Import) New Drugs, issued in 1975 (MHW-Japan).	1975	Requested evaluation of: "… and effects of the test substance on the … central nervous system, peripheral nervous system, sensory organs, respiratory and cardiovascular systems, smooth muscles including uterus, peripheralorgans, … renal function, … and adverse effects observed in clinical studies" [p.71].
Japanese Guidelines for Nonclinical Studies of Drugs Manual 1995, Yakuji Nippo, Limited, Tokyo, 1995.	1995	Studies in Lists "A" and "B" became de facto international blueprints for general/safety pharmacology evaluations until issue of ICH S7A. "*Normally, anesthetized animals are used.*"; thus, anesthetized animal preparation (and particularly the barbiturate-anesthetized dog) became the standard for cardiovascular-respiratory evaluations [p.128].
Committee for Proprietary Medicinal Products (CPMP) (EU). Points to Consider: The assessment of the potential for QT interval prolongation by non-cardiovascular medicinal products (www.emea.eu.int/ pdfs/human/swp/09869en.pdf)	1997	First regulatory document addressing the TdP hazard with pharmaceuticals; credited with generating academic and inter-industrial cooperation to share existing data and to generate collaborative efforts to rapidly produce realistic experimental and clinical approaches for identification of preclinical signals a TdP hazard.
Guidance for Industry. S7A Safety Pharmacology Studies for Human Pharmaceuticals. *U.S. Department of Health and Human Services, Food and Drug Administration, Center for Drug Evaluation and Research, Center for Biologics Evaluation and Research, ICH, July, 2001.* (www.ifpma.org/ich5s.html)	2001	Provides the general study design framework for in vitro and in vivo preclinical evaluations of TdP. Specifically places evaluations addressing risk for repolarization-associated ventricular tachy-arrhythmia within the safety pharmacology domain.

Therapeutic Products Directorate Guidance Document (Canada). Assessment of the QT prolongation potential of non-antiarrhythmic drugs (2001). (www.hc-sc.gc.ca/hpb-dgps/therapeut/htmleng/guidmain.html)	2001	States explicitly that development of a non-antiarrhythmic drug with a preclinical signal (*in vitro* or *in vivo*) of TdP hazard "should be pursued only if it is expected to provide a major benefit for a serious disease or disorder for which safer alternatives are not available, or if the cardiotoxicity is attributable to a metabolite generated in animals, but not in humans."
ICH Guideline on Safety Pharmacology Studies for Assessing the Potential for Delayed Ventricular Repolarization (QT Interval Prolongation) by Human Pharmaceuticals (S7B), Step 2, 2002.	2002	Presents a tiered testing scheme recommending *in vitro* ion current and repolarization evaluations, and an *in vivo* QT assessment in a appropriate species; provides a current assessment of the pros and cons of available techniques while recognizing that this area is in great flux and recommending that new technologies be evaluated and applied as they become available.
The clinical evaluation of QT/QTc interval prolongation and proarrhythmic potential for non-antiarrhythmic drugs. FDA preliminary concept paper, Nov. 15, 2002.	2002	Accepts ICH S7B guidelines, when finalized, fo preclinical assessment of potential TdP hazard; provides a starting point for discussion as to how address preclinical signals within subsequent clinical development.

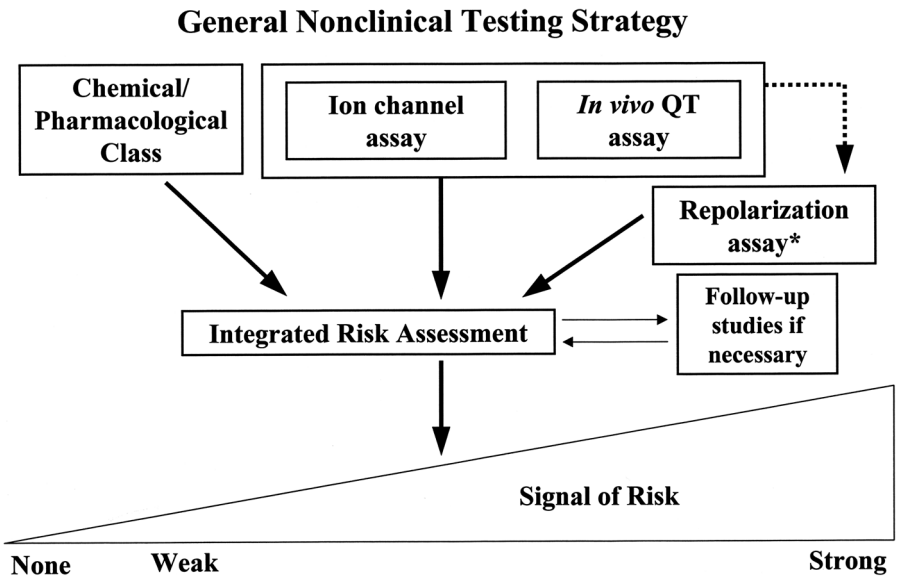


Figure 4.1 General nonclinical testing strategy.

data generated is not so well delineated. An integrative, tiered physiological approach is currently recommended with evaluations directed at the subcellular, cellular, tissue, and intact organism levels (Figure 4.1). Compounds can be examined for effects on a target molecule by evaluating effects on specific ion currents in heterologous expression systems, such as HERG stability expressed in a mammalian cell line. HERG is a logical target molecule; most drugs that prolong QT interval and elicit TdP in humans also block the HERG-related I_{Kr} current. However, HERG is not the only target potentially affecting ventricular repolarization; alteration in the function of IKs (KVLQT1 and minK) or Ina (hH1) may potentially also produce repolarization abnormalities that would predispose toward TdP, though no drug moieties have actually been so identified. Intact ventricular myocyte, isolated tissues (Purkinje fibers) and isolated hearts represent the successive levels of integration; cellular systems allow the study of effects of a compound on all currents during the action potential in aggregate; tissue and organ systems permit addition of cell–cell interactions and interstitial environmental effects. Finally, studies using intact (preferably conscious, unstressed) animals permit evaluation of pharmacodynamic and tissue distribution effects, homeostatic compensatory effects, and effects of metabolites. Surface electrocardiography results in intact animals also provide the bridge into the initial clinical EKG program by providing direct assessments of effects on QT interval and other aspects of EKG morphology (T and U waves, arrhythmias). Hence, it is critical that the intact animal study designs model the intended FIM exposure. In this regard, although there is intense interest in preclinical arrhythmia models with which to evaluate directly treatment-evoked

arrhythmia potential, there is currently no consensus that any of the available models are relevant for predicting TdP hazard in humans.

Remarkable progress has been made toward a rational, consensus approach for preclinical evaluation of TdP hazard. Additional effort is necessary to achieve consensus on specific preclinical study design aspects of the *in vitro* evaluations in heterologous expression systems, isolated cells, tissues, and organs as differences in stimulation frequencies, bathing media, etc., and to produce both qualitative and quantitative result differences. Because the exposures in these isolated systems may approach the solubility of test substances and be many fold greater than therapeutic exposures, guidance as to appropriate margins is still needed. Recent pronouncements for Food and Drug Administration (FDA) figures and experts[20] suggest that weak or negative early findings of QT internal prolongation will allow loss of a monitoring burden in later clinical development. There are several limitations of *in vitro* studies of putative target molecules. A heterologously expressed channel is not in its native cellular environment; this may affect basal function and, even more insidiously, may selectively impact drug action. Such studies do not permit the examination of drug effects in diseased or compromised tissues, nor do they permit assessment of effects peculiar to chronic exposure (e.g., up-regulation or down-regulation). The effects of major human metabolites of the parent drug can only be performed if the metabolites are known. In studies of a single target channel (e.g., HERG), effects of the agent on other ion targets will always be missed. The disadvantages of native cell studies include increased expense and time. The process of isolating cells may change the function of the channel itself and possibly the response to the drug. Microenvironments of cells in artificial culture medium may be different from those in intact tissues and magnify or mask responses. Typically, available human cells have been isolated from diseased hearts that have been exposed to elaborate, uncontrolled drug regimens further complicating the interpretation of the results. Rats and probably mice are known to lack the I_{Kr} currently thought to be responsible for much of the acquired LQTS in humans; hence it is recommended that EKG evaluations not be conducted in these species. On the other hand, anecdotal reports of compound-evoked TdP in rats cause concern that ion currents/combinations other than I_{Kr} can also produce TdP, and suggest a role for rodents EKGs in safety evaluation.

In intact animals, much attention has been paid to appropriate strategies to correct animal QT intervals for underlying changes in heart rate.[21] In an ideal setting, baseline and postdrug EKGs would be collected at identical heart rates; practically, it is nearly impossible to ensure similar heart rates when only small amounts of EKG data (< 1 min) can be collected during each study interval. Bazett recognized this problem in the early 1920s and modeled human QT and RR interval data to produce a QT value corrected to 60 bpm (QTc Bazett).[22] However, Bazett himself recognized the weaknesses of a mathematically modeled, pseudo-physiological parameter and recommended against the widespread use of his QTc formula. QTc Bazett is well

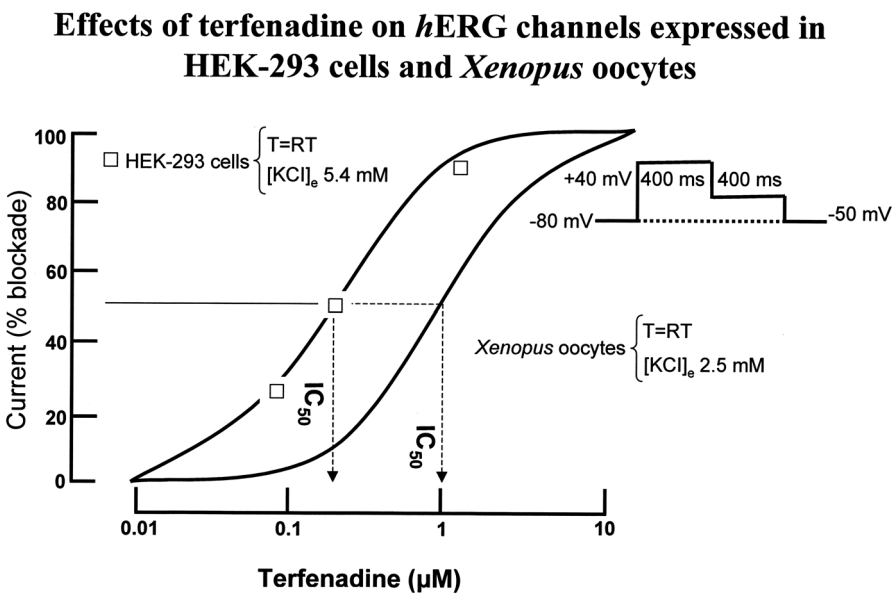


Figure 4.2 Effects of terfenadine on HERG channels expressed in HEK-293 cells and *Xenopus* oocytes. (Modified from Lacerda, A.E. et al., *Eur. Heart J.*, (Suppl. K):K23, 2001. With permission.)

known not to model QT–RR relationships usefully in dogs, and its use has led to misidentification of cardiotoxic effects of drugs, especially those that produce direct or indirect effects on heart rate, in this species.[23,24] With automated data collection and wireless EKG telemetry systems, it is possible to collect vast amounts of EKG data from experimental animals and to develop baseline and postdrug QT–RR relationships in individual animals, obviating the need for QT correction and its inherent assumptions.

In the preceding, several strategies are being considered to generate physiological heart rate ranges for QT–RR relationships. The simplest is to monitor EKGs over 24-h periods taking advantage of circadian heart rate variation.[25]

No single preclinical effect (marker) is yet considered reliable to predict TdP hazard in humans. A blocking effect at the level of HERG *in vitro* does not necessarily translate to an effect on APD or QT in the same, let alone another, species and under other experimental or clinical conditions. Similarly, an effect on APD and QT in animals does not necessarily accurately predict TdP in humans due to species differences in cardiac depolarization/repolarization and in metabolite formation;[17] and the lack of ion current, ADP, or QT signals from preclinical testing does not predict that a compound will be free from such effects in humans. Thus, an integration of signals of unwanted effects in the chain of events from blockade of specific ion currents to generation of actual arrhythmias must be continuously evaluated across the spectrum of preclinical designs

throughout clinical development in humans and in the postdevelopment drug surveillance.

Ultimately, the human risk of TdP is identified through careful assessment of ventricular repolarization and adverse events in relevant clinical populations.[26] Current draft guides state that data on potential QT changes must accompany the safety summary of the regulatory submission from (in aggregate) at least 100 volunteers — preferably including both males and females — on relevant doses of drug and a suitable number of placebo.

But with growing frequency (in part related to the new testing requirements), preclinical signals are being detected from the required testing (particularly the HERG) for almost all moieties at some concentration. Such signals provoke decisions as to (1) additional preclinical investigation, (2) additional clinical safety monitoring and investigations of treatment-evoked TdP hazard in humans, and (3) to continue, delay, or even discontinue development. Thus, for preclinical signals of TdP risk, the following should be considered:

- *Detection and evaluation of signals* — Efforts continue to better identify and understand the significance of signals from *in vitro* and *in vivo* studies; hypothetically, two categories of signals and associated levels of concern are defined:[27]
  - A *strong signal* is an observation for which a current majority (consensus) opinion supports a linkage to a possible TdP risk. A strong signal is a definite cause for concern. Examples of strong signals could include treatment-evoked ventricular arrhythmias (except occasional supraventricular or ventricular actopics) in animals or humans, or association with a chemical or pharmacological class of agents known or suspected to pose a risk of TdP.
  - A *weak signal* is an observation supporting a linkage to a possible TdP risk, but for which there is no consensus opinion. A weak signal is a possible cause for concern. Examples of weak signals could include:
    - An average increase in QT interval (compared with baseline or placebo) outside the normal range
    - Significant change from baseline in rate-corrected QT interval (QTc, using an species-appropriate method)
    - Treatment associated QT and/or QTc absolute values outside of normal ranges identified for specific species
    - Signals from any individual preclinical *in vitro* (HERG channel, action potential duration, etc.) or *in vivo* (EKG evaluation in animals or a pathophysiologic animal model), with the exception of certain treatment-evoked arrhythmias, constitute weak signals because of current uncertainties in extrapolation of nonclinical findings to humans

It is conceivable that two or more weak signals could, in aggregate, constitute a strong signal of TdP risk based upon the consensus criteria. Alternatively, if neither the preclinical testing nor the early clinical testing shows any electrophysiological effects related to delayed repolarization (e.g., signals), the likelihood of the new active substance showing important proarrhythmic effect during its clinical use, is considered remote. In evaluating signals of TdP risk, consideration is given to:

- *Magnitude* — Is the magnitude sufficient to be discerned from the background variability of the model being used?
- *Dose-response* — Are dose- or concentration-response relationships apparent in the data? Evidence of dose-response can help distinguish treatment-evoked effects from experimental variation.
- *Reversibility and reproducibility* — Is the effect reversible upon removal of exposure; can the effect be reproduced following re-exposure (e.g., treatment-evoked)?
- *Therapeutic index (TI)* — What is the ratio of the exposure (dose or concentration) where a signal is first apparent (lowest effect level, LOEL), compared with a projection or measured efficacious unbound plasma concentration (ED90, EC90) associated with the dosing regimen? Although ability to extrapolate nonclinical effects to humans is incomplete, a large TI may provide a basis for reducing concern whereas a TI ≤ 1 may increase concern about a particular signal. Several groups claim to be developing data supporting validation of TI calculations based upon signals generated in HERG expression systems and human TdP risk.

  Another form of this is the calculation of a CSI (cardiac safety index), or ratio between the level where there is clear therapeutic effects (ED90) and that at which there is an indication of minimal HERG activity (such as an EC10). A CSI of 30 is generally considered an absolutely minimum safety margin. A CSI of 100 is a more common margin.
- *Tissue distribution* — Is the drug accumulated or sequestered in cardiac tissues? Drugs that accumulate in cardiac tissue may impart risk that is not reflected by either total or unbound circulating drug concentrations. Drugs that prolong ventricular repolarization in a fashion unrelated to its pharmacokinetic profile or appear at the nadir of drug concentrations should be evaluated for tissue accumulation or for the presence of long-lived metabolites.
- *Metabolites* — With regard to major metabolites, it is important to distinguish whether these are human-specific, or present in sufficient concentrations in human to pose an arrhythmia risk.
- *Species specificity* — Is a signal demonstrable in two or more species? The absence of a signal in multiple species at similar

exposures may lessen concern that a particular signal will be observed in man, or indicate the presence of a species-specific active metabolite.

- *Direct and indirect effects on QT interval* —When QT or QTc interval prolongation or alterations in T-wave morphology are noted in a multiple nonclinical dose study, consideration should be given to whether these findings are a result of direct effects on cardiac conduction or an indirect effect related to toxicity. Consideration should be given as to whether cardiac lesion or alterations in clinical chemistry parameters (e.g., hypokalemia, hypomagnesemia, hypocalcemia). are concurrent with the changes in ventricular repolarization. This distinction is important in recognizing the primary toxicity produced by a drug and monitoring for this event in clinical investigations (e.g., the appropriate choice of biomarker).

In the absence of a validated marker of TdP hazard, differentiation of strong and weak signals of TdP risk is significant due to impact on the clinical program. If a weak signal is present, the clinical program is spurred to fully explore the potential for TdP risk in volunteers and patients with additional risk factors, while perhaps initially increasing the intensity of safety monitoring procedures. The primary objective of these clinical evaluations is relatively straightforward: detect/define the extent of treatment-evoked prolongation of ventricular repolarization, if any, in the relevant patient populations. However, if a strong signal is present, the clinical program must:

- Increase the intensity of safety monitoring
- Define the relationship between the signal and the TdP risk in the intended patient population
- Possibly elect to restrict study entry criteria to exclude patients with additional risk factors (e.g., long basal QT)

The clinical challenge is daunting: to prove that the frequency of an already low-probability adverse event (TdP) is not unduly increased in patients exposed to treatment. Only intensive clinical investigation can reduce TdP concern generated by preclinical signals. It is imperative that this be done prior to initiation of the pivotal efficacy (Phase 3) program, so that safety monitoring and entry criteria are the least restrictive, consistent with patient safety and good clinical practice.

Preclinical aspects of detection and evaluation of signals of TdP hazard have evolved rapidly over the past 10 years and are now codified in the ICH's Safety Pharmacology guidance (S7A and S7B). The latter is a unique document that envisions additional refinements in strategy and technology and encourages their use; however, operationally, it is not clear how tripartite consensus will be maintained as preclinical testing evolves with time in areas not specifically addressed in S7B, so there is no regulatory precedent.

Progress on preclinical markers of TdP is also illustrative of the challenge to develop specific biomarkers of efficacy and toxicity endpoints identified in animals and validated for drugs entering clinical development. Given that knowledge developed over the past century on the physiology and pathophysiology of the heart beat and the intense academic and industrial cooperation since CPMP 1997 (Table 4.3), it is sobering that as yet no single biomarker of TdP hazard has been established from preclinical testing that is uniformly viewed as a predictive surrogate for TdP hazard in humans. QT prolongation appears to be a mechanistically understood candidate marker, common to and easily monitored in both animals and in humans, but the debate continues as to thresholds of prolongation past which the risk of TdP is actually increased in either animals or humans. Rather, there is an expanding list of signals of potential TdP hazard emanating from both preclinical and clinical studies that must be integrated and interpreted for decision-making. The late Dr. Gerhard Zbinden, an early champion of organ function as toxicological targets during an era dominated by morphologic toxicology, is vindicated; today's ideal nonantiarrhythnic drugs should not influence IKr, APD, and QT, which are all organ function targets.

The association between abnormalities of repolarization and life-threatening arrhythmias is stronger than 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,* and each of the congenital long-QT syndromes is associated with an elevated incidence of malignant arrythmias.

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, and so on) and to serum levels of the drug and/or its metabolites. The actual incidence of malignant arrhythmias, even in association with the drugs most commonly 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 alternations in the channels of

* Some QT-prolonging drugs (e.g., amiodarone[28]), 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.

ionic flux through cell membranes. Some cells (e.g., those of the Purkinje system or mid-myocardium) 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, alteration in the propagation of action potentials within a given cell, and alterations in the propagation of action potentials from cell to cell within syncitia 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:

- *In vitro* screening of the drug and its metabolites for effects on ion channels (especially the I_{Kr} calcium channel)
- *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

Some specific techniques that can be employed are described next.

Patch-clamp studies using recombinant cells expressing HERG channels

Most pharmaceuticals associated with TdP inhibit rapidly delayed rectifier current, I_{Kr}. Therefore, particular attention to assays for I_{Kr} is prudent for assessing risk of QT interval prolongation.[29]

Using the voltage-clamp technique, outward or inward ionic currents can be measured from single cell preparations. Because of inherent difficulties associated with recording I_{Kr} in native myocytes, much of the available data for this current has been obtained using recombinant cell lines expressing HERG.

Inhibition of other outward (repolarizing) currents (e.g., I_{to}, I_{Kl}, I_{Ks}) or increase in inward (depolarizing) ionic currents (e.g., I_{Na}) could also lead to QT interval prolongation, and, therefore, should be considered when investigating the mechanisms for QT interval prolongation.[30–32]

HERG protein expression system

When transfected with HERG alone or HERG in association with genes for potential regulating subunits (e.g., MiRP1), appropriate cell systems express a K^+ channel that displays biophysical and pharmacological properties similar to I_{Kr}. Several expression systems have been used to test the activity of test substances on HERG current.

Table 4.4 IC_{50} of Several Drugs against HERG Channel (HEK-293) Conveyed Current Determined at 22°C and 35°C

Drug	IC_{50} (μM) at 22°C	IC_{50} (μM) at 35°C
Terfanadine	0.02	0.3
Loratadine	4.7	7.9
Dephenidramine	3.9	2.7
Fexofenadine	11	13.6
Cisapride	0.03	0.04
Sotalol	810	270
Erythromycin	3158	200

The cell line should be selected based upon appropriate levels of HERG expression.

The cell line should be tested at appropriate intervals to confirm the stability of the current from the expressed channel.

With the presence of endogenous currents, the presence or absence of subunits (directly associated regulatory proteins), and kinases or phosphatases controlling regulatory phosphorylation sites can affect the pharmacology of the expressed channel protein relative to native cardiac ion channels.

Mammalian (e.g., CHO, HEK-293, COS-7) rather than Xenopus oocytes should be selected for expressing HERG because a major limitation of the Xenopus model is that test substances can accumulate in the oocyte yolk, resulting in significant variability and error in potency estimates.

HERG rather than I_{Kr} present murine atrial tumor AT-1 or HL-1 cells derived from transgenic mice should be preferred. The channels expressed by these cells are similar to neonatal/fetal mouse cardiac myocytes raising some concerns about de-differentiation in these tumor-derived cells. Test results obtained are also clearly dependent on culture temperature (see Table 4.4), and must be clearly specified with any provided or published data.

Relevance of HERG to QT prolongation

Compounds associated with ADRs of QT prolongation, and arrhythmias such as TdP 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.

Expression and Recording Systems: HEK-293 cells have been transfected with cDNA for HERG-1 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.

The relevance and relative importance of these measurements in conjunction with all of the other components of cardiovascular safety assessment, is still not clear. Certainly it is influential in making drug development decisions,[21] but should not be a critical determining factor on its own.

Cardiovascular function testing

Traditionally, safety pharmacology screens for cardiovascular function have been performed in the anesthetized dog or rat. In the rat, diastolic, systolic, and mean arterial blood pressures (DAP, SAP, MAP) are recorded via direct cannulation of a major artery (e.g., femoral or carotid), and the test article is injected directly into a vein while the animal remains anesthetized. Heart rate is determined electronically from the pressure signal, and other physiological parameters may be monitored, such as body temperature, respiration rate, and EKG. Often, an escalating dose strategy is employed to determine the minimum dose necessary to see any changes in cardiovascular function.

Cardiovascular monitoring in dogs remains the preclinical gold standard[33] and is usually performed with a thermistor-tipped Swan–Ganz type catheter advanced directly into the pulmonary artery for measurement of cardiac output (CO), pulmonary arterial pressure (PAP), and core body temperature. A second catheter is placed in the caudal vena cava for measurement of central venous pressure (CVP) and for collection of blood for hematology and clinical chemistry. An additional catheter can be advanced into the aorta for measurement of DAP, SAP, and MAP, and for collection of arterial blood for blood gas analysis. A catheter also can be advanced to the left ventricle for measurement of lateral ventricular pressure (LVP) and dp/dt, a measurement of cardiac contractility. Simultaneously, a catheter may be placed in the bladder for continual voiding during surgery and testing, and the urine may be collected for measurements of volume, electrolyte concentration, and protein content or for determination of creatinine clearance. Again, as in the experiments in rats, EKG leads are used to record gross changes in cardiac electrophysiology in anesthetized dogs. This setup permits the simultaneous monitoring of a wide variety of important cardiovascular parameters. The obvious drawback is that these heavily instrumented animals are subject to acute, terminal instrumentation to permit accurate recording of these multiple parameters.

The potential interaction of the anesthetic compound or the anesthetized state with the test article is generally ignored. The rationale for this is the fact that safety pharmacology screens are, by definition, meant to identify gross physiologic changes in response to doses well above the expected therapeutic levels. However, if the anesthetic increases the response threshold to a test article, then we may be underestimating the safety index for that compound. If some parameters of cardiovascular function are depressed by the anesthetic, or if there are unidentified interactions of the test article with the anesthetic, we hope these can be unmasked by escalating the dose off the compound being tested. Traditional toxicity testing attempts to avoid complex drug interactions such as those with an anesthetic.

For basic cardiovascular monitoring in rats, many researchers have switched to tailcuff sphygmonanometry as a noninvasive technique for obtaining DAP, SAP, MAP, and heart rate in awake, nonanesthetized animals. The main drawback in using the tailcuff to measure blood pressure is the need for regular restraint whenever you want to take a reading. Ideally, rats

must be trained beforehand to be comfortable and remain docile in the restraining apparatus. The stress of restraint is well documented to significantly alter physiological state and, in particular, cardiovascular function. In addition, most tailcuff systems work best when the rat is warmer than standard room temperature, requiring the use of heat lamps or other methods to increase local body temperature in the tail. Heating may impact the cardiovascular recording and the functional state of the animal. Cuff sphygmomanometry also can be used with dogs (tail or leg) and monkeys (arm), but these animals must be acclimated for these conscious, restrained blood pressure measurements. Another drawback of this technique is the discontinuity of measuring cardiac function. The rats, dogs, and monkeys can be monitored only while restrained and, in intervening times, must be allowed to move freely. Nevertheless, cuff sphygmomanometry (tail or limb) is a valuable addition and noninvasive alternative allowing for the measurement of basic cardiovascular parameters in nonanesthetized animals.

More recently, researchers have been using radiotelemetric implants to record a number of cardiovascular parameters in awake, freely moving animals.[34] Most often, the lower abdominal aorta is cannulated and connected to a transmitter sutured directly to the wall of the peritoneal cavity. After surgery, the animal is allowed to recover from the anesthesia and returned to its home cage. A receiver unit is placed next to the cage to pick up the radio signals emitted by the telemetry device. Measurements of blood pressure, heart rate, body temperature, and EKGs are routinely taken this way. The major advantage of radiotelemetry is the elimination of external influences because all recordings can be performed from a site remote or hidden from the test animal, and the animal can remain in its home cage throughout the measurement and testing procedures. Of course, drug administration requires a separate setup, instrumentation, or restraint, depending on the route of delivery. In some instances, drug administration can be done from a remote site, too. Telemetry has the potential to record the purest data on the effects of a test article on cardiovascular function because stress from external events and handling are greatly reduced for an animal being studied with telemetry system.

An important advantage of radiotelemetry over cuff sphygmomanometry is the ability to record data continuously. Following implantation, the physiological detectors are always in place, and data storage space is the only practical limitation on data collection throughout an experiment. The major drawback of radiotelemetry is the expense and maintenance of the transmitters The greater expense leads many researchers to do long-term studies, either testing one compound over a very long time (e.g., chronic toxicity testing) or testing multiple compounds with adequate washout periods between them. Of course, with novel test articles, one is never sure of the adequacy of a washout period, particularly if the tissue distribution and half-life information for these novel compounds are not known. If multiple compounds are tested in one animal, then one cannot perform histopathology or tissue distribution evaluations at the same time. Currently with telemetry,

one cannot measure all the parameters that can be monitored via direct cannulation in anesthetized animals. Advances in electronics and miniaturization engineering may change this in the near future.

Radiotelemetry devices are also subject to interference from other nearby electronic instruments, and they have a tendency to display a baseline drift over time. Careful attention to the study setup and design can minimize these problems. To compensate for drift, one may need to periodically calibrate the animal and telemetry system with responses to known agents, but this is not desirable when testing an unknown compound. Lastly, maintenance of a telemetry-instrumented colony can be expensive. Many researchers find that the cost of telemetry implants is not feasible for short-term evaluations. Despite these drawbacks, radiotelemetry may produce the most accurate picture of the effects of a novel compound on cardiovascular function.

Hemodynamics, EKG, and respiration in anesthetized dogs or primates

Anesthetized studies using modular instruments data capture systems to record six-lead EKG (I, II, III, aV_l, and aV_f), 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. EKG intervals are measured from the lead II EKG and QT interval can be corrected for heart rate using Bazett's, Friderecia's, or Van De Water's formulae.[35,36]

Cardiac conduction studies

In addition to the preceding hemodynamic measurements, intra-ventricular, intra-arterial, and atrio-ventricular conduction times and velocities can be measured using epicardial electrodes in the anesthetized and thoracotomized dog.

Conscious rodent, dog, and primate telemetry studies

Effects on blood pressure, heart rate, lead II EKG, core body temperature, and locomotor activity can be explored using DataSciences telemetry implanted devices in rats, guinea pigs, dogs, or primates. Effects on behavior can be captured on video using closed-circuit television (CCTV) for dog and primate studies. Repeated administration and interaction studies can be performed.

Six-lead EKG measurement in the conscious dog

Conscious studies using integrated telemetry systems devices for measurement of blood pressure and six chest lead EKG measurements (V2,

V4, V6, V10, rV2, and rV4). EKG interval analysis is performed on the V2 lead (RR, PR, QT, QTc intervals, QRS duration). QT dispersion can also be measured. Locomotor activity can be monitored and behavior captured on video using CCTV.

- In addition to validated systems for automatic measurement of EKG parameters, EKGs can be reviewed by 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.

Recent concerns over the arrhythmogenic effects of a number of marketed compounds have resulted in the issue of the "Points to Consider" document, CPMP 986/96, by the EMA.

Studies are recommended to assess the effects of compound and any known metabolites on EKG and cardiac action potentials. 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.

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 the subsection inverse use dependent properties of sotalol in dog Purkinje fibers)

Cloned human potassium channels. Assessment of effects on cloned HERG K channels with stable expression in a cell line by measurement of whole cell K current (IKr) using voltage clamp. Other cloned human ion channels (e.g., KvLQT1/minK-IKs currents) are also possible.

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.[37]

Monophasic action potential in anesthetized guinea pigs. Epicardia monophasic action potential recording using suction/contact pressure electrodes is possible according to Carlsson et al.,[7] allowing simultaneous measurement of EKG.

EKG by telemetry in conscious guinea pigs. Lead II EKG recording is performed using DataSciences telemetry device. Repeated administration and interaction studies can be performed in this paradigm.

Hemodynamics and EKG in anesthetized or conscious dogs or primate:

- Conscious studies using DataSciences telemetry for blood pressure and lead II EKG or the ITS system for blood pressure and six chest lead EKG measurements (including QT dispersion)

- Anesthetized studies using MI[2] data capture system with additional measurement of blood flow in selected vascular beds, cardiac output, respiratory, and left ventricular function

Summary

Understanding and adequately and properly evaluating the adverse cardiovascular pharmacological effects of potential new drugs represents the most compelling current area of development and controversy in drug safety evaluation and drug development. The good news is that some very attractive tools and techniques are available for acquiring the necessary information (nonclinical and clinical, *in vitro* and *in vivo*), and many of these are well established and understood.

The situation is, in some ways, analogous to that of mutagenicity testing in the 1970s, however. The challenge lies in understanding all the information that is available and that we can collect, and integrating it into a risk evaluation model that both protects patients and does not inadvertently block or retard the clinical use of therapeutics that are safe and medically valuable.

References

1. Lacroix, P. and Provost, D., Basic safety pharmacology: the cardiovascular system *Therapie,* 55:63–69, 2000.
2. MacKenzie, I., Safety pharmacology requirements for the development of human cardiac/cardiovascular pharmaceuticals, *Drug Devel. Res.*, 55:73–78, 2002.

3. Malik, M. and Camm, A.J., Evaluation of drug-induced QT interval prolongation: implications for drug approval and labeling, *Drug Safety*, 24:323–351, 2001.
4. Cashin, C.H., Dawson, W., and Kitchen, E.A., The pharmacology of benoxaprofen (2-[4-chlorophynl]-α-methyl-5-benzoxazole acetic acid), LRCL 3794, a new compound with antiinflammatory activity apparently unrelated to inhibition of prostaglandin synthesis, *J. Pharm. Pharmacol.*, 29:330–336, 1977.
5. Graf, E. et al. Animal experiments on the safety pharmacology of lofexidine. *Arzneim.-Forsch./Drug Res.*, 32(II)(8a):931–940, 1982.
6. Bramm, E., Binderup, L., and Arrigoni-Martelli, E., An unusual profile of activity of a new basic antiinflammatory drug, timegadine, *Agents Actions*, 11:402–409, 1981.
7. Carlsson, L. et al., Electrophysiological characterization of the prokinetic agents cisapride and mosapride *in vivo* and *in vitro*: implications for proarrhythmic potential? *JPET*, 282:220–227, 1997.
8. Takasuna, K. et al., General pharmacology of the new quinolone antibacterial agent levofloxacin, *Arzneim.-Forsch./Drug Res.*, 42(I)(3a):408–418, 1992.
9. ICH, Safety Pharmacology Studies for Assessing the Potential for Delayed Ventricular Repolarization (QT Interval Prolongation), by Human Pharmaceuticals, 2002.
10. Thomas, S.H.L., Drugs, QT interval abnormalities and ventricular arrhythmias. *Adverse Drug React., Toxicol.*, Rev. 13; 77–102, 1994.
11. De Ponti F., Poluzzi, E., and Mantanaro, N., Organizing evidence on QT prolongation and occurrence of Torsades de Pointes with nonantiarrhythmic drugs: a call for consensus, *Eur. J. Clin. Pharmacol.*, 57; 185–209, 2001.
12. Moss, A.J., The QT interval and torsade de pointes, *Drug Safety*, 21; 5–10, 1999.
13. Ackerman, M.J., The long QT syndrome: ion channel diseases of the heart, *Mayo Clin. Proc.*, 73, 250–269, 1998.
14. Viskin, S., Long QT syndromes and torsade de pointes, *Lancet*, 354: 1625–1633, 1999.
15. Haverkamp, W. et al., The potential for QT prolongation and proarrhythmia by non-antiarrhythmic drugs: clinical and regulatory implications, *Cardiovasc. Res.*, 47:219–233, 2000.
16. Anderson, M.E. et al., Cardiac repolarization: current knowledge, critical gaps and new approaches to drug development and patient management. *AHJ*, 144, 2002.
17. Hammond, T.G. et al., Methods of collecting and evaluating non-clinical electrophysiology data in the pharmaceutical industry: results of an international survey, *Cardiovascular Res.*, 49:741–750, 2001.
18. Altman, E.M., *Cardiovascular Therapeutic*, 2nd ed., Philadelphia: W.B. Saunders Company, 2002.
19. Hardman, J.C. and Limbird, L.E., *Goodman & Gilman's The Pharmacological Basis of Therapeutics*, 9th ed., New York: McGraw-Hill, 1996.
20. Anon., *QT Internal Early Negative Signal Would Lesson Burden in Later Clinical Trial*, The Pink Sheet, 13–14, 27 January 2003.
21. Bonate, P. and Russell, T., Assessment of QTc prolongation for non-cardiac-related drugs from a drug development perspective, *J. Clin. Pharmacol.*, 39:349–358, 1999.
22. Bazett, H.C., An analysis of time — relations of electrocardiograms, *Heart*, 7:353–370, 1920.

23. Turner, R.A., *Screening Methods in Pharmacology*, vols. I and II, New York: Academic, 42–47, 60–68, 27–128, 1965.
24. Domer, F.R., *Animal Experiments in Pharmacological Analysis*, Springfield, IL: Charles C Thomas, 98, 115, 155, 164, 220, 1971.
25. Walker, M.J.A. and Pugsley, M.K., *Methods in Cardiac Electrophysiology*, Boca Raton, FL: CRC Press, 1998.
26. Temple, R., Are surrogate markers adequate to assess cardiovascular disease drugs? *JAMA*, 282:790–795, 1999.
27. Kinter, L.B. and Valentin, J.P., Safety pharmacology and risk assessment, *Fundamental Clin. Pharmacol.*, 16:175–182, 2002.
28. Hohnloser, S.H., Klingenheben, T., and Singh, B.N., Amiodarone-associated proarrhythmic effects, *Ann. Intern. Med*, 121(7): 529–535, 1994.
29. Cavero, I. and Crumb, W.J., Mechanism-designed assessment of cardiac electrophysiology safety of pharmaceuticlas using human cardiac ion channels *Business Briefing: Pharma. Tech.*, 1–9, July 2001.
30. Lacerda, A.E. et al., Comparison of block among cloned cardiac potassium channels by non-antiarrhythmic drugs, *Eur. Heart J.*, Suppl. K:K23–K30, 2001.
31. Rampe, D. et al. A mechanism for the proarrhythmic effects of cisapride (Propulsid): high affinity blockade of the human cardiac potassium channel HERG, *FEBS Lett.*, 417:28–32, 1997.
32. Fenichel, R.P. and Koerner, J., Development of drugs that alter ventricular repolarization, Internet-based draft, 1999.
33. Detweiler, D.K., Electrocardiography in toxicologic studies, in *Comprehensive Toxicology*, Sipes, I.G., McQueen, C.A., and Gadndolfi, J.A., Eds., New York: Pergamon Press, 95–114, 1997.
34. Abernathy, F.W., Flemming, C.D., and Sonntag, W.B., Measurement of cardiovascular response in male Sprague-Dawley rats using radiotelemetric implants and tailcuff sphymomanometry: a comparative study *Toxicol. Methods*, 5:89–98, 1995.
35. Osborne, B.E. and Leach, G.H.D., The Beagle Electrocardiogram, *Fd Cosmet. Toxicol.*, *9*:857–864, 1971.
36. Rosendorff, C. *Essential Cardiology: Principles and Practice*, Philadelphia: W.B. Saunders Company, 2001.
37. Burns, B.D. and Paton, W.D.M., Depolarization of the motor end-plate by decamethonium and acetylcholine, *J. Physiol.*, *(*London), 115:41–73, 1951.

chapter 5

Central nervous system

One characteristic that distinguishes central nervous system (CNS) safety pharmacology from discover or efficacy pharmacology, is that CNS safety pharmacology is generally conducted in normal animals. The aim is to see whether the new drug induces adverse changes in normal function, not whether the drug can have potential therapeutic effects on abnormal function. Not to be included in CNS safety pharmacology studies, therefore, would be models of pathology (e.g., depression, anxiety), unless it is expected that therapeutic use of the drug might present particular problems in a specific group of patients. Indeed, the notion of safety might differ radically depending on the intended therapeutic application. For example, when developing an NMDA antagonist for acute stroke, it would not be an important safety factor if the drug induced a reversible psychotomimetic state, providing the patient's life was saved. On the other hand, if such a drug were to be used preventively in patients suffering repeated minor strokes, the presence or absence of psychotomimetic potential could become a crucial safety consideration.

A further characteristic of CNS safety pharmacology is that studies are almost exclusively carried out *in vivo*, using conscious animals. In contrast to other organ systems, such as the cardiovascular system, CNS safety cannot rely on *in vitro* techniques because the neurobiological mechanisms of CNS function are less well understood.

Considering these points, what kinds of studies come within the scope of CNS safety pharmacology? This topic was extensively developed in the Japanese Guidelines,[1] which divided recommended CNS safety pharmacology studies into two categories: A and B. Category A included core battery studies such as general behavioral observation, measures of spontaneous motor activity, general anesthetic effects and eventual synergism/antagonism with general anesthetics, effects on convulsions (proconvulsant activity and synergy with convulsive agents), analgesia, and body temperature. Category B included effects of the test substance on the electroencephalogram (EEG), the spinal reflex, conditioned avoidance response, and locomotor coordination. The governing notion of Category A was that the studies

Table 5.1 CNS Safety Pharmacology Evaluation

Irwin test[a]
General assessment of effects on gross behavior and physiological state
Locomotor activity[b]
Specific test for sedative, excitatory effects of compounds
Neuromuscular function[b]
Assessment of grip strength
Rotarod
Test of motor coordination
Anaesthetic interactions
Test for central interaction with barbiturates
Anti/pro-convulsant activity
Potentiation or inhibition of effects of pentylenetetrazole
Tail flick[b]
Tests for modulation of nociception (also hot plate, Randall Selitto, tail pinch)
Body temperature[b]
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

[a] Usually a functional observational battery (FOB) is integrated into a rodent (rat) repeat dose toxicity studies to meet this requirement.

[b] **Properly designed and executed FOB can also meet their requirements.**

mentioned were obligatory (known as core battery under International Conference on Harmonization [ICH]), whereas those mentioned in Category B were to be carried out when necessary, but the concept of necessity remained undefined. More recently, the European Agency for the Evaluation of Medicinal Products has proposed a new set of guidelines, ICH S7A[2,3] in operation since June 2001, which have also been adopted in the U.S. and Japan. These guidelines (Table 5.1) are much less specific than the Japanese, and include motor activity, behavioral changes, coordination, sensory/motor reflex responses, and body temperature as core battery CNS studies with the remark that "… the central nervous system should be assessed appropriately…". Follow-up studies should include behavioral pharmacology examinations, etc.

In general, core battery studies should be carried out prior to first administration in humans, whereas the follow-up studies should be carried out prior to product approval. Such core battery studies should be carried out in full accordance with Good Laboratory Practice (GLP), whereas follow-up studies, because of their unique characteristics, require only assurance of data quality and integrity. Comparison of the Japanese Guidelines and ICH S7A suggests a clear intent by the European authorities to free safety pharmacology from the constraints of a cookbook approach. On the other hand,

the general terms — as with all the other portions of safety pharmacology guidelines except for those governing QT interval correction (QTc) internal prolongation — do not provide a clear and precise idea of what could or should be done.

The study designs presented here should serve to evaluate the CNS safety pharmacology profile of new entities and excipients while also fulfilling the requirements of the ICH S7A guidelines. The aim, however, is to provide concrete examples of how established techniques in CNS pharmacology can give body to the bare outlines represented by ICH S7A. After the presentation of the core battery tests, the follow-up procedures within the different domains are presented in order of their methodological complexity that can determine their priority of execution in the drug development process.[4]

Historically, there are four broad classes of approaches to assessing nervous system effects of drugs in animals. We shall consider those (behavioral/observational) that address core battery requirements first.

Core battery CNS procedures

Core battery CNS procedures are simple tests by design, using traditional techniques, which can be carried out rapidly in a routine fashion. They are the first techniques to be employed in safety assessment and are frequently applied at the beginning of the discovery process as a screen to eliminate substances with a potential for CNS risk. Because of their use early in the safety evaluation process, such studies are conducted almost exclusively in the rat or the mouse.

The tests described next generally cover the topics included in the ICH S7A guidelines. Three further tests are described, which are not mentioned in ICH S7A, but were included in the Japanese guidelines: the convulsive threshold, pain sensitivity, and interaction with barbiturates. It appears, to the present authors, that detection of eventual proconvulsant activity or hyperanalalgesic effects are topics of eminent concern for CNS safety. As for interaction studies with barbiturates, although hardly novel, these provide a very simple and sensitive means for unmasking eventual sedative or anti-sleep effects that are not always detectable using simple observation tests such as the Irwin or the functional observation battery (FOB).

General behavioral observation

A first approach to assessing the global behavioral profile of a novel substance could be whether the primary observation procedure originally described by Irwin[5] or the FOB originally Gad[6] more recently described by Haggerty[7] or Mattson et al.[8] Both are mentioned in ICH S7A, but the FOB is more specifically used for assessing neurotoxicity.

According to the basic Irwin procedure, rats or mice are given the test substance and are repeatedly observed over a 2- to 3-h period followed by daily observations up to 72 h using a standardized observation grid contain-

ing most or all of the following items: mortality, sedation, excitation, stereotypes, aggressiveness, reaction to touch, pain sensitivity, muscle relaxation, loss of righting reflex, changes in gait and respiration, catalepsy, ptosis, corneal reflex, pupil diameter, and rectal temperature. Animals are usually given a relatively high dose first, and then the other doses are selected on the basis of the effects observed. The aim is to establish, by means of a limited number of doses, the highest dose that can be given without inducing observable effects, the active dose-range, and the first lethal dose. This test also permits an estimate of the duration of activity and the kinds of behavioral effects observed.

Functional observational battery

Neurobehavioral evaluations are an important component of testing for the neurotoxic potential of chemicals. Observations made during standard toxicity studies or specialized neurotoxicity studies can provide information important for identifying and characterizing neurotoxic effects. A protocol that includes a framework for the systematic recording of observations and manipulations, such as a FOB, is an integral part of neurobehavioral screening. A neurobehavioral test battery can be composed of a variety of endpoints, usually chosen to assess an array of neurological functions, including autonomic, neuromuscular, sensory, and excitability.

The protocols in this unit are divided into (1) observational assessments and (2) manipulative tests. Each protocol is further subdivided into specific tests or endpoints (Table 5.2). These endpoints may be combined into a battery of tests for neurobehavioral screening. Most or all of these protocols/endpoints should be used in the context of a broad neurobehavioral test battery, whereas judicious selection of specific endpoints may be appropriate for more focused neurological testing. Originally developed by Gad, the method using rats is described in excellent detail in Moser.[9]

Observational assessments

This section provides observational assessments used to characterize neurological functions, a description of the behaviors being observed, and possible ranking scales to be used, where applicable. The observed behaviors are innate; that is, they do not need to be taught or shaped. Thus, observations require little or no interaction between the observer and the subject, with the possible exception of holding the rat. Because the observer makes judgments regarding these behaviors, the assessments are subjective.

The rat may be observed briefly in the home cage, but such observations are quite constrained and therefore limited. The home cage may prevent clear observations due to a variety of factors including available light, position of the cage, and clarity of the cage material. For these reasons, most of these evaluations are made in an open field or arena. The observer should have a clear, unobstructed view as the rat moves about. Choose an open field that is large enough for the rat to explore, has a nonslippery surface, and has a raised

Table 5.2 Endpoints in a Neurobehavioral Screening Battery

Observational Assessments	Manipulative tests
Activity levels:	*Neurological reflexes/reactions:*
Home-cage observations	Pupil response
Open-field observations	Palpebral reflex
Rearing	Pinna reflex
Reactivity/excitability:	Extensor thrust reflex
Reactivity	*Neuromuscular tests and postural reactions:*
Arousal	Grip strength
Gait and postural characteristics:	Landind food splay
Gait descriptions	Hopping righting reaction
Postural descriptions	*Sensory responses:*
Involuntary/abnormal motor movements:	Visual test—approach response
Tremors	Visual test—visual placing
Fasciculations	Somatosensory tests—touch response
Clonus	Auditory test—click response
Tonus	Nociceptive test—tail/toe pinch
Sterotypy	Nociceptive test—flexor reflex
Bizarre behaviors	Proprioceptive positioning test
Clinical signs:	Olfactory test
Lacrimation	
Salivation	
Hair coat	
Palpebral closure	
Ocular abnormalities	
Muscle tone/mass	

Note: Generally, a sample size of five subjects per sex per treatment is sufficient. It is essential that control animals be treated exactly the same as the other groups, except that they are administered only vehicle or formulation (minus drug).

border to prevent the subject from escaping or falling over the edge. Typical open fields include the top of a laboratory cart with a rim, or a bench top with metal sides inclosing the area. Cover the open field with clean absorbent paper that can be changed after each rat (or used to clean the area after each rat) to eliminate interfering olfactory cues. Environmental conditions (e.g., lighting and temperature) should be held constant from day to day.

A single observer may evaluate the endpoints described in here simultaneously during the handling phase and open-field observation period. It is suggested that the open-field period be sufficiently long for the observer to score the animal on all the associated endpoints (e.g., 2 or 3 min).

Locomotor activity

Locomotor activity can be quantified in rodents by a variety of means: interruptions of photoelectric beams, activity wheels, changes in electromagnetic fields, Doppler effects, video-image analysis, and telemetry.[7,13] Animals are administered the test substance and are placed in standardized enclosures for

a limited observation period (10 to 30 min), a fixed time after administration of the test substance. In contrast to the Irwin test, which is labor-intensive, most activity tests are automated to permit fairly large numbers of animals to be tested simultaneously. As the behavior measured occurs spontaneously, particular care has to be taken to ensure constant experimental conditions — time of day, temperature, noise level, lighting, apparatus cleanliness — to obtain reproducible results. Although different authors quantify different aspects of locomotion (small displacements, large displacements, rearing), the basic information yielded by activity meter tests is whether a test substance increases or decreases locomotion. Furthermore, data obtained should be correlated with that obtained from direct observation (Irwin test or FOB) to ensure that apparent decreases are not due to motor incapacity or even to overexcitation.

Motor coordination

Motor coordination is most commonly assessed using a rotarod.[10,11] Rats or mice are placed onto a rod rotating at either a fixed or constantly increasing speed. The time taken for the animal to fall off the rod, or the number of animals remaining on the rod over a set duration, is measured. To decrease test variability, the animals can be given prior habituation to the rotarod before receiving the test substance. Usually, several animals are tested simultaneously on the same rod, separated physically and visually by partitions. As such, the test lends itself readily to automation.

In contrast to most of the other core battery tests, the rotarod is unidirectional, detecting only the capacity of substances to decrease neuromuscular coordination. On the other hand, when used in conjunction with locomotor activity tests, it provides a useful quantification of the margin of safety between doses of test substances which alter spontaneous activity and those that disturb motor function.

Pain sensitivity

Nociception (pain perception) can be measured in rodents by a variety of procedures whereby aversive stimulation is applied externally either by heat or electrical stimulation to the tail or paws (tail flick, hot plate, plantar test), or internally by injection into the peritoneum of chemicals (acetic acid, phenylbenzoquinone). Inflammation pain can also be induced by injecting chemicals (carrageenan, formalin) into the paws. More complex approaches induce neuropathic pain by surgical lesions, usually to the sciatic or spinal nerve. For safety pharmacology purposes, usually only the simpler procedures are used. For example, with the hot plate test,[12] the animal is placed onto a heated metal plate (54°C) within a vertical cylinder, and the latency to licking its front paws is counted over a short period.

Analgesic effects are clearly demonstrated with major analgesics such as morphine, at doses that do not change spontaneous locomotion, but false positives can also be obtained with any drug that inhibits locomotion unless

this activity is quantified separately. A more interesting application of pain sensitivity tests is for detecting increased pain sensitivity. Drug-induced hyperalgesia could constitute an important risk factor for novel substances.

Convulsive threshold

Convulsions can be induced in rodents either by electric shock applied directly across the cerebrum (known as electrocerebral silence, or ECS), by chemical agents administered peripherally, or even by exposure to noise in specific strains. In safety pharmacology, it is particularly important to detect proconvulsant activity. On the other hand, although anticonvulsant activity does not in itself constitute a risk, many substances with anticonvulsant activity — for example, benzodiazepines — induce sedation and memory impairment, which have obvious implications for CNS safety. Both kinds of activity can be seen with chemically induced convulsions such as with pentylenetetrazol[10] where a shorter or longer latency to convulsions and deaths can be observed with pro- and anticonvulsants, respectively. A variant of the electroconvulsive method, whereby successive animals are subjected to increasing or decreasing intensities of ECS depending on the occurrence of convulsions with the preceding shock level, permits a sensitive measure of the convulsive threshold that can vary in both directions and is therefore particularly useful for safety pharmacology purposes.

Sleep induction and interaction with hypnotics

Sleep induction in rodents can be evaluated by loss of the righting reflex. Intrinsic sleep-inducing activity can be observed during the Irwin test or FOB (see the "Core battery CNS procedures" section). A more sensitive index of sleep enhancement can, however, be obtained by observing the interaction of the test substance with sleep induced by standard hypnotics such as barbiturates. A further advantage of this kind of procedure is that it can also reveal antisleep activity (decreased duration or abolition of barbiturate-induced sleep). These studies, therefore, provide a useful complement to studies of general behavior and spontaneous locomotion. Drug interaction procedures have the disadvantage because the effects observed could arise through pharmacokinetic factors such as changes in metabolism, rather than through a true pharmacodynamic interaction. For this reason, interaction tests using barbital as standard barbiturate are to be recommended because barbital itself undergoes no metabolism, in contrast to most other barbiturates.

Higher cognitive function

More complex tests of higher cognitive function find their place during the later phases of CNS safety assessment because they are more time-consuming to perform and cannot therefore be performed on a routine screening basis. The term cognition includes learning, memory and attention,

and general intellectual activity. It is important that drugs should be devoid of impairing effects on these functions, whatever their indications. Indeed, package inserts for many kinds of drugs contain warnings about the use of the drug while driving, working, or engaging in various daily occupations. It is, therefore, essential that CNS safety pharmacology provide procedures for evaluating these effects in animal studies. There is no standardized set of procedures in this area. The following provides an indication of the kinds of experiment that could be undertaken. The list is by no means exhaustive.

Passive avoidance

One of the simplest procedures for looking for adverse effects on learning/memory is the so-called one-trial passive avoidance task.[13] Variants of this procedure exist, but the principal is that a rat or a mouse receives an aversive stimulation in the recognizable environment and on a later occasion shows it has remembered by not going there (passive avoidance). For example, a rat placed into the lighted compartment of a two-compartment box will explore the apparatus and eventually enter the dark compartment where it receives a brief electric shock.[14] When placed again into the lighted compartment 24 or 48 h later, the rat will avoid going into the dark compartment despite a natural preference for the dark. Amnesia-inducing drugs (benzodiazepines, anticholinergics, NMDA antagonists) administered before the first exposure will attenuate the animal's memory for the shock as shown by a decreased latency to enter the dark compartment on the test day.

Passive avoidance is a fairly simple and rapid procedure because it involves learning obtained in a single trial. It is, therefore, suitable as a first screen for potential cognition impairing activity. On the other hand, it is not clearly interpretable in terms of the cognitive function implicated since it is not possible to distinguish a drug's effects on learning or memory, or whether the drug influenced performance by impairing attention or even pain sensitivity during the learning trial.

Morris maze

Another fairly simple procedure, which allows a greater degree of interpretation, is the Morris water maze. The water maze task has been most extensively used to investigate specific aspects of spatial memory. This task is based upon the premise that animals have evolved an optimal strategy to explore their environment and escape from the water with a minimum amount of effort (i.e., swimming the shortest distance possible). The time it takes a rat to find a hidden platform in a water pool after previous exposure to the setup using only available external cues is determined as a measure of spatial memory. Studies with pharmacologic agents are initiated when performance is stable and the water maze task is particularly sensitive to the effects of aging.[15] Alternatively, one can study the effects of a drugs or

lesions upon the acquisition of this task. Drugs can be administered (or lesions produced) prior to training to assess their effects either upon or after acquisition in order to assess their effects upon performance. A water maze can easily be built or purchased by an investigator, so the cost for this equipment can be quite low. Using the maze can be labor-intensive, requiring that a tester be present or nearby throughout the task. This maze is typically used for rats, but it can be scaled down in size for use with mice.

Materials

- Rats
- Pharmacologic or toxicologic agents (optional)
- Water maze apparatus
- Tracking system and software (such as provided by Columbus Instruments, HVS Image, San Diego Instruments, or CPL Systems)

Set up apparatus and begin acquisition testing

1. Set up water maze:
   - A water-tight pool, painted white, should be positioned in a room with various external cues that are visible to a rat swimming in the pool (e.g., a doorway, overhead lights and camera [if desired], and large simple designs on the walls).
   - Make water opaque by adding powdered milk or nontoxic white paint to the water.
   - The pool should be designed so that it can be easily drained on a regular basis.
2. Insert platform into one quadrant of the pool.
3. Place rat into water with its head pointed toward the side of pool. The starting position should be at a different, and randomized, location each day of testing (e.g., north, south, etc.).
4. Record time (in seconds) it takes the rat to find the submerged platform. Guide rat to platform on first few trials if it requires > 120 sec.

A tracking camera, positioned 200 cm above the center of the pool, can be used to quantify the distance swam on each trial and thereby determine swimming speed when combined with latency measurements. The tracking system can also display swim path and distance and proved additional information on search efficiency and exploration patterns during acquisition and probe trials. This equipment and associated computer software can be obtained from several commercial manufacturers.

5. Allow rat to remain on platform for 10 to 15 sec. This allows the experiment time to return to an appropriate place at the side of the water pool in order to be ready for step 6.
6. Remove rat from pool. Wait 5 min.

7. Release rat into pool from same location with platform in same location. Record time for rat to find platform.
8. Give each rat four trials on the first day.

Perform trials

9. On second day, insert platform in same location as on the first day.
10. Release rat with its head pointed towards the side of the water pool.
11. Record time it takes rat to find platform.
12. Give rat eight to ten trials per day with 5-min intertribal intervals for several days until performance is stable and latency to find the platform is low (< 5 to 7 sec).
13. Perform data analysis. Performance is expressed as the average time it takes each rat to find the submerged platform. The data are best presented as a line drawing comparing the latency to find the platform for each group versus daily test sessions. Data from 2 or 4 days of testing can be averaged into blocks.
14. Begin studies with test and control standard drug.

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 a classical pharmacologist normally performs these bioassays, a good technician can be trained to conduct them. The required equipment consists of a Mangus or similar style tissue bath[12,16,17] a physiograph or kymograph, force transducer, glassware, a stimulator, and a bench spectrophotometer. The assays utilized in the screening battery are listed in Table 5.3, 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 5.3) 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[11] 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

Table 5.3 Isolated Tissue Pharmacologic Assays

Assay system	Endpoint	Standards (agonist/ antagonist	References
Rat ileum	General activity	None (side-spectrum assay for intrinsic activity)	Domer[10]
Guinea pig vas deferens	Muscarinic nicotinic or Muscarinic	Methacholine/atropine Methacholine/ hexamethonium Methacholine/atropine	Leach[23]
Rat serosal strip	Nicotinic	Methacholine/ hexamethonium	Khayyal et al.[24]
Rat vas deferens	Alpha adrenergic	Norepinephrine/ phenoxybenzamine	van Rossum[25]
Rat uterus	Beta adrenergic	Epinephrine/propranol	Levy and Tozzi[26]
Rat uterus	Kinin receptors	Bradykinin/none	Gecse et al.[27]
Guinea pig tracheal chain	Dopaminergic	Dopamine/none	Domer[10]
Rat serosal strips	Tryptaminergic	5-Hydroxytryptamine (serotonin)/dibenzyline or lysergic acid dibromide	Lin and yeah[28]
Guinea pig tracheal chain	Histaminergic	Histamine/benadryl	Castillo and DeBeer[29,30]
Guinea pig ileum (electrically stimulated)	Endorphin receptors	Methenkephaline/none	Cox et al.[31]
Red blood cell hemolysis	Membrane stabilization	Chloropromazine (not a receptor-mediated activity)	Seeman and Weinstein[32]
Frog rectus abdominis	Membrane depolarization	Decamethonium iodide (not a receptor-mediated activity)	Burns and Paton[33]

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 Tyrodes 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. This is a classic nervous system pharmacology approach.

Electrophysiology methods

A number of electrophysiological techniques are 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.[18] 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 briefly overviewed.

Data collection and communication in the nervous system occurs 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 (that is, intracellular and extracellular) or multiple cell (e.g., evoked potentials [Eps], slow potentials). These potentials may be recorded in specific central or peripheral nervous system areas (e.g., visual cortex, hippocampus, sensory and motor nerves, muscle spindles) during various behavioral states or in *in vitro* preparations (e.g., nerve-muscle, retinal photoreceptor, brain slice).

CNS function: electroencephalography

The 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.

Because the EEG is 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.[13] 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 brainstem 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 after-discharge 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 because they are cheap, resist infection during chronic electrode and cannulae 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 or chronic EEG experiments. Limitations of the former are that most drugs that produce general anesthesia can 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).

Excellent reviews of these electrophysiology approaches can be found in Fox et al.[19] and Takeuchi and Koike.[20]

Neurochemical and biochemical assays

Though elegant methods are now available to study the biochemistry of the brain and nervous system, none have yet discovered any generalized marker chemicals, which serve as reliable indicators or early warnings of neurotoxic actions or potential actions. Some useful methods are available, however. Before looking at these, one should understand the basic problems involved.

Normal biochemical events surrounding the maintenance and functions of the nervous system centers on 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. Measurement 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 Damstra and Bondy.[21,22]

References

1. MHW, *Guidelines for Toxicity Studies of Drugs*, 1999.
2. CPMP, Note for Guidance on Safety Pharmacology Studies in Medicinal Product Development, 1998.
3. ICH, Safety Pharmacology Studies for Human Pharmaceuticals, 2000.
4. Porsolt, R.D. et al., New perspectives in CNS safety pharmacology, *Fundamental Clin. Pharmacol.*, 16:197–207, 2002.
5. Irwin, S., Comprehensive observational assessment: la. A systematic, quantitative procedure for assessing the behavioral and physiologic state of the mouse, *Psychopharmacologia* **(Berlin)**, 13:222–257, 1968.
6. Gad, S.C., A neuromuscular screen for use in industrial toxicology, *J. Toxicol. Environ. Health*, 9:691–704, 1982.
7. Haggerty, G.C., Strategies for and experience with neurotoxicity testing of new pharmaceuticals, *J. Am. Coll. Toxicol.*, 10:677–687, 1991.
8. Mattson, J.L., Spencer, P.J., and Albee, R.R., A performance standard for clinical and functional observational battery examination of rats, *J. Am. Coll. Toxicol.*, 15:239, 1996.
9. Moser, V.C., Neurobehavioral screening in rodents (Unit 11.2), in *Current Protocols in Toxicology*, Maines, M.D., Ed., New York: John Wiley & Sons, 1999.
10. Domer, F.R., *Animal Experiments in Pharmacological Analysis, Springfield, IL:* Charles C Thomas: 1971, 98, 115, 155, 164, 220.
11. Bondy, S.C., 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, R.M. and Frankos, V.H., Eds.,Washington, D.C.: Office of Health Affairs, FDA, 1979, 133–143.
12. Turner, R.A., *Screening Methods in Pharmacology*, vols. I and II. New York: Academic, 1965, 42–47, 60–68, 27–128.
13. Norton, S., Toxic responses of the central nervous system, in *Toxicology: The Basic Science of Poisons*, 2nd ed. Doull, J., Klaassen, C.D., and Amdur, M.O., Eds. New York: Macmillan, Inc., 1980.
14. Brady, J.V. and Lukas, S.E., *Testing Drugs for Physical Dependence Potential and Abuse Liability.* NIDA Research Monograph 52, 1984.
15. Brandeis, R., Brandys, Y., and Yehuda, S., The use of the Morris water maze in the study of memory and learning, *Int. J. Neuroscience*, 48:29–69, 1989.
16. Offermeier, J. and Ariens, E.J., Serotonin I., Receptors involved in its action, *Arch. Int. Pharmacodyn. Ther.*, 1966, 64:92–215.
17. Nodine, J.H. and Seiger, P.E., *Animal and Clinical Pharmacologic Techniques in Drug Evaluation*, **Chicago:** Year Book Medical Publishers, Inc., 1964, 36–38,
18. Seppalaninen, A.M., Applications of neurophysiological methods in coocupational medicine: **a review.** *Scand. J. Work Environ. Health*, 1:1–14, 1975.

19. Fox, D.A., Lowndes, H.E., and Bierkamper, G.G., Electrophysiological techniques in neurotoxicology, in *Nervous System Toxicology,* Mitchell C. L., Ed., New York: Raven Press, 1982, 299–336
20. Takeuchi, Y. and Koike, Y., Electrophysiological methods for the *in vivo* assessment of neurotoxicology, in *Neurotoxicology,* Blum, K. and Manzo, L., New York: Marcel Dekker, Inc., 1985, 613–629.
21. Damstra, T. and Bondy, S.C., The current status and future of biochemical assays for neurotoxicity, In *Experimental and Clinical Neurotoxicology,* Spencer, P.S. and Shaumburg, H.H., Eds., Baltimore: Williams & Wilkins, 1980, 820-833.
22. Damstra, T. and Bondy, S.C., Neurochemical approaches to the detection of neurotoxicity, in *Nervous System Toxicology,* Mitchell, C.L., Ed. New York: Raven Press, 1982, 349–373.
23. Leach, G.D.H., Estimation of drug antagonisms in the isolated guinea pig vas deferens, *J. Pharm. Pharmacol.,* 8:501, 1956.
24. Khayyal, M.T. et al., A sensitive method for the bioassay of acetylcholine, *Eur. J. Pharmacol.,* 25:287–290, 1974.
25. van Rossum, J.M., Different types of sympathomimetic β–receptors, *J. Pharm. Pharmacol.,* 17:202, 1965.
26. Levy, B. and Tozzi, S., The adrenergic receptive mechanism of the rat uterus, *J. Pharmacol. Exp. Ther.,* 142:178, 1963.
27. Gecse, A., Zsilinksky, E., and Szekeres, L., Bradykinin antagonism, in *Kinins; Pharmacodynamics and Biological Roles*, Sicuteri, R., Back, N., and Haberland, G., Eds., New York: Plenum Press, 1976, 5–13.
28. Lin, R.C.Y. and Yeoh, T.S., An improvement of Vane's stomach strip preparation for the assay of 5-hydroxy-tryptamine, *J. Pharm. Pharmacol.,* 17:524–525, 1965.
29. Castillo, J.C. and De Beer, E.J., The guinea pig tracheal chain as an assay for histamine agonists, *Fed. Proc.* 6:315, 1947a.
30. Castillo, J.C. and De Beer, E.J., The tracheal chain, *J. Pharmacol. Exp. Ther.,* 90:104, 1947b.
31. Cox, B.M. et al., A peptide-like substance from pituitary that acts like morphine 2, Purification and properties, *Life Sci.,* 16:1777–1782, 1975.
32. Seeman, P. and Weinstein, J., Erythrocyte membrane stabilization by tranqilizers and antihistamines, *Biochem. Pharmacol.*, 15:1737–1752, 1966.
33. Burns, B.D. and Paton, W.D.M., Depolarization of the motor end-plate by decamethonium and acetylcholine. *J. Physiol.* **(London)**, 115:41–73, 1951.

chapter 6

Respiratory system

As early as 1964, it became apparent that β-adrenergic blocking agents could lead to bronconstriction (and possible death) in asthmatics.[1] 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 6.1, 6.2, and 6.3. Resulting worldwide regulatory requirements (Tables 6.4 and 6.5) support the need for conducting respiratory evaluations in safety pharmacology.

The objective of such studies is to evaluate the potential for drugs to cause secondary pharmacologic or toxicologic effects that influence respiratory function. Changes in 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 pressures (work) required for lung inflation and deflation.

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. Hyperventilation 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), the appropriate depth (tidal volume), and frequency of breathing. The depth and frequency of breathing required for alveolar ventilation are determined primarily by the 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

Table 6.1 Agents Known to Cause Pulmonary Disease

Chemotherapeutic	Analgesics
Cytotoxic	Heroin*
Azathioprine	Methadone*
Bleomycin*	Noloxone*
Busulfan	Ethchlorvynol*
Chlorambucil	Propoxyphene*
Cyclophosphamide	Salicylates*
Etoposide	Cardiovascular
Melphalan	Amiodarone*
Mitomycin*	Angiotensin-converting enzyme inhibitors
Nitrosoureas	Anticoagulants
Procarbazine	Beta-blockers*
Vinblastine	Dipyridamole
Ifosfamide	Fibrinolytic agents*
Noncytotoxic	Protamine*
Methotrexate*	Tocainide
Cytosine arabinoside*	Inhalants
Bleomycin*	Aspirated oil
Procarbazine*	Oxygen*
Antibiotic	Intravenous
Amphotericin B*	Blood*
Nitrofurantoin	Ethanolamine oleate (sodium morrhuate)*
Acute*	Ethiodized oil (lymphangiogram)
Chronic	Talc
Sulfasalazine	Fat emulsion
Sulfonamides	Miscellaneous
Pentamidine	Bromocripitine
Anti-inflammatory	Dantrolene
Acetylsalicylic acid*	Hydrochlorothiazide*
Gold	Methysergide
Methtrexate	Oral contraceptives
Nonsteroidal antiinflammatory agents	Tocolytic agents*
Penicillamine*	Tricyclics*
Immunosuppressive	L-Tryptophan
Cyclosporin	Radiation
Interleukin-2*	Systemic lupus erythematosus (drug-induced)*
	Complement-mediated leukostasis*

* Typically cause acute or subacute respiratory insufficiency.

Sources: Touvay, C. and Le Mosquet, B., Therapie, 55:71–83, 2000; Akoun, G.M. et al., in *Drug Induced Disorders, Volume 3: Treatment Induced Respiratory Disorders*, New York: Elsevier Scientific Publishers B.V., 1989, 3–9; Dorato, M.A., *Drugs Pharmaceutical Sci.*, 62:345-381, 1994; Lalej-Bennis, D. et al., *Diabetes Metab.*, 2001, 27(3):372–377; Mauderly, J.L., *Concepts in Inhalation Toxicology*, New York: Hemisphere Publishing Corp., 1989, 347–401; and Rosenow, E.C. et al., *Chest*, 102:239–250, 1992.

Table 6.2 Drugs That Adversely Affect Respiratory Function

Drugs known to cause or aggravate bronchospasm	Agents Associated with Pleural Effusion
Vinblastine	Chemotherapeutic agents
Nitrofurantoin (acute)	Nitrofurantoin (acute)
Acetylsalicylic acid	Bromocriptine
Nonsteroidal antiinflammatory 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	
Agents associated with acute-onset pulmonary insufficiency[a]	**Agents that cause subacute respiratory failure**
Bleomycin plus O_2	Chemotherapeutic agents
Mitocycin	Nitrofurantoin (chronic)
Bleomycin[b]	Amiodarone
Procarbazine[b]	L-Trytophan
Methotrexate[b]	Drug inducing systemic lupus erythematosus
Amphotericin B	
Nitrofurantoin (acute)[c]	
Acetylsalicylic acid[c]	
Interleukin-2[c]	
Heroin and other narcotics[c]	
Epinephrine[c]	
Ethchlorvynol[c]	
Fibrinolytic agents	
Protamine	
Blood products[c]	
Fat emulsion	
Hydrochlorothiazide	
Complement-mediated leukostasis	
Hyskon (Dextran-70)[c]	
Tumor necrosis factor[c]	
Intrathecal methotrexate	
Tricyclic antidepressants[c]	
Amiodarone plus O_2	
Naloxone	

[a] Onset at less than 48 h.

[b] Associated with hypersensitivity with eosinophilia.

[c] Usually reversible within 48–72 h, implying noncardiac pulmonary edema rather than inflammatory interstitial pneumonitis.

Sources: McNeill, R.S., *Lancet*, 21:1101–1102, November 1964; Borison, H.L, *Pharmacol. Ther.* B., 3:211-226, 1997; Eldridge, F.L. and Millhorn, D.E., *Ann. Rev. Physiol.*, 3:121–135, 1981; Tattersfield, A.E., *J. Cardiovasc. Pharmacol.*, 8 (Suppl. 4):535–539, 1986; Heymans, C., in *Reflexogenic Areas of the Cardiovascular Systems*, London: Churchill Ltd., 1958, 192–199; Illum, L. (Reprint), Davis, S.S., *Clinical Pharmacokin.*, 23:30–41, 1992; Keats, A.S., *Ann. Rev. Pharmacol. Toxicol.*, 25:41–65, 1985; Mueller, R.A. et al., *Pharmacol. Rev.*, 34:255–285, 1982; Shao, Z. et al., *Pharmaceutical Res.*, 9(9):1157–1163, 1992; and Shao, Z. and Mitra, A.K., *Pharmaceutical Res.*, 9:11184–1189, 1992.

Table 6.3 Drugs Known to Influence Ventilatory Control

Depressants	Stimulants
Inhaled anesthetics	Alkaloids
Barbituates	Nicotine
Benzodiazepines	Lobeline
Diazepam	Piperdine
Temazapan	Xanthine analogs
Chlordiazepoxide	Theophyline
Serotonin analogs	Caffeine
Methoxy-(dimethyl)-tryptamine	Theobromine
Dopamine analogs	Analeptics
Apomorphine	Doxapram
Adenosine analogs	Salicylates
2-Chloroadenosine	Progestrone analogs
R-Phenylisopropyl-adenosine (R-PIA)	Almitrine
	Glycine analogs
N-Ethylcarboxamide (NECA)	Strychnine
B-Adrenergic antagonists	γ–aminobutyric acid (GABA) antagonists
Timolol maleate	Picrotoxin
GABA analogs	Bicuculline
Muscimol	Serotonin synthesis inhibitors
Baclofen	p-Chlorophenylalanine
Opiates	Reserpine
Morphine	
Codeine	
Methadone	
Meperidine	
Phenazocine	
Tranquilizers/Analgesics	
Chlorpromazine	
Hydroxyzine	
Rompum (xylazine)	
Nalorphine	

the rate of breathing can influence the efficiency of ventilation.[19] The inspiratory and expiratory phases of individual breath rates of airflow and durations are distinct and independently controlled.[20] 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.[19] 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 neurological components of the pumping apparatus that are located in the CNS and include the medullary central pattern generator (CPG) as well as integration centers

Table 6.4 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 anaesthetized animals

Table 6.5 Regulatory Documents Recommending Respiratory Function Testing in Safety Pharmacology Studies

U.S.	FDA Guideline for the Format and Content of the Nonclinical Pharmacology/Toxicology Section of an Application (Section IID, p. 12, Feb. 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, Jan. 1991
Australia	Guidelines for preparation and presentation of Applications for Investigational Drugs and Drug Products under the Clinical Trials Exemption Scheme (pp. 12, 15)
Canada	RA5 Exhibit 2, Guidelines for Preparing and Filing Drug Submissions (p. 21)
U.K.	Medicines Act 1968, Guidance Notes on Applications for Product Licenses (MAL 2, p. A3F-1)

Results

Parameters	Theophyline 10 mg/kg PO	Pentobarbital 35 mg/kg IP	Diazapam 35 mg/kg IP	Codeine 100 mg/kg IP
F (beat/min)	+ + +	- - -	- - -	No change
TV (ml)	No change	No change	No change	-
Ti (s)	- -	+ +	+ +	+
Te (s)	- -	+ + +	+ +	-
PIF (ml/s)	+ +	-	-	-
PEF (ml/s)	+ +	No change	+	-
Penh	-	+	+	No change

Source: Data from Toura, C. and Le Mosquet, B., *Therapie*, 55:71–83, 2000.

located in the medulla, pons, hypothalamus, and cortex of the brain that regulate the output of the CPG.[20] The major neurologic inputs from the peripheral nervous system that influence the CPG are the arterial chemoreceptors.[20] Many drugs stimulate or depress ventilation by selective interaction with the central nervous system[4,9,15] or arterial chemoreceptors.[11,12]

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 toxicology studies (rats and dogs) because 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.[21–25]

The key questions in safety pharmacology of the respiratory system are:

- 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?

Plethysmography

The classic approach to measuring respiratory function in laboratory animals is plethysomography. It has two basic governing principles:[20,26–28]

1. The animal (mice, rat, or dog), anesthetized or not, restrained or not, is placed in a chamber (simple or double) with pneumotachographs.
2. The variations of pressure in chambers at the time of the inspiration and the expiry make it possible to obtain the respiratory flow of the animal.

Three main types of body plethysmographs are used: constant volume, constant pressure, and pressure volume. The constant-volume body plethysmograph is a sealed box that detects volume change by 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 chane in plethysmographic pressure and the appropriate calibration factor. The plethysmograph is calibrated by injecting or withdrawing the change in box pressure. To avoid an adiabatic artifact, the rate of air injection or withdrawal is kept the same as that of chest expansin, indicated by the same dp/dt.

The constant-pressure body plethysmograph is a box with a pneumotachograph port built into its wall. This plethysmograph detects volume change via integration of the flow rate, 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 detection of plethysmographic 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 assuracy 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.[29] Similar plethysmographs can also be provided for use with large mammals.

A third type of pressure-volume plethysmograph has the mixed characteristics of the two types of body box mentioned previously. 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 pressurer-volume plethysmograph has excellent frequency-response characteristics and a wide range of sensitivities.[30]

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, 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 6.6 defines the parameters that are typically determined by these methods, while Figure 6.1 illustrates how they actually appear in tracings.

Selection of the proper reference values for interpretation of findings is essential.[31,32]

Table 6.6 Pulmonary Variables from the Maximal Expiratory Flow-Volume Curve

Test description	Recommended term	Symbol
1. Volume of gas expired after full inspiration expiration being as rapid and complete as possible (i.e., forced)	Forced vital capacity	FVC
2. Peak expiratory flow (liter/min or liter/s) measured by various instruments	Peak expiratory flow, qualified by name of instrument used	PEF
3. Volume of gas exhaled over a given time interval during a complete forced expiration	Forced expiratory volume, qualified by time interval in s	FEV_t
4. FEV_t expressed at % of FVC	Percent of FVC expired in time interval t	FEV_t/FVC%
5. Volume of air exhaled over a specified volume range of the FVC divided by the time to exhale this volume, expressed as liter/min or liter/s	Mean forced expiratory flow between two designated volume points in FVC	
Examples:		
Volume between 0.2 and 1.2 liters of the FVC/time	Mean forced expiratory flow, 0.1–1.2 1	$FEF_{0.2–1.2\ liters}$
Volume between 25 and 75% of the FVC time	Mean forced expiratory flow, 25–75% of FVC	$FEF_{25–75\%}$
6. Maximal expiratory flow at a specific volume during FVC expressed in liter/min or liter/s	Maximal expiratory flow qualified at XX percent of VC (Note: 100% VC is a TLC; 0% is at RV)	
Examples:		
Flow at point when 75% of FVC remaining	Maximal expiratory flow, 75% of VC	$Vmax_{75}$
Flow at point when 50% of FVC remaining	Maximal expiratory flow, 50% of VC	$Vmax_{50}$
Flow at point when 25% of FVC remaining	Maximal expiratory flow, 25% of VC	$Vmax_{25}$

Design of respiratory function safety studies

General considerations

The objective of a safety pharmacology evaluation of the respiratory system is to determine whether a drug has the potential to produce a change in respiratory function. Because 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

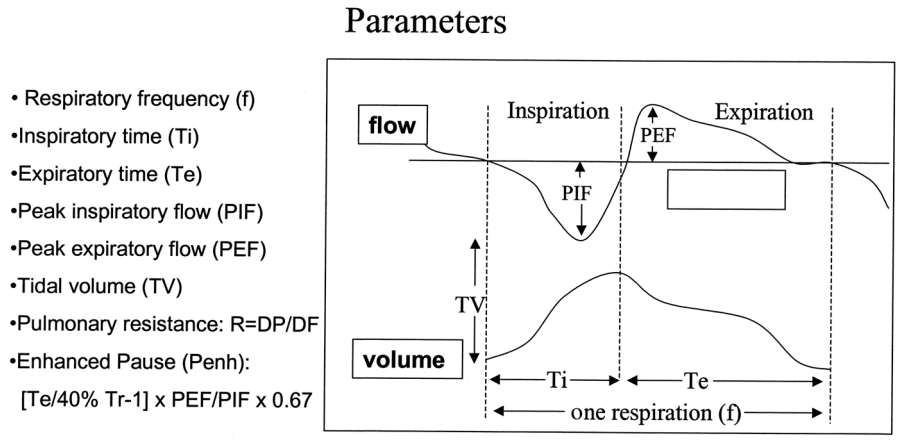


Figure 6.1 Diagrammatic representation of normal respiratory pattern.

properties of the lung in anesthetized/paralyzed animals. Together, these evaluations are used to determine (1) if drug-induced changes in the total respiratory system have occurred and (2) whether these changes are related to pulmonary or extra-pulmonary factors.

Study design

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 drug (T_{max}), at least one time before and one after T_{max}, and 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 to 30 min postdose. Time intervals of the 1, 2, 3, and 24 h are also monitored to evaluate possible delayed effects. If administered by inhalation or intravenous infusion, ventilatory parameters would generally be monitored continuously during the exposure period and at the 1-, 2-, 3-, and 24-h time intervals after dosing.

The time interval showing the greatest ventilatory change is selected for evaluating lung mechanics. If no ventilatory change occurred, however, the T_{max} is 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 CO_2 measurements, or responses to CO_2 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.

Capnography

The measurements of rates, volumes, and capacities provided by plethysmograph measurements have a limited ability to detect and evaluate some ventilatory disorders.[21]

Detection of hypo- or hyperventilation syndromes requires measurement of the partial pressure of arterial CO_2 ($PaCO_2$). In humans and large animal models, this can be accomplished by collecting arterial blood with a catheter or needle and analyzing for $PaCO_2$ 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 $PaCO_2$ is the measurement of peak expired (end-tidal) CO_2 concentrations. This technique has been successfully used in humans,[33] and recently has been adapted for use in conscious rats.[34] Measuring end-tidal CO_2 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 CO_2 concentrations. End-tidal CO_2 values in rats are responsive to ventilatory changes and accurately reflect changes in $PaCO_2$.[34]

A noninvasive procedure in conscious ratsthat is used to help distinguish between the central and peripheral nervous system effects of drugs on ventilation, has been developed. Exposure to CO_2 gas stimulates ventilation primarily through a central mechanism.[8] In contrast, a bolus injection of NaCN produces a transient stimulation of ventilation through a mechanism that involves selective stimulation of peripheral chemoreceptors.[12] 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% CO_2 gas and a bolus IV injection of 300 μg/kg of NaCN. A central depressant (e.g., morphine sulfate) inhibits the CO_2 response and has little effect on the NaCN response.

Study design considerations

Dose selection

Safety pharmacology studies are designed to evaluate acute functional changes and, thus, involve single administration of test drugs. Further, because pharmacologic and toxicologic effects can be route-specific, the test drugs are given by the intended clinical route. To evaluate dose response relationships, two or more doses are used. Selection of the high dose is based on acute toxicologic findings and is generally the lowest dose that produces evidence of toxicity (minimally toxic dose). This dose, however, must not be associated with toxicologic changes in secondary organs systems that may

compromise the respiratory function measurements. The middle and low doses are generally log or half-log decrements of the high dose, with the low dose no less than the primary pharmacologic or anticipated human dose.

Species selection

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: the pharmacokinetic data generated in these species can be used to define the test measurement intervals, and 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 because rats are readily available and techniques for measuring pulmonary function are well established in this species.

Summary

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 function evaluations in safety pharmacology. Safety pharmacology studies of the respiratory system should include evaluations of both the total respiratory system and the mechanical properties of the lung. Changes in ventilatory patterns of conscious animals are used to evaluate the total respiratory system and the parameters measured include respiratory rate, tidal volume, minute volume, peak (or mean) inspiratory flow, peak (or mean) expiratory flow, and fractional inspiratory time. Changes in lung airflow and elasticity are used to evaluate the mechanical properties of the lung. The functional endpoints used to evaluate lung airflow include PEF (peak expiratory flow), FEF_{25} (forced expiratory flow-25%), FEF_{75}, FEV (forced expiratory volume), and dynamic resistance (or conductance), while the endpoints used to evaluate lung elasticity include TLC (total lung capacity), IC (inspiration capacity), FRC (forced respiratory capacity), and compliance.

The species recommended for use in safety pharmacology studies are those used in toxicology studies because pharmacokinetic and toxicologic/pathologic data would be available. 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.

References

1. McNeill, R.S., Effect of a β-adrenergic blocking agent, propranol, on asthmatics, *Lancet*, 21:1101–1102, November 1964.
2. Touvay, C. and Le Mosquet, B., Systeme respiratore et pharmacology de securite, *Therapie*, 55:71–83, 2000.

3. Akoun, G.M. et al. Natural history of drug-induced pneumonitis, in Akoun, G.M. and White J.P., Eds., *Drug Induced Disorders Volume 3: Treatment Induced Respiratory Disorders*, New York: Elsevier Scientific Publishers B.V., 1989, 3–9.
4. Dorato, M.A., Toxicological evaluation of intranasal peptide and protein drugs, *Drugs Pharmaceutical Sci.*, 62:345–381,1994.
5. Lalej-Bennis, D. et al., 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.*, 2001, 27(3): 372–377.
6. Mauderly, J.L., Effects of inhaled toxicants on pulmonary function, in McClellan, R.O. and Henderson, R.F., Eds., *Concepts in Inhalation Toxicology*, New York: Hemisphere Publishing Corp, 1989, 347–401.
7. Rosenow, E.C. et al., Drug-induced pulmonary disease: an update, *Chest*,102:239–250, 1992.
8. Borison, H.L., Central nervous system depressants: control-systems approach to respiratory depression, *Pharmacol. Ther. Bull.*, 3:211–226, 1977.
9. Eldridge, F.L. and Millhorn, D.E., Central regulation of respiration by endogenous neurotransmitters and neuromodulators, *Ann. Rev. Physiol.*, 3:121–135, 1981.
10. Tattersfield, A.E., Beta adrenoreceptor antagonists and respiratory disease, *J. Cardiovasc. Pharmacol.*, 8 (Suppl. 4):535–539, 1986.
11. Heymans, C., Action of drugs on carotid body and sinus, *Pharmacol. Rev.*, 7:119–142, 1955.
12. Heymans, C. and Niel, E., The effects of drugs on chemoreceptors, In Heymans, C. and Neil, E., Eds., *Reflexogenic Areas of the Cardiovascular Systems*, London: Churchill, Ltd., 1958, 192–199.
13. Illum, L. (Reprint); Davis., S.S., Intranasal insulin — clinical pharmacokinetics, *Clinical Pharmacokin.*, 23:30–41, 1992.
14. Keats, A.S., The effects of drugs on respiration in man, *Ann. Rev. Pharmacol. Toxicol.*, 25:41–65, 1985.
15. Mueller, R.A. et al., The neuropharmacology of respiratory control, *Pharmacol. Rev.*, 34:255–285, 1982.
16. Shao, Z., Krishnamoorthy, R., and Mitra, A.K., Cyclodextrins as nasal absorption promoters of insulin-mechanistic evaluations, *Pharmaceutical Res.*, 9(9):1157–1163, 1992.
17. Shao, Z. and Mitra, A.K., Nasal membrane and intracellular protein and enzyme release by bile salts and bile salt fatty-acid mixed micelles-correlation with facilitated drug transport, *Pharmaceutical Res.*, 9:1184–1189, 1992.
18. Cherniack, N.S., Disorders in the control of breathing: hyperventialtion syndromes, in *Textbook of Respiratory Medicine*, Murray, J.F. and Nadal, J.A., Eds., Philadelphia: W.B. Saunders Co., 1988, 1861–1866.
19. Milic-Emili, J., Recent advances in clinical assessment of control of breathing, *Lung*, 160:1–17, 1982.
20. Boggs, D.F., Comparative control of respiration, in Parent, R.A., Ed., *Comparative Biology of the Normal Lung, Vol I.* Boca Raton, FL: CRC Press, 1992, 309–350.
21. Murphy, D.J., Safety pharmacology of the respiratory system: techniques and study design, *Drug Dev. Res.*, 32:237–246, 1994.
22. Amdur, M.O. and Mead, J., Mechanic of respiration in unanesthetized guinea pigs, *Am. J. Physiol.*, 192:364–368, 1958.

23. Diamond, L. and O'Donnell, M., Pulmonary mechanics in normal rats, *J. Appl. Physiol.: Respir. Environ. Exercise Physiol.*, 43:942–948,1977.
24. King, T.K.C., Measurement of functional residual capacity in the rat, *J. Appl. Physiol.*, 21:233–236, 1966.
25. Mauderly, J.L., The influence of sex and age on the pulmonary function of the beagle dog. *J. Gerontol.*, 29:282–289, 1977.
26. O'Neil, J.J. and Raub, J.A., Pulmonary function testing in small laboratory mammals, *Environ. Health Perspect.*, 53:11–22, 1984.
27. Palecek, F., Measurement of ventilatory mechanics in the rat, *J. Appl. Physiol.*, 27:149–156, 1969.
28. Brown, L.K. and Miller, A., Full lung volumes: functional residual capacity, residual volume and total lung capacity, in Miller, A., Ed., *Pulmonary Function Tests: A Guide for the Student and House Officer,* New York: Grune & Stratton, Inc., 1987, 53–58.
29. Sinnet, E.E. et al., Fast integrated flow plethysmograph for small mammals, *J. Appl. Physiol.*, 50:1104–1110, 1981.
30. Leigh. D.E. and Mead, J., *Principals of Body Plethysmography,* Bethesda, MD: National Heart and Lung institute, National Institutes for Health, 1974.
31. American Thoracic Society, Lung function testing: selection of reference values and interpretative strategies, *Am. Rev. Respir. Dis.*, 144:1202–1218, 1991.
32. Drazen, J.M., Physiological basis and interpretation of indices of pulmonary mechanics, *Environ. Health Perspect.*, 56:3–9, 1984.
33. Nuzzo, P.F. and Anton, W.R., Practical applications of capnography, *Resp. Ther.*, 16:12–17, 1986.
34. Murphy, D.J., Joran, M.E., and Grando, J.C., Microcapnometry: a non-invasive method for monitoring arterial CO_2 tension during ventilatory measurements in conscious rats, *Toxicol. Methods*, 4:177–187, 1994.

chapter 7

Renal function

The kidney is a very important and common target for numerous classes of drugs. Unfortunately, adverse renal pharmacologic effects may elude detection by classical regulatory toxicological studies. It is the responsibility of safety pharmacologists to detect it by using validated functional evaluation methods that are well established. Table 7.1 presents a partial list of known nephrotoxic drugs.

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 glomurelar filtration rate appears to be of major interest because it is associated with the definition of acute renal failure. Measurement of the renal blood flow, proteinuria, enzymuria, and fractional excretion of sodium are 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. Parameters may vary according to the species, sex, strain, and age. Because 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, co-administration of pharmacologic agents, etc.).

The kidney is a privileged target for toxic agents because of:

1. Physiological properties:
   - The most irrigated organ per gram of tissue (~400 ml/100 g), and therefore it is more exposed to exogenous circulating toxins than many other organs
     - Tubular mechanisms of ion transport acts to facilitate drug entry into renal tubular cells
2. Pharmacokinetic properties:
   - Involved in filtration, secretion and reabsorption of drugs
   - Kidney concentrate urine → intratubular drug concentration may be much higher than plasma concentration
   - high metabolic rate

Table 7.1 Nephrotoxic Drugs

- Antibiotics: aminoglycosides, sulfonamides, methicillin, cephaloridin, polymyxins, etc.
- Non-steroidal antiinflammatory drugs
- Iodinated contrast media
- Immunosuppressive drugs
- Angiotensin-converting enzyme inhibitors
- Chemotherapeutic agents (cisPT, methotrexate, etc.)
- Heavy metals (inorganic Hg + salts, Cd, Fe, As, Bi, Th, etc.)
- Fluorinated anaesthetics
- Dextrans

Sources: From Fillastre, J.P. et al., *J. Pharmacol.* (Paris), 1986, 17 (Suppl. I):41–50 and Thatte, L. and Vaamonde, C.A., *Postgrad. Med.,* 1996, 100:83–100.

Major functions of the kidney

1. To maintain a stable chemical and physical environment for cells (this entails regulation of water, electrolytes and acid/base rations)
2. To excrete metabolic wastes, including:
   - Uric acid derived from nucleic acids
   - Creatinine from muscle creatine
   - Bilirubin derived from hemoglobin
   - Urea nitrogen derived from dietary and endogenous proteins
3. To excrete many foreign chemicals, including drugs
4. To regulate arterial blood pressure (Na^+ and water balance, rennin/angiotensin/aldosterone system, vasoactive PGs, etc.)
5. To produce the active form of vitamin D
6. To degrade or excrete hormones (gastrin, insulin, PTH, GH, CCK, ADH, secretin, glucagons, etc.)
7. Be capable of gluconeogenesis during a prolonged fast
8. To release renal erythropoeitic factor which acts on a liver globulin to produce erythropoietin, in case of decreased O_2 delivery to the kidney

Glomerular filtration rate (GFR) is the best global estimate of renal function. Inulin clearance is the gold standard for GFR measurement. Creatinine clearance may also be used, although its value exceeds the exact value for GFR because creatinine is not exlusively excreted by the tubules but also secreted by the tubules. This discrepancy increases as GFR falls. Serum creatinine measurement does not reflect abnormal renal function until after GFR has been reduced to at least 50% of the baseline value.[1]

Other parameters are of great interest for a global approach of renal safety. These include fractional excretion of Na^+ and K^+, renal blood flow (PAH clearance or ultrasonic transit-time flowmetry), enzymuria (which could allow differential location of toxic injuries), proteinuria, glucosuria, diuresis, and concentrating ability of the kidney (measurement of urine osmolality).

Under normal conditions, GFR is submaximal because of adaptive increases in single nephron GFR following loss of damaged nephrons. Sensitized animal models that mimic risk factors commonly found in patients with drug-induced acute renal failure are advisable. This need should stimulate research in the field of safety pharmacology. The choice of the species, strain, and sex of test animals should take into account physiological and pharmacotoxicological specificities.

Acute renal failure (ARF)

Deterioration of renal failure over a period of hours to days results in the failure to excrete nitrogenous waste products and the inability to maintain fluid and electrolyte homeostasis.

ARF due to toxic or ischemic injury is a clinical syndrome referred to as acute tubular necrosis, and a common disease with high overall mortality (approximately 50%). Little progress has been made in treatment since the advent of dialysis more than 30 years ago.[2] Table 7.2 summarizes the known risk factors for ARF, which should always be kept in mind during drug development, especially in clinical trials.

Functional reserve of the kidney

Adaptive increases in single nephron GFR tend to obscure renal injury until a considerable amount of kidney parenchyma is irreversibly lost. Under normal conditions, GFR is submaximal. Because of the marked amount of functional reserve in the kidney, no decrease is seen in impaired mammal animals until significant impairment has occurred. Such impairments, although not detectable in normal populations, can be quite significant in individuals who already suffer from renal disease. As a result, consideration should be given to the use of animals that have been made more sensitive to functional effects by having their existing capacity reduced. Means of doing this would include surgical removal of one kidney as a measured renal toxic insult.[3,4]

Table 7.3 presents the basic guidance provided by International Conference on Hormization (ICH) for safety pharmacology evaluation of new drugs. These evaluations are noncore and therefore not required prior to first-in-man (FIM) studies. It is essential that intact animals be used in such evaluations, but it must be kept in mind that there are species differences in responses,[5,6] and that renal function is readily influenced by anesthesia.[7,8]

Clearance

Clearance techniques provide a convenient quantitative measure of renal function and are usually based on simultaneous analysis of plasma and urine. The clearance concept possesses the great advantage of permitting assessment of renal function in terms of a physiologically meaningful

Table 7.2 Risk Factors for Drug-Induced ARF

Patient-related risk factors
• Age, sex, race
• Preexisting renal insufficiency
• Specific diseases (diabetes mellitus, multiple myeloma, lupus, diseases associated with proteinuria, etc.)
• Sodium-retaining states (cirrhosis, heart failure, nephrosis)
• Dehydration and volume depletion
• Hyperuricemia, hyperuricosirua
• Sepsis, shock
• Renal transplantation
Drug-related risk factors
• Inherent nephrotoxic potential
• Dose
• Duration, frequency and form of administration
• Repeated exposure
Drug interaction
• Associated use of diagnostic or therapeutic drugs with added or synergistic nephrotoxic potential

Sources: Thadani, R. et al., *New Engl. J. Med.,* 334:1448–1460, 1996 and Lieberthal, W. and Nigam, S.K., *Am. J. Physiol.,* 278:F1–F12, 2000.

Table 7.3 ICH Renal System Primary Endpoints

Renal function	Measurement of effects on urine excretion in saline loaded rats
Renal dynamics	Measurement of renal blood flow, GFR, and clearance

quantity — the virtual volume of plasma cleared of a solute by the kidney in unit time. The significance of the term *virtual* lies in the fact that the same clearance value would be calculated whether a solute is completely extracted from 100 ml plasma per unit time, or only 50% cleared from 200 ml. Clearance, in other words, only corresponds to a discrete physical volume if a single pass through the kidneys leads to complete removal of the solute under study.

Conventionally, renal clearance has been equated to urinary clearance. In this case, the amount of the plasma solute excreted in urine per minute is given by the product of urine flow V (in ml/min) and the concentration of the solute in urine (U). If P denotes the steady concentration of that solute in arterial plasma, then the expression UV/P describes the volume of plasma from which the solute has been extracted into the urine over a period of 1 min. This defines the urinary clearance (C) of a solute as summarized by:

$$C(ml/min) = UV/P$$

Strictly speaking, of course, renal accumulation or metabolism of a plasma solute also constitutes renal clearance from plasma. Thus, the kidney

readily filters plasma Cd-metallothionein but almost completely reabsorbs it at low concentrations; little is excreted in urine[9] and the reabsorbed protein is not returned to blood. A more general expression for renal clearance from plasma clearance in ml/min is therefore given by the product of renal plasma flow (RPF) and E, the percent extraction calculated as (A-V)/A, where A and V represent the arterial and venous plasma concentrations of the solute under study. Note that the simplifying assumption is implicitly made that solute fluxes between red cells and plasma do not affect A or V. The general formula for clearance is thus as:

$$C = E \times RPF$$

Measuring clearance by determination of renal plasma flow and the A-V difference is, of course, more invasive and less convenient than basing the calculation, when possible, on urinary excretion. In further discussion of clearance, we shall, therefore, restrict ourselves to renal clearance as classically defined in terms of urinary excretion of plasma solutes. Note, therefore, that the clearance concept cannot be applied to urinary excretion of a solute like ammonia, which is synthesized in the kidney, or cadmium and whose excretion may result, in part, from release of metal sequestered for long periods of time in the kidney. Always keep in mind that only if the solute were completely removed from plasma during a single passage through the kidneys (i.e., if E above equaled 100%) with subsequent excretion into the urine, would the calculated clearance correspond to a specific physiological parameter, namely the renal plasma flow. However, even for solutes like *p*-aminohippuric acid (PAH) that most closely approach these conditions, extraction from plasma seldom exceeds 90%.

By definition, the clearance of a solute like inulin, which is freely filterable at the glomerulus and which is neither secreted nor reabsorbed in the tubules, must equal the GFR. Note the demand for complete filterability of the quantity of solute expressed by P in the initial equation. Thus, a major fraction of heavy metals in plasma is bound to plasma protein and can therefore not be filtered; accordingly, P in the expression UV/P must be corrected for nonfilterable metal for the clearance calculation to lead to a physiologically meaningful result. In practice, ultrafilterability of a solute circulating in plasma may be difficult to measure. First, properties of pores in synthetic ultrafiltration membranes may differ greatly from those of glomerular pores; a different charge distribution *in vitro* is likely to alter the filtration characteristics of ionic macromolecules from those observed *in vivo*. A second difficulty, sometimes overlooked, is related to the fact that anticoagulants, for instance by chelating heavy metals, may alter their filterability.[10,11]

Renal clearance of extracellular solutes that are not otherwise excreted or metabolized can be calculated from their plasma disappearance curves.[1] Such techniques routinely utilize isotopically labeled compounds and possess the advantage of not requiring either urine collections or prolonged equilibrating infusions. When urinary clearance is conventionally

determined from the U/P ratio, the arterial plasma concentration of the solute should ideally have reached a steady state. For an exogenous solute, the classical clearance technique, therefore, requires as long as 1 h of equilibrating infusion, often preceded by a priming injection. Alternatively, and more conveniently, giving an intramuscular or subcutaneous depot injection can achieve reasonably stable plasma levels of solutes. Even if absolute constancy is not achieved in this manner, a mean concentration is slow and constant.

The steady-state solute concentration, however achieved, must always lie well below the half-saturating concentration (K_m) of any nonlinear reabsorptive or secretory process in the renal tubule. A well-known illustration of this fact is the glucosuria of diabetes mellitus, which reflects not damage to the proximal tubules but saturation of their capacity to transport the excess sugar. Another prerequisite for significant clearance value is a urine flow that is sufficiently high enough to permit essentially complete collections during accurately timed clearance periods. Even carefully controlled experiments in the laboratory cannot always achieve such ideal conditions. As a result, renal clearance values seldom possess confidence limits of better than ± 10%. Such accuracy may be adequate for most purposes; it implies, however, that as much as 20% depression of a renal solute clearance following exposure to a compound believed to act directly on the kidney does not necessarily indicate a nephrotoxic effect.

Endogenous creatine is a special case. This is normally produced at a constant rate as the end product of muscle metabolism and, therefore, circulates in plasma at a constant concentration; it is freely filtered at the glomerulus, but neither reabsorbed nor effectively secreted in the tubule. Secretion has been reported in some species, but it is sufficiently small so that it does not significantly contribute to creatine excretion except at low filtration rates. Similar to inulin clearance, that of creatine can serve as a convenient measure of the GFR. Creatinine possesses a great advantage over insulin in that its clearance can be measured without an equilibrating infusion.[11,12]

Measurement of the renal clearance of an exogenous solute by conventional techniques consumes relatively large amounts of this solute in order to reach and maintain the required steady plasma levels; this is especially true of determinations of maximum tubular transport activity T_m at high plasma concentrations (see the section on Measurement of T_m below). Factors such as cost of the solute, time required for its equilibration, toxicity, radiation exposure in the case of labeled solutes, and metabolic breakdown or transformation by, for instance, transamination of an infused amino acid into a competing amino acid, all constitute disadvantages of normal T_m measurements.

These problems can be minimized with the use of an arterial gradient infusion procedure in which the arterial plasma concentration of the solute in strongly diuresing animals is rapidly raised in a stepwise fashion. It can then usually be kept approximately constant for several minutes by decreasing the

rate of infusion in a manner equal and opposite to the rate at which recirculation would have raised plasma levels at a constant infusion rate. It is possible, in this manner, to carry out a complete clearance determination in 3 to 4 min.[13] The procedure has been successfully applied to dogs, rabbits, and rats, but has not been widely adopted, perhaps because of occasional difficulties encountered in maintaining constant plasma concentrations.

Clearance determinations discussed so far require measurement of concentrations in carefully timed urine and plasma samples. In addition, useful approximations to relative solute clearance values can be obtained by simplified procedures. The best known of these simply takes the plasma level of urea, or preferably creatinine, as a measure of the GFR. Indeed, if creatinine excretion (UV in g/day) is constant, the GFR (= UV/P) theoretically is inversely proportional to P_{Cr}, the creatinine concentration is plasma; any increase in P_{Cr} above a normal level of around one mg/dl should, therefore, reflect a corresponding fall in GFR. In practice, the method is not very sensitive in the normal range of plasma creatinine levels (< 1.4 mg/dl); a better correlation between measured creatinine clearance (C_{Cr}) and that predicted on the basis of P_{Cr} is obtained at higher plasma levels (i.e., lower C_{Cr} values).[1,14]

Because its rate of production is influenced by changes in protein metabolism and its excretion by the rate of urine flow, the level of urea in plasma varies more widely than that of creatinine. Although it is, therefore, a less reliable indicator of glomeruler function than is plasma creatinine, plasma urea is frequently used to follow changes in GFR. Note that blood and plasma or serum urea are used interchangeably because this solute freely diffuses into and out of blood cells.

Considerable information can also be obtained on changes in clearance values by analysis of only randomly collected urine. This approach is also based on the characteristics of creatinine excretion, as discussed previously, and consists of normalizing excretion of a solute by that of creatinine. A change in the solute/creatinine concentration ratio in urine then reflects a corresponding change in the excretion of the solute under study.

Normalization of urinary solute excretion on the basis of creatinine — or, less frequently, on the basis of urinary specific gravity — possesses other advantages. Thus, normalization significantly diminishes the large fluctuations often encountered in urinary solute concentrations. It also justifies, as discussed earlier, the analysis of spot samples of urine, thus avoiding the inconvenience of having to collect 24-h urine samples. Because of diurnal variations in the GFR, the spot samples should be collected at the same time each day.

Gross changes in renal plasma flow, or a reduction in the efficiency (E) of extraction of secreted solutes from blood due to cytotoxic effects, can also be demonstrated by changes in the time required for excretion of a standard injected dose of compounds like ortho-iodohippurate, the dye phenol red (PSP), and x-ray contrast media like Diodrast™. The normal kidney, with a high value of E, efficiently extracts each of these substances so that their

excretion provides a composite measure of effective renal plasma flow, total tubular mass, and the functional integrity of the tubule cells. Some elegant methods are now available for measuring renal blood flow, such as ultrasound flowmetry.[15] Additional information on changes in tubular mass and activity can be obtained from determination of maximum tubular capacity T_m as discussed below.

A rapid and powerful approach to the study of renal function in control or poisoned animals is the double indicator dilution technique, introduced into renal studies by Chinard.[16] It consists essentially of a rapid arterial bolus injection of a test solute, together with a glomerular and extracellular marker such as creatinine, followed by rapid collection of sequential urine or renal venous blood samples. Although the procedure is too invasive for routine or other than terminal function assessment, it can provide extensive information on the integrity of solute transport processes and on their mechanisms, as well as on the site of action of various inhibitors.

The indicator dilution technique possesses the disadvantage of requiring arterial injection and ureteral cannulation in the anesthetized animal. On the other hand, it is rapid and consumes relatively little solute, thus permitting convenient and repeated measurements of relative clearances. Care must be taken in all renal transit experiments to use sufficiently low bolus concentrations of the test solute so that even at its peak plasma concentration renal transport mechanisms do not become saturated.

Saturation of a reabsorptive system will, by definition, increase the fractional excretion of its substrate. As an example of the important use to which such observations can be put, reference may be made to the report that at high bolus concentrations of metallothionein (MT) the characteristics of its transit from artery to urine approach those of insulin. The implications of this finding are that MT is freely filterable and that it is normally reabsorbed by a saturable mechanism. In addition, obviously little nonfiltered MT can have been transported directly from blood into urine (secretion), a conclusion further confirmed by the similar renal vascular transit curves of insulin and of MT.

Free water clearance and renal concentrating ability

One of the most important functions of the normal kidney is its ability to respond to changes in the state of hydration by excreting dilute or concentrated urine. Appropriate water excretion or retention provides a very sensitive measure of general renal integrity.

The simplest measure of the ability of the kidney to conserve water is provided by osmolality of urine excreted by a water-deprived subject. The osmolality of plasma changes relatively little and averages close to 300 mOsm; analysis of the osmolar concentration of urine thus leads directly to the (U/P) osmol ratio. Simple and convenient osmometers are available for such determinations. Human urine may range from osmolalities below 100 mOsm to a maximum of perhaps 1800 (i.e., to a urine sixfold more concentrated than plasma). The rat, in contrast, can readily concentrate its urine nine- to tenfold.

A more quantitative measure of urine concentration ability is the clearance of free water (C_{H_2O}), derived in turn from the osmolal clearance, (UV/P) osmol. The latter, in effect, represents the volume that would be required to excrete total urinary solute in isosmotic solution (i.e., at a concentration of 300 mOsm). If filtrate were excreted as such, the osmolal clearance would equal the GFR. If the urine is more dilute than the filtrate, the dilution may be visualized as that volume of filtrated from which solute was removed without reabsorption of water; this volume is defined as the free water clearance (C_{H_2O}), or free water excretion, and is given next, where V stands as usual for the urine volume excreted per unit time:

$$C_{H_2O} = V - (UV/P)\text{osmol}$$

Dilution of urine (i.e., reabsorption of solute without water), primarily occurs in the so-called diluting segment in the thick ascending loop of Henle. The water permeability of this segment is very low and as NaCl is reabsorbed, the remaining fluid is rendered dilute. Additional solute reabsorption occurs in the more distal nephron and contributes to a high interstitial solute concentration in the medulla. In the presence of antidiuretic hormone, water can move out of the lumen along the osmotic gradient, so that the final osmality of the urine reflects the osmolal concentration in the medulla. Excretion of concentrated urine, by analogy with the formation of dilute urine discussed previously, implies that a certain volume of solute free water must have been removed from the isosmotic filtrate, in addition to the large fraction of filtered water and solute normally reabsorbed isosmotically in the proximal tubule. In this case, the clearance of solute-free water is negative, and appropriately referred to as negative free water clearance; it is calculated as (UV/P)osmol-V.

Several intrarenal factors are involved in the production of concentrated urine, including especially the attainment and maintenance of high interstitial osmolality in the medulla as a result of solute filtration and reabsorption and of the countercurrent concentrating mechanism, which depends, in turn, on normal medullary blood flow. Urinary concentrating ability can, therefore, not be associated with a specific aspect of renal function, but serves as reflection of its general integrity and can thus prove useful for screening purposes.

Two standard materials are used in the evaluation of clearance: insulin, which is freely filtered and neither reabsorbed nor secreted into the tubules, and creatinine, which is not exclusively filtered, but also secreted into the tubules. Creatinine's clearance may exceed the exact value for GFR, and this discrepancy increases as GFR falls. We may, therefore, consider the advantages and disadvantages of current clearance measurement methods in animals:

Advantages

- Can be used in about all species
- No need for anesthesia or surgery that may alter renal function
- Link between animal and human studies
- Functional approach

Disadvantages

- Cannot distinguish internephron variation
- Depends upon circadian changes in rats and humans

Renal blood flow

Plasma clearance of para-aminohippuric acid (PAH) is almost equal to a plasma flow.

PAH is secreted into the tubules on a single pass and 10 to 15% of the RBF goes to nonfiltering portions of the kidney. This nonfiltered plasma cannot therefore lose its PAH, and thus PAH clearance = effective renal plasma flow = .5–90% of total RPF.

Fractional excretion of sodium

The percentage of filtered Na^+ load excreted in the final urine. Also, the clearance of Na^+ expresses as a% of GFR: FE_{Na} (%) = (U_{Na} x V)/(P_{Na} x GFR) x 100

Theoretically, if drug-induced damage of the medullopapillary portion of the nephron occurs, decreases in reabsorption of Na^+ and water lead to an increase in FE_{Na}.

Clinical chemistry measures

Enzymuria

Enzymuria depends upon:

- Circadian and infradian rhythms
- Urinary flow rate
- Urinary pH
- Age
- Sex
- Environmental pollution

Proteinuria

In general, albuminuria leads to increases in the permeability of the glomerular capillary wall, whereas urinary release of low-molecular weight proteins (β-2-mg) suggests an impairment in tubular reabsorption.

Glucosuria

If there is normal serum glucose concentration, it may reveal proximal tubular damage (such as with gentamicin and maleic acid).

Table 7.4 Clinical Chemistry Measures of Specific Renal Effects

Origin or urinary enzymes	Enzymes
Brush border	• Alanine-aminopeptidase • γ-glutamyl-transferase • Trehalase
Lysosomes	• β-glucuronidase • N-acetyl-β-D-glucosaminidase • Acid phosphatese • β-galactosidase
Cytosol	• Lactate dehydrogenase • Leucine aminopeptidase • β-glucosidase • Fructose-1,6 biphosphatase (proximal tubule) • Pyruvate kinase (distal tubule)

Urine concentration test

Food and water are withheld for 24 h. Osmolality of the collected urine is measured. Some toxic agents: ↓ maximal concentrating capacity (cis-dichlorodiammineplatinum, etc.)

Animal models

The rat

Although there are no specific scientific reasons to select the rat for evaluating renal safety, two factors should be kept in mind:

1. Easy handling
2. Availability of numerous strains with metabolic and/or physiological specificities: BB; Munich–Wistar; Lewis–DA; hypertensive salt-sensitive Dahl strain; SHR

When comparing GFR and RPF between rats of different strains, it is appropriate to correct these parameters for body weight because they are

Table 7.5 GFR in Healthy and Diabetic Rats

	Control rats[a]		Diabetic rats[b]	
	Conscious	Anaesthetized	Conscious	Anaesthetized
Inulin clearance ml/min	3.2 ± 0.2	1.8 ± 0.2	3.6 ± 0.2	3.1 ± 0.6
Inulin clearance ml/min/kg BW	10.9 ± 0.05	6.8 ± 0.8	15.9 ± 0.8	10.9 ± 1.7

[a]Sprague–Dawley female rats: BW 270–300 g; anesthesia: Inactin 100 mg/kg.

[b]Diabetic rats: streptozotocin, single dose 60 mg/kg, 2–4 months prior to the study.

genetically correlated, but high within-strain variation of kidney weight not correlated with renal function; do not correct GFR for kidney weight.

The dog

- Ethical limitations
- Easy surgical accessibility to kidney and vessels
- Single nephron GFR = 60 nl/min closer to that of humans (60–65 nl/min) than that of rats (about 33 nl/min)
- Renal medulla = simple type with small, cone-shaped vascular bundles as in humans

Cautions:

- All anesthetics induce hemodynamic changes.
- Barbiturate can reduce GFR by 50% in rats.
- Thiobutabarbital (Inactin) impairs RBF autoregulation in dogs and rats; pentobarbital does not.

One example is Wistar rats,[17] where the following procedure is used:

- Salt depletion: After four daily injections of furosemide (2 mg/kg), rats are fed with boiled rice (Na^+ 2 mg/100 g) for 7 d before test-compound.
- Indomethacin 10 mg/kg 1 h before test-compound.
- Uninephrectomy 3–5 weeks before test-compound.
- Dehydration the day of test-compound injection.

Another example is that dehydration greatly enhances the sensitivity of rats to the nephrotoxicity of aminoglycosides.

Examples of species differences in drug sensitivity

Ethacrynic acid
- Drastic diuretic in humans, monkey, dog, mouse
- Lesser effect in rats

Ouabain
- Effective inhibitor of Na^+/K^+ ATPase in humans, dog
- Less effective in guinea pig, rabbit; very low activity in rat

Amino-glycosides
- Higher nephrotoxic threshold in animals (10- to 60-fold) than in humans

Iodinated contrast media
- Less toxic in rats than man

References

1. Cockroft, D.W. and Gault, M.H., Prediction of creatinine clearance from serum creatinine. *Nephron.*, 1976, 16:31–41.
2. Brezis M., Rosen S., and Epstein, F.H., Acute renal failure, in *The Kidney*, Brenner, B.M. and Rector, F.C., Eds., Philadelphia: W.B. Saunders, 1991, pp.993-1061.
3. Idée, J.M., Renal safety pharmacology: value of sensitised experimental models, *Thérapie.*, 2000, 55:91–96.
4. Lauwerys, R.R. and Bernard, A., Early detection of the nephrotoxic effects of industrial chemicals: state of the art and future prospects, *Am. J. Ind. Med.*, 11:275–285, 1987.
5. Chiu, P.J.S., Models used to assess renal function, *Drug Dev. Res.*, 1994, 32:247–255.
6. Craddock, G.N., Species differences in response to renal ischemia. *Arch. Surg.*, 111:582–584, 1976.
7. Conger, J.D. and Burke, T.J., Effect of anesthetic agents on autoregulation of renal hemodynamics in the rat and dog, *Am. J. Physiol.*, 230:652–657, 1976.
8. Buchardi, H. and Kaczmarczyk, G., The effect of anaesthesia on renal function. *Eur. J. Anaesth.*, 11:163–168, 1994.
9. Nomiyama, I.K. and Foulkes, E.C., Reabsorption of filtered cadmium metallothionein in the rabbit kidney, *Proc. Soc. Exp. Biol. Med.*, 156:97–99, 1977.
10. Foulkes, E.C., Mechanisms of renal excretion of environmental agents, in Lee D.H.K., Ed., *Handbook of Physiology, Section 9: Reactions to Environmental Agents.* Washington, D.C.: American Physiological Society, 1977, 495–502
11. Harvey, A.M. and Malvin, R.L., Comparison of creatinine and inulin clearances in male and female rats, *Am. J. Physiol.*, 209:849–852, 1965.
12. Tepe, P.G. et al., Comparison of measurements of glomerular filtration rate by single sample, plasma disappearance slope/intercept and other methods, *Eur. J. Nucl. Med.*, 13:28–31, 1987.
13. Foulkes, E.C., On the mechanism of chlorothiozide-induced kaliuresis in the rabbit, *J. Pharmacol. Exp. Ther.*, 406–413, 1965.
14. Jobin, J. and Bonjour, J.P., Measurement of glomerular filtration rate in conscious, unrestrained rats, *Am. J. Physiol.*, 248:F734–F738, 1985.
15. Evans, R.G. et al., Chronic renal blood flow measurement in dogs by transit-time ultrasound flowmetry, *J. Pharmacol. Toxicol. Meth.*, 38:33–39, 1997.
16. Chinard, F.P., Relative renal excretion patterns of *p*-aminohippurate (PAH) and glomerular substances. *Am. J. Physiol.*, 185:413–417, 1956.
17. Heyman, S.N. et al., Acute renal failure with selective medullary injury in the rat, *J. Clinical Invest*, 82:401–412, 1988.
18. Fillastre, J.P. et al. Détection de la néphrotoxicité médiicamenteuse, *J. Pharmacol.* (Paris), 1986, 17 (Suppl. I):41–50.
19. Thatte, L. and Vaamonde, C.A., Drug-induced nephrotoxicity, *Postgrad. Med.*, 1996, 100:83–100.
20. Thadani, R., Pascual, M., and Bonventre, J.V., Acute renal failure, *New Engl. J. Med.*, 334:1448–1460, 1996.
21. Lieberthal, W. and Nigam, S.K., Acute renal failure, II. Experimental models of acute renal failure: imperfect but indispensable, *Am. J. Physiol.*, 278:F1–F12, 2000.
22. Hierholzer, K., Value and necessity of animal research in nephrology, *Expl. Biol. Med.*, 1982, 7:88–101.

chapter 8

The gastrointestinal system

Alterations in gastrointestinal (GI) function are common side effects of many drugs, not limited to those administered orally.[1] Changes in bowel habits (constipation or diarrhea) and gastric mucosal irritation are the most common side effects covered, at least in part, by the classical technical approach involving the evaluation of gastric emptying, intestinal and colonic transit, as well as direct gastric mucosal damage score evaluation.[2,3] For example, increased small intestinal permeability caused by nonsteroidal antiinflammatory drugs (NSAIDs) is probably a prerequisite for NSAID enterpathy, a source of morbidity in patients with rheumatoid arthritis. These results were supported by the 51Cr EDTA/L-rhamnose urine excretion ratios, which reflect changes in intestinal permeability.[4]

Effects of drugs on intestinal tone has appeared recently to be of importance since compounds having a relaxatory effect on colonic muscular tone may favor the occurrence of ischemia. There is now evidence that other functions of the gut may be affected, however, giving rise to more chronic alteration and/or reactivity to oral pathogen or irritant or alteration in digestion and presence of gut hypersensitivity to mechanical stimuli.[4] Several other gut functions, such as epithelial barrier regulating both transcellular and paracellular absorption, may be altered as well as enzymes, hormones, and ion secretion. Intestinal and colonic microflora may be also affected by orally administered drugs. Visceral sensitivity of the gut is affected by the immune status of the mucosa. Evidence has accumulated that orally administered drugs may modify this immune status giving rise to sensitization of sensory nerve terminals, but a direct action after epithelial absorption on receptors located on terminals of primary afferent neurons is possible. These effects may trigger abdominal pain.

Increased paracellular permeability such as that induced by the majority of nonselective NSAIDs favors the entry of pathogens, allergens, and bacterial translocation, giving rise to enteropathy and septic shock. Simple tests performed in animals by measurement of permeability to macromolecules in basal and stimulated conditions may provide relevant information. Similarly, drugs may directly affect both functional and nociceptive

Table 8.1 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

sensitivity of the gut and subsequently inhibit enzyme secretion, many upstream inhibitory reflexes, and initiate abdominal pain in response to normal mechanical stimuli.

Not only antibiotics but also other drugs orally administered may have an effect on the colonic microflora, here again, altering the immune balance through the gut mucosa. Substances administered orally may also affect the population of immunocytes within the gut. Among them, mast cells are the most often affected population. This effect may be related or unrelated to sensitization through gut allergic reaction. Increased numbers of mast cells or their contents may also contribute to sensitization of primary afferent terminals, here again affecting gut sensitivity. All this information suggests that it should be relevant to test the influence of new drugs on colonic sensitivity to distension through classical models used to evaluate the sensitivity of the gut to luminal distension.

Finally, safety pharmacology, particularly for drugs administered orally, also must evaluate the effect of drugs on other important functions of the gut with specific attention to paracellular permeability, enzyme secretion, immune status of the mucosa, and sensitivity.

Drug-induced alterations of GI transit or motility

- Gastric emptying: → Dyspeptic symptoms
  - Early satiety
  - Postprandial fullness
  - Regurgitations, nausea
- Intestinal transit: → IBS symptoms
  - Abdominal fullness
  - Gas retention
  - Abdominal pain
- Colonic transit: → IBS symptoms
  - Constipation
  - Diarrhea
  - Bloating, abdominal pain

Colonic tone is the association of reduced colonic muscular tone and the slowing of colonic transit may sometimes initiate ischemic colitis (e.g., 5-HT3 antagonists-alosetron).

GI function

Drugs can affect the following GI functions:

- Transit (motility)
- Absorption (trans and paracellular transport)
- Digestion (enzymes)
- Secretion (ions, hormones, enzymes)
- Microflora (colon)
- Immunity (mucosa)
- Viscerosensitivity (abdominal pain)

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/or gastric emptying in the rat are examples recommended for safety pharmacology evaluation. Intestinal motility — assessed by the transit of carmine dye in the mouse and gastric motility, assessed by stomach weight in the rat — was examined using a range of clinical drugs or potent pharmacological agents known to affect gastrointestinal 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 gastrointestinal 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 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.[4a] 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 intravenous 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 Mason Laboratories, we tested the ability of a parasympatholytic agent, intravenous atropine sulfate, to inhibit GI transit. In a dose-dependent fashion, 30 and 50 mg/kg atropine sulfate significantly decreased the distance traveled by the charcoal meal.

Appropriate methodology to evaluate effects on GI transit and motility include:

- Gastric emptying
  - Nutritive radiolabeled meal
  - Solid and liquid phases
  - Rats, dogs
- Intestinal transit
  - Radio-labeled meal
  - Determination of the geometric center for 10 equal intestinal segments
- Colonic transit
  - Rats with intracolonic catheter
  - Accustomed to eat in a given time (3–4 h)
  - ^{51}Cr-Na administered intracolonically
  - Collection of feces by 30 min period during 24 h

Assessment of intestinal transit

A second major function of the intestinal tract depends on its contractility. Intestinal motility is responsible for appropriate mixing of ingested materials with endogenous secretions required for digestion and for appropriate delivery of substances to the site of absorption elimination.

One approach to the study of intestinal motility *in vivo* is the use of an intraluminal marker substance whose transit through the gut lumen can be quantitated.[5] Substances used as valid markers should not be absorbed from the intestine, nor adsorbed onto the mucosal surface. In additions, ideal markers do not affect any aspect of intestinal function. Furthermore, since the contractile properties of the intestinal musculature may have different effects on solid as opposed to liquid components of the gut contents, markers should be chosen to correspond to the physical composition of the endogenous substance of greatest interest. This point is especially relevant to the study of gastric emptying.[6] Markers used in the tracing of solid substances include, for example,^{99m}Technitium incorporated into a chicken liver meal.[6] Markers such as this gamma emitter can be readily monitored for stomach content in humans using a gamma camera. Other solid markers, used primarily in analysis of total gastrointestinal transit time in humans, include raiopaque pellets made from polythene impregnated with barium sulfate.[7]

A technique developed by Summers et al.[8] permits investigation of intestinal transit in animals by administering markers into the duodenum. This procedure allows interpretation of effects on intestinal transit without any influence of gastric emptying. Permanent, indwelling catheters are surgically implanted in the duodenum of rats and exteriorized behind the head. Markers such as ^{51}Cr can be administered directly into the small intestine without the disruptive effect of anesthesia or oral intubation procedures. In animal studies, the content of markers in the intestine can be determined by analyzing sequential gut segments after the animals have been killed.

Other approaches to the *in vivo* analysis of the motor function of the gut[27] include determination of intraluminal pressure changes. These changes

can be measured using small balloons or fluid-filled open-tipped tubes connected to external strain gauges, or internal miniaturized strain gauges monitored by telemetry or by means of exteriorized wires.[9] Analysis of the function of the intestinal musculature may also entail measurement of its electrical activity, which in a chronic *in vivo* preparation can be carried out with a recording electrode surgically implanted on or in the intestinal wall or intubated into the intestinal lumen using a balloon to ensure its juxtaposition to the mucosal surface.[10]

The contractile properties of intestinal smooth muscle can also be assessed using *in vitro* preparations of this organ. Such *in vitro* techniques are valuable in screening potentially toxic substances for effects on intestinal contractility and for elucidating mechanisms of effect on propulsion observed with *in vivo* methodology. Use of *in vitro* techniques to study intestinal motility has certain advantages over *in vivo* procedures. Generally, they are technically simpler to execute. They isolate the tissue from extrinsic neural and hormonal influences. The tissue can be directly exposed to the test substance. These advantages are at the expense of loss of prediction of *in vivo* effects of a test substance.[11]

The choice of a particular *in vitro* technique depends on the specific aim of the experiment. Techniques most commonly used differ in several ways. First, the species from which the intestinal segment is taken markedly affect the basal contractile activity. The rabbit jejunum, for example, maintains rhythmic contractions *in vitro* and, therefore, is especially useful for analysis of substances suspected of having inhibitory effects on intestinal smooth muscle. The guinea pig ileum, in contrast, exhibits little spontaneous activity *in vitro*. This preparation is, therefore, widely used in the bioassay of agents causing contraction of intestinal smooth muscle. To test for depressant effects, the investigator must induce contraction of this tissue as with electrical stimulation. Second, there are differences in the responses of the smooth muscle, depending on the site within the intestine under investigation. This limits the investigator's ability to generalize from an experiment carried out with a muscle preparation from a single region of the intestine and reinforces the importance of strictly controlling the tissue region studied in a series of experiments.

Determination of intestinal absorption

Study of the absorptive function of the intestinal tract can be carried out with numerous methods. Among the primary considerations in choosing an experimental technique to assess an aspect of the absorption process are the following:

1. The test species to be used (i.e., Must the study be conducted in humans with all the accompanying complications, or is there an appropriate experimental animal model?)
2. The aspect of the absorption process of interest (e.g., Is it the overall absorption from the gut lumen to the systemic circulation and tissues,

or is it the process of transport across the brush border or basolateral membrane of the intestinal mucosal cell?)
3. Which experimental or physiological variables should be controlled (e.g., Control the presence of anesthetic agents, the electrochemical potential difference across the gut wall, or the pH of the luminal gut contents?)

The answers to these questions will determine the particular method that may be chosen from among the *in vitro* methods available for studying absorptive function.[12,13] These methods can be categorized according to (1) the method by which the test substance is administered, and (2) the method for assessing the extent and/or rate of absorption.

Methods of administering test substance

Among the various techniques for administering the test drug in an *in vivo* study are the following:

1. Incorporating the test substance into the diet, which is then administered to the subject. This procedure may be especially relevant to the analysis of the absorption of drugs that are contaminants of the diet.
2. Intubating the test substance into the stomach, a procedure that allows more precise control of the total dose administered. With both these methods, the rate and possibly the extent of absorption of the test substance may be markedly affected by the gastric emptying pattern of the subject.
3. Directly administering the test substance into the intestinal lumen, which eliminates the influence of gastric emptying. In human studies, substances can be administered through small-bore intubating tubes localized to particular sites by radiographic techniques.[14] In animal experimentation, test substances may be perfused through the gut lumen as a single pass, analogous to the perfusion method in man, or recirculated in the perfusate. Such a technique requires cannulation of the intestine, an external heating device for maintaining the perfusate at body temperature, and a pump for maintaining constant flow. An advantage of perfusion procedures, over techniques described below, is that the influence of flow rate on absorption kinetics can be directly determined. Analyses have indicated that for many substances, such as long chain fatty acids, bile acids, and cholesterol,[15] diffusion through an unstirred water layer overlying the mucosal surface is a rate-limiting step in the overall absorption process.

Another method of direct administration in experimental animals consists of placing a test substance into a segment of the intestine that is closed by ligatures both proximally and distally.[16] The construction of the closed segment and the injection of the test dose do require the use of anesthetic

agents. However, the animal, typically small animals such as rats, can be allowed to recover from anesthesia and to become ambulatory for the majority of the absorption period. This procedure has the additional advantage of not requiring perfusion pumps or heating devices.

Substances may also be administered directly into the intestinal lumen following surgical creation of exteriorized fistulas.[12] Studies are then carried out in unanesthetized larger animals, typically dogs. This approach, of historical importance, has the disadvantage of requiring considerable surgical manipulation.

Methods for quantitating degree of absorption

Appearance in systemic fluids

In addition to the method of administering the test substance, the second critical aspect of a technique is the sampling procedure for quantitating the extent and/or rate of absorption. With *in vivo* methods, the least invasive techniques entail the collection of blood, urine, or breath samples for determining the appearance of the absorbed test substance and its metabolites in body fluids. Comparing the time course of plasma concentrations or excretory rates in urine or breath after oral administration with the results after intravenous administration may permit quantitation of the extent of absorption of the test substance and the rate constant of this process.[17] This approach is relatively imprecise and may be confounded by numerous factors such as the first pass effect, the enterohepatic circulation of the agent, and the status of elimination processes such as hepatic and renal function. Nevertheless, this technique is useful in the diagnosis of malabsorption syndromes associated with gastrointestinal diseases.[18] For example, one test of the transport capacity of the small intestine for carbohydrates entails oral administration of the pentose sugar D-xylose followed by its determination in plasma or in urine. Similarly, assessment of intestinal lactase, the disaccharide that cleaves lactose into the absorbable sugars glucose and galactose, involves an oral lactose load followed by determination of blood sugars. A test of fat absorption can include determination of serum carotene. One test of ileal absorptive function entails the oral administration of radio-labeled vitamin B_{12} with determination of its urinary recovery.

Another approach for analyzing intestinal function is the sampling of excretory products in breath.[19] This technique entails analyzing breath for hydrogen — which is generated by the body exclusively by the action of intestinal bacteria on unabsorbed carbohydrates or for carbon dioxide, which is derived from metabolism of an orally administered isotopically labeled test substance — can be used to detect bilary, pancreatic, and mucosal cell malfunction as well as bacterial overgrowth in the small intestine. For example, the bile salt glycocholic acid is normally absorbed intact from the ileum and reexcreted in the bile. However, in patients with impaired ileal function or with bacterial overgrowth in the small intestine, there is increased bacterial deconjugation of glycocholic acid with release

of glycine. Glycine is then metabolized to CO_2, primarily by bacterial enzymes. Consequently, the administration of glycocholic acid, labeled in the glycine moiety with ^{14}C, results in an increased excretion of $^{14}CO_2$ in patients with ileal disease or bacterial overgrowth. Generally, clinical validation of CO_2 breath tests has been carried out using ^{14}C radioisotopes; however, use of the stable ^{13}C analogs with quantitation of $^{13}CO_2$ by mass spectroscopy also has been employed.[20]

The technique of administering an oral load and of sampling body fluid not only has clinical diagnostic importance, but is also a useful approach for determining the overall rate and extent of absorption of environmental contaminants. Such determinations may be important in the theoretical prediction of systemic concentrations of toxic substances following various ingestion rates. Analyses of this sort, referred to by the recently coined term toxicokinetics, apply mathematical tools extensively used to describe the disposition of pharmacologic agents.

A more direct approach for analyzing absorption characteristics than the sampling of systemic or excreted body fluids entails the sampling of portal blood[21] or the collection of the mesenteric blood draining the sites of absorption of the test substance. Such a procedure requires considerably more complicated surgical techniques than that of sampling systemic blood, urine, or breath. Furthermore, transfusions of blood into the animal may be required. An important advantage of this procedure is the capacity to determine *in vivo* the kinetics of metabolism of a test substance by intestinal tissues. In certain studies, the appearance of the test substance in lymph may be critical, as in the absorption of fats. Cannulation of the mesenteric lymphatic vessel may be carried out even in a small animal such as the rat.[22]

Another approach to quantitating absorption entails monitoring the disappearance of a test substance from the intestine after its administration into the lumen. The perfusion method, for example, monitors differences in the amount infused from the amount appearing at a site distal to the area of infusion. Such perfusion techniques have been exceedingly useful in determining the transport of electrolytes and nutrients in man. The determination of the amount unabsorbed in the sample taken at a distal site is made possible by the use of a marker substance that is neither metabolized nor absorbed, commonly polyethylene glycol 4000. This particular marker has the advantages of lack of adsorption to the gut, high water solubility, stability during frozen storage, and ease of determination by radioisotopic or spectrophotometric methods.[23] However, one drawback to techniques in which only the luminal contents are sampled is that retention of the test substances in the intestinal mucosa is not quantitated.

In the closed segment procedure referred to earlier, the extent of absorption is calculated on the basis of disappearance of the test substance from both the lumen and the intestinal tissue. With this method, at the end of the absorption period, the entire segment, both intestinal wall and contents, is assayed quantitatively for the amount of the test substance remaining. A disadvantage to this technique, when compared

with perfusion methods, is that sequential samples cannot be taken from a single animal; however, this method is readily used in small animals and can therefore be relatively economical.

In both the perfusion and the closed segment procedures, equating loss of a test substance with its absorption requires verification that disappearance does not result as a consequence of metabolism in the intestine. If metabolism of the substance does occur, then assay of the intestine alone is inadequate for a description of its absorption kinetics, unless the metabolite is poorly absorbed and can be completely recovered in the intestinal samples.

Another important safety assay of the gastrointestinal system is the influence of the test article on the formation of ulcers.[24] 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 gastrointestinal mucosa.[25]

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 at a preset time after administration of the test article.[26] 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 with 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.[27,28] 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.[28]

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

correlate with the movement of the HC through the different regions of the GI tract.[29] 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 four 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 the 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.[29,30] Although it was 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.[29] 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.[31] Dogs are readily trainable to accept pills and to wear a receiver belt, and could be tested after administration of a test compound.[32,33] 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 gastrointestinal system could also be measured with the HC.[34] The major drawback for using the HC for safety pharmacology screening is the price of the capsules and the receiver system.

Effects of drugs on gut immune system (jejunum, ileum, colon)

- Histological damage score
  - Macroscopic
  - Microscopic
- Resident and attracted immunocytes
  - Mast cell numbers/degranulation *in vitro*
  - Neutrophil activation
    - Tissue myeloperoxidase (MPO)
    - Fecal calprotectin
- Lymphocyte infiltration
- Cytokine profile
    - TH1 (IL2, IFNγ)/Th2 (IL4, IL5)
    - IL1β, TNFα

Candidate drugs to evaluate for effects on gut immune system

- PDE inhibitors (PDE III & IV):
  - Rolipram:
    - Gastric glandular mucosa inflammation

    - Reduced neutrophilic attraction (colon)
    - Intestinal edema
    - Increased mucus secretion
- Other candidates:
    - Kinase inhibitors (Rho and MAP kinase)
    - Herbal preparations
    - NO (nitrous oxide) donors

Specific methods also exist for using the rat to evaluate the potential of drugs to cause or influence GI inflammation and ulceration.[25]

Conclusions

1. GI side effects of drugs are frequent and not limited to alterations in GI transit or motility or histological damage observed in toxicology studies.
2. Influence of drugs on mucosal barrier and particularly on paracellular permeability may have long-term effects on the gut immune system and immune equilibrium with colonic microflora (e.g., food allergy or intolerance, leakage of immunocytes, macromolecules, bacterial infiltrations).
3. Direct or indirect effects of drugs on sensitivity may trigger chronic symptoms similar to FBD and particularly IBS affecting the quality of life.

References

1. *Compendium of Pharmaceuticals and Specialties*, 30th ed. Ottawa: Canadian Pharmaceutical Association 2002.
2. Bueno, L., Gastrointestinal safety pharmacology: exploring more than just gastrointestinal motility, *MDS Pharma Safety Pharmacology Symp.*, Lyon, France, December 2002.
3. Mortin, L.I., Horvath, C.J., and Wyand, M.S., Safety pharmacology screening. practical problems in drug development, *Int. J. Toxicol.*, 16:41–65, 1997.
4. Bjarnason, T. et al., Importance of local vs. systemic effects of non-steroidal antiinflammatory drugs in increasing small intestinal permeability in man, *Gut*, 32(3):275–277, 1991.

4a. Janssen, P.A.J. and Jageneau, A.H., A new series of potent analgesics: Dextro 2: 2-diphenyl-3-methyl-4-morpholino-butyrylpyrrolidine and related amides, *J. Pharm. Pharmacol.* 9:381–400, 1957.

5. Lacroix, P. and Guillaume, P., Current Protocols in Pharmacology, 5.3.1–5.3.8, Gastrointestinal Models: Intestinal Transit and Ulcerogencic Activity in the Rat, New York: John Wiley & Sons, 1998.
6. Lavigne, M.E. et al., Gastric emptying rates of solid food in relation to body size, *Gastroenterology*, 1978, 74:1258–1260.
7. Hinton, J.M., Lennard-Jones, J.E., and Young, A.C., A new method for studying gut transit times using radio-opaque markers. *Gut*, 10:842–847, 1969.

8. Summers, R.W., Kent, T.H., and Osborne, J.W., Effects of drugs, ileal obstruction and irradiation on rat gastrointestinal propulsion, *Gastroenterology*, 59:731–739, 1970.
9. Scott, L.D. and Summers, R.W., Correlation of contractions and transit in rat small intestine, *Am. J. Physiol.*, 230:132–137, 1976.
10. Bass, P., In vivo electrical activity of the small bowel, in *Handbook of Physiology, Sect. 6: Alimentary Canal, Vol. IV: Motility*, Code, C.F., Ed., Washington, D.C: American Physiological Society, 1968, 2051–2076
11. Scultz, S.G., Frizzell, R.A., and Mellans, H.M., Ion transport by mammalian small intestine. *Ann. Rev. Physiol.*, 36:51–91, 1974.
12. Parsons, D.S., Methods for investigation of intestinal absorption, in *Handbook of Physiology, Section 6, Alimentary Canal, Vol. III, Intestinal Absorption*, Code, C.F., Ed., Washington, D.C.: American Physiological Society, 1968, pp. 1177–1216
13. Levine, R.R., Intestinal absorption, in *Absorption Phenomena*, Rabinowitz, J.L. and Myerson, R.M., Eds., 1971. New York: Wiley-Interscience, pp. 27–96.
14. Fordtran, J.S. et al., Permeability characteristics of the human small intestine, *J. Clin. Invest.*, 44:1935–944, 1965.
15. Thomas, A.B.R., and Dietschy, J.M., Intestinal absorption: major extracellular and intracellular events, in *Physiology of the Gastrointestinal Tract*, Vol. 2, New York: Raven Press, 1981, 1147–1220.
16. Levine, M.E. and Pelikan, E.W., The influence of experimental procedures and dose on the intestinal absorption of an onium compound, benzomethamine. *J. Pharmacol. Exp. Ther.*, 131:319–327, 1961.
17. Wagner, J.G., *Fundamentals of Clinical Pharmacokinetics*, Hamilton, IL: Drug Intelligence Publications, Inc., 1975, p. 173
18. Gray, G. Maldigestion and malabsorption-clinical manifestations and specific diagnosis, in *Gastrointestinal Disease: Pathophysiology, Diagnosis, Management*, Sleisenger, M.H. and Fordtran, J.S., Eds., Philadelphia: W.B. Saunders, 1978.
19. Schwabe, A.D. and Hepner, G.W., Breath tests for the detection of fat malabsorption. *Gastroenterology*, 76:216–218, 1979.
20. Watkins, J.B. et al., C-trioctanoin: a nonradioactive breath test to detect fat malabsorption, *J. Lab. Clin. Med.*, 90:422–430, 1977.
21. Pelzmann, K.S. and Havemeyer, R.N., Portal vein blood sampling in intestinal drug absorption studies, *J. Pharm. Sci.*, 60:331, 1971.
22. DeMarco, T.J. and Levine, R.R., Role of the lymphatics in the intestinal absorption and distribution of drugs, *J. Pharmacol. Exp. Ther.*, 169:142–151, 1969.
23. Soergel, K.H., Inert markers, *Gastroenterology*, 54:449–452, 1968.
24. Shay, H. et al., A simple method for the uniform production of gastric ulceration in the rat, *Gastroenterology* 5:43–61, 1945.
25. Whitely, P.E. and Dabrymple, S.A., Current Protocols in Pharmacology, 10.2.1–10.2.4, Models of Inflammation: Measure Gastrointestinal Ulceration in the Rat, 1998.
26. Megens, A.A.H.P., Awouters, F.H.L., and Niemegeers, C.J.E., General pharmacology of the four gastrointestinal motility stimulants bethanechol, metoclopramide, trimebutine, and cisapride, *Arzneim.-Forsch./Drug Res.*, 41(I)(6):631–634, 1991.
27. Hightower, N.C., Motor action of the small bowel, in *Handbook of Physiology, Section 6, Alimentary Canal, Vol. IV, Motility*, Code, C.F., Ed., Washington, D.C.: American Physiological Society, 1968, 2001–2024.

28. Tamhane, M.D. et al., Effect of oral administration of *Terminalia chebula* on gastrix emptying: an experimental study. *J. Postgrad. Med.*, 43(I):12–13, 1997.
29. Mojaverian, P. et al., Gastrointestinal transit of a solid indigestible capsule as measured by radiotelemetry and dual gamma scintigraphy, *Pharm. Res.* 6:717–722, 1989.
30. Mojaverian, P. et al., Mechanism of gastric emptying of a nondisintegrating radiotelemetry capsule in man, *Pharm. Res.* 8:97–100, 1991.
31. Itoh, T. et al., Effect of particle size and food on gastric residence time of nondisintegrating solids in beagle dogs, *J. Pharm. Pharmacol.*, 38:801–806, 1986.
32. Lui, C.Y. et al., 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, 1986.
33. Vashi, V.I. and Meyer, M.C., Effect of pH on the *in vitro* dissolution and *in vivo* absorption of controlled-release theophylline in dogs, *J. Pharm. Sci.* 77:760–764, 1988.
34. Youngberg, C.A. et al., Radio-telemetric determination of gastrointestinal pH in four healthy beagles, *Am. J. Vet. Res.*, 46:1516–1521, 1985.

chapter 9

Safety pharmacology of the immune system

Introduction

The immune system is a highly complex system of cells involved in a multitude of functions including antigen presentation and recognition, amplification, and cell proliferation with subsequent differentiation and secretion of lymphokines and antibodies. The end result is an integrated system responsible for defense against foreign pathogens and spontaneously occurring neoplasms that, if left unchecked, may result in infection and malignancy. 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, it requires regulation through cell-to-cell communications and precise recognition of self vs. nonself. Immune active agents can upset this balance if they are lethal to one or more of the cell types or alter membrane morphology and receptors. Several undesired immune system responses 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)

The safety pharmacology evaluation of potential drugs, unlike its close relative immune toxicology, is limited in scope and not part of the testing required before a new drug is administered to man. This is probably because the separate immunotoxicology guidelines cover much ground and an overlap is undesirable. As reviewed in Chapter 2, the initial and only required evaluation of a drug in clinical development is limited to an evaluation of potential to induce passive cutaneous anaphylaxis (PCA). Immune modulated responses to drugs (drug allergies) are a major problem and cause of discontinuance of use by patients who need access to the therapeutic benefits,[1] and no adequate preclinical methodology remains for identifying/predicting these responses to orally administered small molecule drugs.[2]

As a discipline, immunopharmacology involves the study of the effects that xenobiotics have on the immune system. 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[3] and drugs that are immunoenhancing at low doses may be immunotoxic at high doses. Although, in general, therapeutic agents are not endogenously produced, immunologically active biological response modifiers that naturally occur in the body should also be included because many are not known to compromise immune function when administered in pharmacologically effective doses.[4]

Although the types of immunological responses to various therapeutics may be similar, the approach taken for screening potential immunological activity will vary depending on the route of administration of the compound. 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 altering the function of 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, immune function testing of new pharmaceutical products has become a growing concern.

Until recently, immune function evaluation in pharmaceutical safety assessment has been poorly addressed by both 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. 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

Table 9.1 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 the market
Aminopyrine	Agranulocytosis	Approx. 1900	75
Phenacetin	Interstitial nephritis	Approx. 1900	83
Dipyrone	Agranulocytosis	Approx. 1930	47
Clioquinol	Subacute myelo-optic neuropathy	Approx. 1930	51
Oxyphenisatin	Chronic active hepatitis	Approx. 1955	23
Nialamide	Liver damage	1959	19
Phenoxypropazine	Liver damage	1961	5
Mebanazine	Liver damage	1963	3
Ibufenac	Hepatotoxicity	1966	2
Practolol	Oculo-mucocutaneous 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: Data from Bakke, O.M. et al., *Clin. Pharmacol. Therapy*, 35:559–567, 1984.

immunosuppression in man.[5] However, hypersensitivity (allergy) and autoimmunity are frequently observed and are serious consequences of some drug therapies.[1,6–8] 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 9.1) and accounts for approximately 15% of adverse reactions to xenobiotics.[9] In addition, adverse immune responses such as this — usually urticaria and frank rashes — are the chief unexpected findings 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 immunotoxicology testing is to help us detect these adverse effects earlier in development, before they are found in clinical trials.

PCA test for potential antigenicity of compound

PCA evaluation is a screen for a form of potential immune stimulation, and can be performed using mice, rats, or guinea pigs. Historically, primates have also been used.

Anaphylaxis is an immediate-type, hypersensitivity reaction involving specific IgE antibodies.[10] Following a sensitizing contact, IgE binds to high-affinity receptors on mast cells and basophils. After a subsequent contact, the reaction between a divalent antigen and bound IgE results in the degranulation of target cells with the immediate release of stored vaso-active mediators (histamine) and the synthesis of arachidonic acid derivatives (e.g., prostaglandins and leucotrienes). These mediators exert a wide array of biological effects, which accounts for the clinical symptoms of anaphylaxis such as urticaria, angioedema, bronchospasm, and shock.

Drug-induced anaphylaxis is a relatively rare but life-threatening event.[11] It has been reported with protein-derived drugs (i.e., heparin, insulin, etc.), penicillin, curates, and miscellaneous drugs.

The IgE-mediated PCA reaction in the rat was one of the first *in vivo* animal models in which cromolyn was shown to be effective and has since been used extensively in screening for similar compounds. In this method, inflammatory mediators released by immediate hypersensitivity reactions in the skin produce a local increase in capillary permeability. Measurement of this response provides an estimate of the intensity of the cutaneous anaphylactic reaction.[12]

Rats are generally injected intradermally with serum containing IgE antibodies to antigens such as ovalbumin. After a latent period of 24 to 72 h to permit sensitization of the cutaneous mast cells, the animal is injected intravenously with a mixture of antigen and marker dye. The animal is subsequently killed and the dorsal skin deflected to reveal the undersurface. Extravasation of the dye is then estimated by determining the diameter of the odernatous lesion or by extraction and spectrophotometric determination of the marker. Alternatively, extravasation of a high molecular weight radioactive tracer can be employed. Test drugs are usually injected intravenously together with the antigen or, where appropriate, administered orally before hand.

Both cromolyn and nedocromil are effective inhibitors in this system and exhibit a comparable potency. The two compounds exhibit tachyphylaxis and cross-tachyphylaxis[45] (see below), which is normally taken to imply that they are acting through a common mechanism.

At first sight, the rat PCA test has many apparent advantages for the screening of novel antiallergic compounds. It is rapid, simple, and large numbers of experimental drugs can be readily monitored; however, it depends on the ability of these compounds to inhibit degranulation of rat skin mast cells. Given the heterogeneity of mast cell responses to antiallergic agents, it is by no means clear to what extent these findings can be extrapolated to man. Certainly, many compounds have been developed that are much more potent than cromolyn in the rat PCA test, but that have no value in the management of clinical asthma.

The guinea pig has been considered to be an appropriate test species because of the histologic similarities that exist between the lungs of antigen-exposed animals and asthmatic patients and because the cavy can

exhibit early and late onset airway obstruction, bronchial eosinophilia, and acquired bronchial hyper-reactivity.

Bronchial challenge of ovalbumin-sensitized guinea pigs produces a triphasic reduction in specific airways conductance (sGaw), with maximal reductions at 2, 17, and 72 h, accompanied by infiltration of the airways with neutrophils at 17 h and eosinophils at 17 and 72 h. Nedocromil inhaled before challenge blocks the 2 and 17 h sGaw response but not the neutrophil influx indicating that these effects are unrelated. The sGaw response and the eosinophil accumulation at 72 h are also inhibited by nedocromil given at this time.

Repeated exposure to ovalbumin aerosol produces a significant increase in the number of epithelial eosinophils in the airways of all sizes and induces bronchial hyperreactivity as assessed by an increased pulmonary resistance to inhaled acetylcholine. These changes are also inhibited by nedocromil.

Test method[13]

1. Shave and depilate naïve guinea pigs 24 h prior to use. One guinea pig can serve as the recipient animal for two serum samples; each serum sample must be evaluated on two animals because two antigens (PA-GPSA and GPSA) are used at challenge.
2. Using an indelible ink marker, mark seven injection sites along each shaved side of the animals; the sites should be approximately 5 cm apart.
3. Dilute the test sera 1:4, 1:8, 1:16, 1:32, 1:64 and 1:128 in physiological saline; minimal volume needed is 300 µl.
4. Using the 26-gauge needle, i.d. inject 100 µl of each dilution of serum sample at the six marked sites along one side of the guinea pig; dilutions of a second serum sample are injected at the marked sites along the other side of the animal. The seventh marked site on each side of the animal is injected into two naïve guinea pigs.
5. Prepare the challenge antigen solutions by making 500- µl/ml solutions of PA-PGSA or GPSA in 1.0% Evans blue dye.
6. Weigh each animal and lightly ether anesthetize 4 h after the i.d. injections.
7. Using the 21-gauge needle, i.c. inject 2.0 ml/kg of the challenge antigen solutions intracutaneously into the appropriate animals (i.e., a 400-g guinea pig is i.c. injected with 0.8 ml of challenge solution).
8. Examine the injected skin sites for signs of bluing 15 to 30 min after the antigen challenge. The PCA titer is defined as the reciprocal of the highest dilution of serum to yield a significant blue response at the injection site. This response must be greater than the bluing observed at the saline injected site. A significant antibody response can be typified by solid blue coloration extending as a circle beyond the injection site or a continuous blue ring around the injection site (halo effect) or a discontinuous blue arc around the injection site. The blue response that occurs at the saline sites should not exceed more than 2 to 3 mm beyond the puncture site. Significant antibody to PA

is defined as a PCA titer in PA-GPSA challenged PCA animals but not in GPSA challenged PCA animals. If an endpoint titer cannot be reached at the 1:128 dilution of serum, repeat the PCA test with additional dilutions of sera that range from 1:64 to 1:2048 (or higher if required).

Several approaches are used to analyzing the presenting the data from the PCA test. One simple approach is to express the percent of PCA positive sera in each test group. A second approach is to calculate the mean PCA titer for each dose group and use a simple statistical package for determining significant differences in titer versus injected dose of chemical. A third approach is to convert the PCA dilution titers into the logarithm to the base 2 of the reciprocal of the endpoint dilution (i.e., 1:4 dilution equals a titer of 2 and 1:1024 dilution equals a titer of 10) and determine the significant differences from the log base 2 values.

The general pitfalls in the PCA test are comparable to the pitfalls encountered in the ACA (active cutaneous anaphylaxis) test. There is the risk of anesthesia and the potential loss of animals from the i.c. injection. In addition, on occasion, the PCA recipient animals can be dermatographic and develop a mottled blue appearance on the skin. These nonspecific skin responses can interfere with evaluating the PCA response at the serum injection sites. The subjectiveness in reading the PCA titer is not as great as it is in the ACA due to the differences in the appearance of the skin responses. In the ACA test, skin reactions occur as solid blue spots that decrease in size with decreasing concentrations of antigen. In the PCA test, significant antibody at the lower dilutions yields solid blue spots the size of a dime (1.8 cm), whereas antibody at the higher dilutions yields responses that are not solid blue in the center, but are fully or partially ringed. Therefore, it is easier to distinguish between a true allergic response in the skin versus irritation due to saline injections in the PCA test as compared with the ACA test.

CDER guidance for investigational new drugs

Center for Drug Evaluation and Research (CDER)'s recently (2001) promulgated draft guidance for pre-Investigational New Drug Application (INDA) immunotoxicity[14] has been open for comment and is certain to be modified some, but clearly establishes the framework for the Food and Drug Administration's (FDA) approach. 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.[15] 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 should 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 9.1) and for follow-up studies for exploring mechanisms (Figure 9.2).

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 9.2. 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 specific, or adaptive, immune system 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 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.[16]

Activation of the immune system is thought to occur when antigen-presenting cells (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 9.3). Macrophages release soluble mediators such as interleukin 1 (IL-1), which stimulate T cells to proliferate. Antigen-presenting cells must present antigen to T cells in conjunction with the class II major histocompatibility complex (MHC) proteins that are located on the surfaces 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,

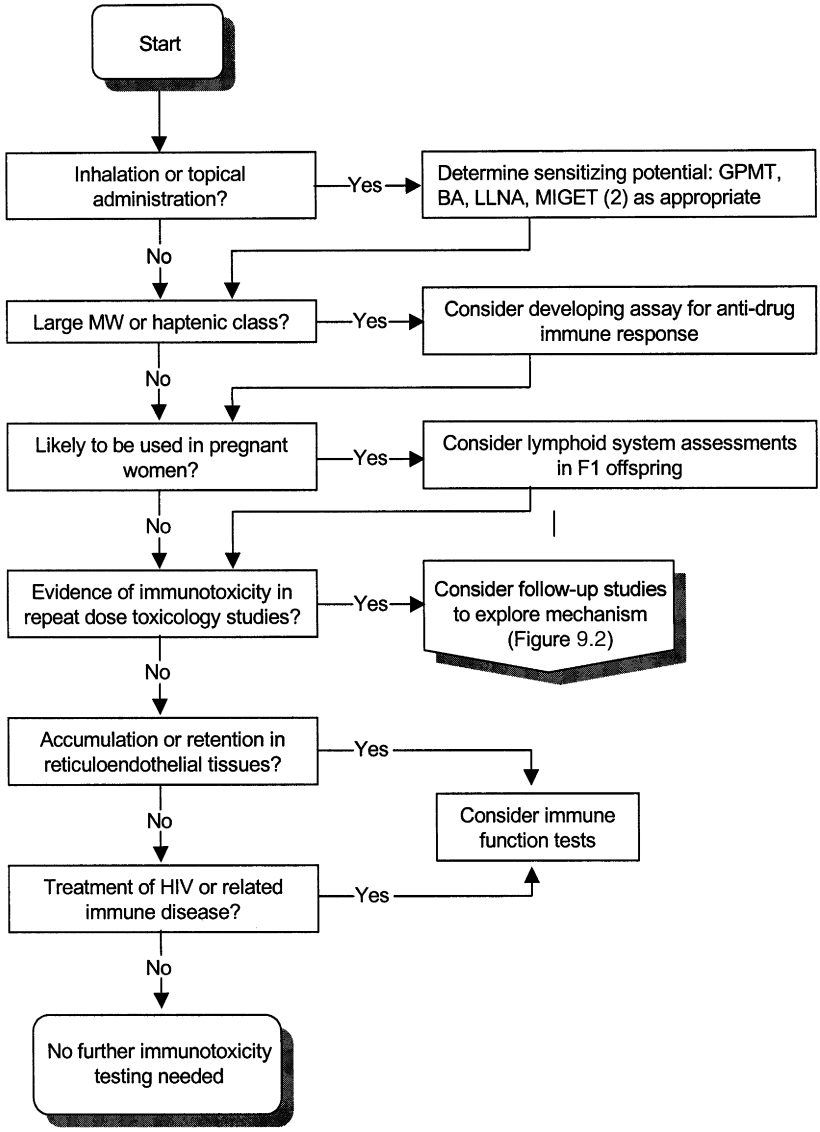
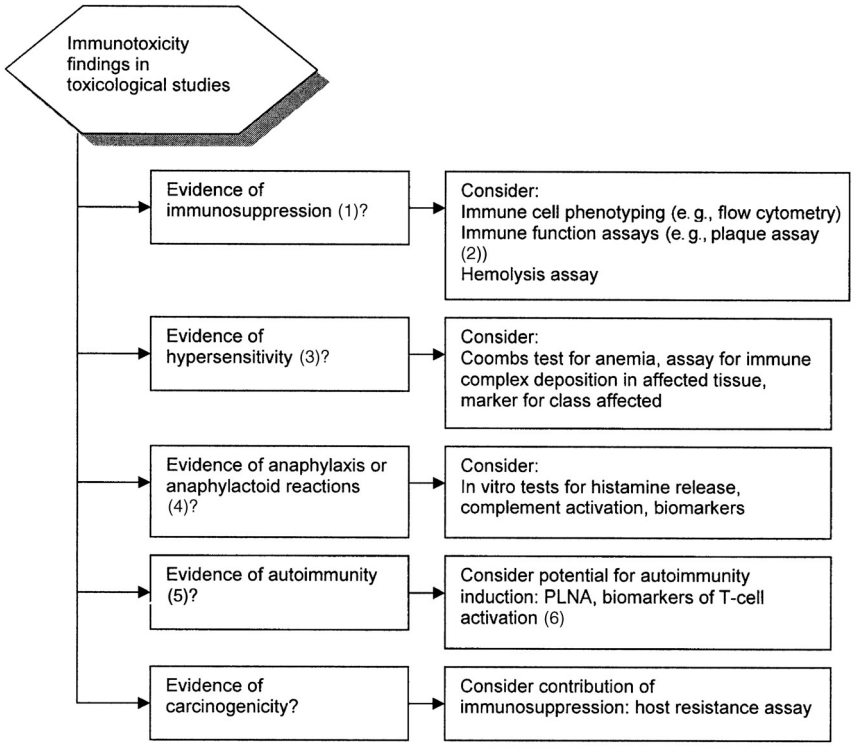


Figure 9.1 CDER flowchart for determining when to conduct specific immunotoxicity testing. GPMT—guinea pig maximization test; BA—Buehler assay (Buehler patch test); LLNA—local lymph node assay; MIGET—mouse IgE test. (There is only a relatively small database available for assessing the usefulness of the MIGET for drug regulatory purposes.)

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.



1. Examples include myelosuppression, histopathology in immune associated tissues, increased infection, tumors, decreased serum Ig, phenotypic changes in immune cells.
2. Other acceptable assays include drug effect on NK cell function in vitro bastogenesis, cytotoxic T-cell function, cytokine production, delayed-type hypersensitivity, host resistance to infections or implanted tumors.
3. Examples include anemia, luekopenia, thrombocytopenia, pnuemonitis, vasculitis, lupus-like reactions, glomerulonephritis.
4. Examples include cardiopulmonary distress, rashes, flushed skin, swelling of face or limbs.
5. Examples include vasculitis, lupus-like reactions, glomurelonephritis, hemolytic anemia.
6. There are no established assays that reliably assess potential for autoimmunity and acute systemic hypersensitivity. The popliteal lymph node assay (PLNA) has only a relatively small database available for assessing its usefulness for drug regulatory purposes.

Figure 9.2 Follow-up studies to consider for exploring mechanisms of immunotoxicity.

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 9.3. Multiple genes (polymorphisms) 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.[17]

Table 9.2 Cellular Components of the Immune System and Their Functions

Cell subpopulations	Markers[a]	Functions
Nonspecific immunity		
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)		
Specific immunity		
Humoral immunity		
Activated B cells	CD19; CD23	Proliferate; form plasma cells
Plasma cells		Secrete antibody; terminally differentiated
Resting		Secrete IgM antibodies (primary response)
Memory		Secrete IgG antibodies (secondary response)
Cell-mediated immunity		
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

[a]Activation surface markers detected by specific monoclonal antibodies; can be assayed with clow cytometry

The immune system is regulated, in part, by feedback inhibition involving complex interactions between the various growth and differentiation factors listed in Table 9.4. Because antigen initiates the signal for the immune response, elimination of antigen will decrease further stimulation.[16] 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 cells (NK cells) are

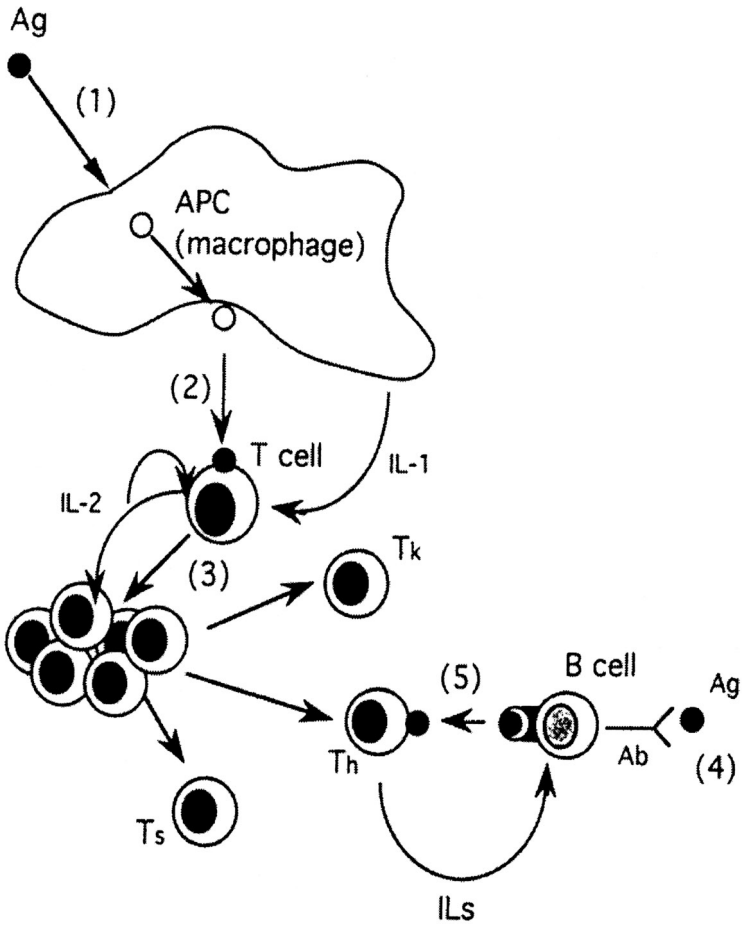


Figure 9.3 A simplified schematic of the immunoregulatory circuit that regulates the activation of T cells and B cells involved in humoral (T-cell dependent) and cell-mediated immunity. (1) Antigen (Ag) is processed by the APCs expressing class II MHC molecules. (2) Antigen plus class II MHC is then 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 interleukins (ILs) that stimulate clonal expansion and differentiation of the antigen-specific B cell.

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 various types of cells involved in cell-mediated cytotoxicity and their mechanisms of action are outlined in Table 9.5.

Table 9.3 Antibodies Involved in the Humoral Immune Response

Antibodies	Serum concentration mg ml^{-1} (%)	Characteristics/functions
IgG	10–12 (80%)	Monomeric structure (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: Data from Clark, W.R., *The Experimental Foundations of Modern Immunology,* 2nd ed., New York: John Wiley & Sons, 1983.

Table 9.4 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 and 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
	Interleukins[a]	
IL-5	T cells (T_h)	Stimulates antibody secretion (IgA), proliferation of B cells, and eosinophil differentiation

Table 9.4 (continued) Growth and Differentiation Factors of the Immune System

Factors	Cell of origin	Primary immune functions
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)	
TNF	Macrophage, B and T cells	Catectin; promotes tumor cytotoxicity; activates macrophages and neutrophils; enhances IL-2 receptor expression on T cells; inhibits antibody secretion
TNF	T cells (T_h)	Lymphotoxin; promotes T-cell-mediates cytotoxicity
	NK cells	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

[a] Includes lymphokines, monokines, and cytokines produced by T cells, macrophages, and other cells, respectively.

Source: From Golub, E.S. and Green, D.R., *Immunology: A Synthesis,* Sunderland, MA: Sinauer, 1991.

Table 9.5 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.

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. Because 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 leukemia, autoimmune disease, and chronic inflammatory diseases and to prevent transplant rejection; 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. 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.[3] Likewise, dimethylnitrosamine, a nitrosamine detected in some foods, has been shown to have both suppressing and enhancing effects on the immune system.[18]

Immunosuppression

The various cells of the immune system may differ in their sensitivities to a given xenobiotic. Thus, immunosuppression may be expressed as varying degrees of reduced activity of a single cell type of 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 has 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 humeral 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.

Because 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. T cells, macrophages, and NK cells involved in complex compensatory networks and secondary changes mediate cellular immunity. 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 (Table 9.2). Viruses are particularly susceptible to cytolysis by T cells because 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.[19] 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 also result in increased malignancy and decreased viral resistance through indirect mechanisms 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 9.3).[20] In addition to the common cell types described in Table 9.3, at least two other types of cytotoxic effector cells of T-cell origin have been identified; each has a unique lytic specificity phenotype and activity profile.[21] Of these, both LAK and TIL cells have been shown to lyse a variety of different tumor cells. However, TIL cells have 50 to 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.[18]

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.[22] 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.

Immunosuppressive drugs

Table 9.6 lists numerous types of drugs that are immunosuppressive and describes their immunotoxic effects. Several classes of drugs that characteristically depress the immune system are further discussed next.

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

Table 9.6 Immunosuppressive Drugs and Their Effects

Drugs	Biological activity and indications	Immunotoxic effects
	Hormones and antagonists	
Corticosteroids (prednisone)	Antiinflammatory; 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
	Antibiotics	
Cephalosporins	β-lactam antimicrobial	Granulocytopenia; cytopenia
Chloramphenol	Wide-spectrum antimicrobial	Pancytopenia, leukopenia
Penicillins	β-lactam antimicrobial	(idiosyncratic)
Rifampin	Macrocyclic antibiotic	Granulocytopenia; cytopenia
Tetracyclines	Antimicrobial	Suppresses T-cell function Decreased migration of granulocytes
	Chemotherapeutics and immunomodulators	
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
Carmutin and Lomustin (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

(*continued*)

Table 9.6 (continued) Immunosuppressive Drugs and Their Effects

Drugs	Biological activity and indications	Immunotoxic effects
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
Penostatin	Adenosine analog; T-cell leukemia	Inhibits adenosine deaminase; suppresses T and B cells
Zidovudine (AZT)	Antiviral (HIV)	Decreases T_h cells and granulocytes
	Miscellaneous	
Colchicine	Antimitotic; gout; antiinflammatory	Inhibits migration of granulocytes; leukopenia; agranulocytosis
Diphenylhydantoin (phenytoin)	Antiepileptic	Leukocytopenia; neutrapenia
Indomethacin (indocin)	Nonsteroidal antiinflammatory; analgesic; antipyretic	Neutrapenia
Procainamide	Antiarrhythmic	Agranulocytosis; leukopenia (rare)
Sulfasalazine	Antimicrobial antiinflammatory; ulcerative colitis/inflammatory bowel diseases	Suppresses NK cells; impaired lymphocyte function

Source: Data from Gilman, A.G. et al., *The Pharmacological Basis of Therapeutics*, 8th ed., New York: Pergamon Press, 1990.

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[23] 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.[24] 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.[25] 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. At clinical dosages, however, only the T-cell response is affected, without an apparent decrease in T-cell numbers.[26]

Pentostatin (2′-deoxycoformycin) is an adenosine analog that is a potent inhibitor of adenosine deaminase. Pentostatin is particularly useful for treating T-cell leukemia because 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.[27] 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.[28] Likewise, myelosuppression is the major toxicity associated with bolus-dose regimens of 5-FU.

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.[29] 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.[30] 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.[31] Glucocorticosteroids induce immunosuppression and antiinflammation as a result of the inhibition of specific leukocyte functions such as lymphokine activity. Glucocorti-

coids 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.

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[32] primarily by inhibiting production of IL-2,[33] and at higher doses it may inhibit expression of IL-2 receptors.[34] Cyclosporine does not prevent the stimulation of helper T cell clonal expansion by IL-2, only its activation. Because 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-host disease.[35] Cyclosporine has also been used as an antihelminthic and as an antiinflammatory agent to treat rheumatoid arthritis and other autoimmune-type diseases.

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 in 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,[36] 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.[37] Treatment with cyclophosphamide sup-

presses all classes of lymphoid cells, which may result in reduced lymphocyte function as well as lymphopenia and neutropenia.[38] Thus, it has also been administered as a large single dose prior to bone marrow transplants to suppress cellular immunity and subsequently inhibit rejection.[39]

Estrogens

β-estradiol[40,41] and therapeutics with estrogenic activity, such as diethylstilbestrol (DES), have also been shown to be immunosuppressive.[42] 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.[43] 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*.[44,45] 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.[46] Dean et al.[47] 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[48] 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.

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.[49] Clinically, mild to moderate myelosuppression may also be evident with transient leukopenia and thrombocytopenia.

Likewise, injectable gold salts, such as gold sodium thiomalate, affect a variety of immune responses in man.[50] 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 it is 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 human administered parenteral preparations.[51] 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).

Antibiotics

β-lactam-containing antibiotics such as the cephalosporins may also induce significant immunosuppressive effects[52] 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.[53] 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[53] and suppression of bone marrow stem cell activity appear to be antibody-mediated.[54]

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.[6] 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 that results in an immediate anaphylactic response to one that 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 9.7.

Hypersensitivity

The four types of hypersensitivity reactions as classified by Coombs and Gell[55] are outlined in Table 9.8. The first three types are immediate antibody-mediated reactions, whereas the fourth type is a cellular-mediated

Table 9.7 Drugs That Produce Immunostimulation

Drug	Type of response
Antibiotics	**Hypersensitivity**
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,
Isoprinosine	IL-1, and IL-2
Indomethacin	Delayed-type hypersensitivity increases T-lymphocytes
Quinidine	Rash, urticaria, asthma, granulocytopenia
Salicylates	Fever, anaphylaxis, asthma
	Rash, urticaria
	Autoimmunity
Amiodarone	Thyroiditis
Captopril	Autoimmune hemolytic anemia, pemphigus,
Chlorpromazine	granulocytopenia
Halothane	Granulocytopenia
Hydralizine	Autoimmune chronic active hepatitis
	Autoimmune hemolytic anemia, drug-induced SLE, myasthenia gravis, pemphigus, glomerulonephritis, Goodpasture's disease
Methyldopa	Autoimmune hemolytic anemia, leukopenia,
Nitrofurantoin	drug-induced SLE, pemphigus
D-Penicillamine	Peripheral neuritis
	Autoimmunity drug-induced SLE, myasthenia gravis, pemphigus, glomerulonephritis, Goodpasture's disease
Propranolol	Autoimmunity
Procainamide	Autoimmunity, drug-induced SLE, rash, vasculitis, myalgias
Pyrithioxine	Pemphigus
Antibiotics	**Hypersensitivity and Autoimmunity**
Isoniazid	Rash, dermatitis, vasculitis, arthritis, drug-induced SLE
Penicillins	Anaphylaxis, dermatitis vasculitis, serum sickness,
Sulfonamides	hemolytic anemia
	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 9.8 Types of Hypersensitivity Responses

Type and designation	Agents: Clinical manifestations	Components	Effects	Mechanism
I, Immediate (reaginic)	Food additives: GI allergies anaphylactic Penicillin: uticaria and dermatitis	Mast cells IgE	Anaphylaxis, asthma, urticaria, rhinitis, dermatitis	IgE binds to mast cells to stimulate release of humoral factors
II, Cytotoxic	Cephalosporine: hemolytic anemia Quinidine: thrombocyto-penia	IgG, IgM	Hemolytic anemia, Goodpasture's disease	IgG and IgM bind to cells (e.g., RBCs), fix complement (opsinization), then lyse cells
III, Immune complex (arthus)	Methicillin: chronic glomeruleno-phritis	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: Data from Coombs, R.R.A. and Gell, P.G.H., in *Clinical Aspects of Immunology*, P.G.H. Gell, R.R.A. Coombs, and D.J. Lachman, Eds., Oxford: Blackwell Scientific Publications, 1975, 761.

delayed-type response that may require 1 to 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 GI 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 GI emptying (ingestion), and systemic anaphylactic shock (parenteral).

Type I

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 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 9.9) 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.

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%[56] 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 con-

Table 9.9 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: Data from Clark, W.R., *The Experimental Foundations of Modern Immunology,* 2nd ed., New York: John Wiley & Sons, 1983.

jugates can then act as haptens to form the determinants for antibody induction. Although most patients who receive penicillin produce antibodies against the metabolite benzylpenicilloyl, only a fraction experience allergic reaction,[57] which suggests a genetic component to susceptibility.

Type II

Type II cytolytic reactions are mediated by IgG and IgM antibodies that can fixcomplement, opsonize particles or induce antibody-dependent cellular cytolysis 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, methyldopa), leukopenia, thrombocytopenia (quinidine), and 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.

Type III

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.[58] In addition, immunoglobulin complexes have been observed in rats treated with gold and autologus immune complex nephritis has been observed in guinea pigs.[59] 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 deposit of antigen-antibody complexes through the circulatory system results in a syndrome referred to as serum sickness, which was quite prevalent prior to 1940,[60] 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 the deposit 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 antilymphocyte serum or globulins may also exhibit serum sickness.

Type IV

Delayed-type hypersensitivity reactions (DTH) 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 to 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.[60] 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 delayed-type hypersensitivity reaction in rats.[61]

A second form of delayed-type hypersensitivity 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 to 8 h after systemic exposure and can result in fatality within 24 h following intravenous ore 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 one to two weeks to form.

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 UF 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 (i.e., 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. Because 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 because it is lipophilic and can be stored for extended periods in the body fat.[62] In addition, it is quite common for cross-reactions to occur between structurally related drugs of the same class.

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, delayed-type hypersensitivity 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.[63] 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.[64] 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 deposit 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, systemic lupus erythematosus (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.[65]

One possible mechanism for xenobiotic-induced autoimmunity involves xenobiotic binding to autologus molecules, which then appear foreign to the immuno-surveillance 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.[66] 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 nontissue-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.[67]

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.[5] 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[65]). Sulfonamides were one of the first classes of drugs identified to induce an autoimmune response, while to date, more than 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, chlorpropaminde, captopril, cefalexin, penicillin, and methyldopa.[68,69] 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. Methlydopa 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.[70] 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[65] 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[65,71] 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.[65]

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). Because gold is not observed at the site of the lesions,[72] 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 immediately after immunization with thyroglobulin in lithium-treated rats compared to controls; 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.[73]

Some drugs such as penicillin have been shown to induce autoimmunity as well as anaphylaxis.[5] 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.[74]

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 sulphamethazine. A polymorphism, which is known to exist for the genes responsible for expression of hepatic *N*-acetyl transferase 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.[75] Other predisposing factors, such as HLA phenotype (HLA-DR4 ant/or C4 allele), may also play a genetic role in determining susceptibility to hydralazine-induced SLE.[76]

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.[77,78] 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 to 5 years after implantation in a small percentage of people who 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.

References

1. Patterson, R. et al., Drug allergies and protocols for management of drug allergies, *NER Allergy Proc.* 7:325–242, 1986.
2. Hastings, K.L., Pre-clinical methods for detecting the hypersensitivity potential of pharmaceuticals: regulatory consideration, *Toxicology*, 158:85–89, 2001.
3. Hutchings, P., Nador, D., and Cooke, A., Effects of low doses of cyclophosphamide and low doses of irradiation on the regulation of induced erythrocyte autoantibodies in mice, *Immunology*, 54:97–104, 1985.
4. Koller, L.D., Immunotoxicology today, *Toxicol. Pathol.*, 1987, 15:346–351, 1987.
5. Gleichmann, E., Kimber, I., and Purchase, I.F.H., Immunotoxicology: suppressive and stimulatory effects of drugs and environmental chemical son the immune system, *Arch. Toxicol.*, 63:257–273, 1989.
6. DeSwarte, R.D., Drug allergy: an overview, *Clin. Rev. Allergy*, 4:143–169, 1986.
7. Choquet-Kastylevsky, G., Vial, T., and Descotes, J., Drug allergy diagnosis in humans: possibilities and pitfalls, *Toxicology*, 158:1–10, 2001.
8. Pieters, R., The popliteal lymphnode assay: a tool for predicting drug allergies, *Toxicology*, 158:65–69, 2001.
9. DeWeck, A.L., Immunopathological mechanisms and clinical aspects of allergic reactions to drugs. in *Handbook of Experimental Pharmacology: Allergic Reactions to Drugs*, A.L. deWeck and H. Bundgaard, Eds., New York: Springer-Verlag, 1983, 75–133.
10. Atkinson, T.P. and Kaliner, M.A., Anaphylaxis Medical Elimination, North America, 76:841–855, 1992.
11. Van der Klann, N.M., Wilson, J.H.P, and Stricker, B.H., Drug assisted anaphylaxis: 20 years reporting in the southern lands, *Elim. Europ Allergy*, 26:1355–1363, 1996.
12. Middleton, E.P. et al., *Allergy: Principles and Practice*, 4th ed., Baltimore: Shosby, 2002.

13. Sarlos, K. and Clark, E.D, Evaluating chemicals as respiratory allergens: using the Tieh Approach for Risk Assessment, in *Methods in Immunotoxicology,* Vol. 2, G.R. Burlesson, J.H. Dean, and A.E. Munson Eds., New York: John Wiley & Sons, 1995, 411–421.
14. CDER *Guidance for Industry: Immunotoxicology Evaluation of Investigational New Drugs*. U.S. Department of Health and Human Services, 2001.
15. Ader, R. and Cohen, N., Psychoneuroimmunology: conditioning and stress. *Ann. Rev. Psychol.,* 44:53–85, 1993.
16. Male, D., Champion, B., and Cooke, A., *Advanced Immunology,* Philadelphia: J.B. Lippincott, 1982
17. Roitt, I.M., Brostoff, J., and Male, D.K., *Immunology,* St. Louis: C.V. Mosby, 2001.
18. Yoshida, S., Golub, M.S., and Gershwin, M.E., Immunological aspects of toxicology: premises not promises, *Reg. Toxicol. Pharm., 9*:56–80, 1989.
19. Penn, I., Development of cancer as a complication of clinical transplantation, *Transplant Proc.,* 9:1121–1127, 1977.
20. Burnet, F.M., The concept of immunological surveillance, *Progr. Exper. Tumor. Res.,* 13:1–27, 1970.
21. Merluzzi, V.J., Comparison of murine lymphokine, activated killer cells, natural killer cells, and cytotoxic T lymphocytes, *Cell Immunol.,* 95:95–104, 1985.
22. Volkman, A., *Mononuclear Phagocyte Function,* New York: Marcel Dekker, 1984.
23. Hunter, T. et al., Azathioprine in rheumatoid arthritis. A long-term follow-up study, *Arthr. Rheum.,* I 8:15–20, 1975.
24. Hadden, J.W., Cornaglia-Ferraris, P., and Coffey, R.G. Purine, Analogs as immunomodulators, in *Progress in Immunology IV,* Y. Yamamura and T. Tada, Eds., London: Academic Press, 1984, 1393–1407.
25. Calabresi, P. and Chabner, B.A., Antineoplastic agents, in *The Pharmacological Basis of Therapeutics,* A.G. Goodman et al., Eds., New York: Pergamon Press, 1990, 1209–1263.
26. Spreafico, F. and Anaclerio, A., Immunosuppressive agents, in *Immunopharmacology 3,* J. Hadden, R. Coffey, and R. Spreafico, Eds., New York: Plenum Medical Book Company, 1977, 245–278.
27. Alper, J.C. et al., Rationally designed combination chemotherapy for the treatment of patients with recalcitrant psoriasis, *J. Am. Acad. Dermatol.*, 13:567–577, 1985.
28. Barnett, M.J. et al., High-dose cytosine arabinoside in the initial treatment of acute leukemia, *Semin. Oncol.*, 12:133–138, 1985.
29. Elion, G.B. and Hitchings, G.H., Azathioprine, in *Antineoplastic and Immunosuppressive Agents*, A.C. Sartorelli and D.G. Johns, Eds., Berlin: Springer-Verlag, 1975, 403–425.
30. Hollenberg, S.M. et al., Colocalization of DNA-binding and transcriptional activation functions in the human glucocorticoid receptor, *Cell,* 49:39–46, 1987.
31. Wallner, B.P. et al., Cloning and expression of human lipocortin, a phospholipase A2 inhibitor with potential antiinflammatory activity, *Nature,* 320:77–80, 1986.
32. Kay, J.E. and Benzie, C.R., Rapid loss of sensitivity of mitogen-induced lymphocyte activation to inhibition by cyclosporin A, *Cell Immunol.,* 87:217–224, 1984.
33. Elliot, J.F. et al., Induction of interleukin 2 messenger RNA inhibited by cyclosporin A, *Science,* 226:1439–1441, 1984.

34. Herold, K.C. et al., Immunosuppressive effects of cyclosporin A on cloned T cells, *J. Immunol.*, 136:1315–1321, 1986.
35. Kahan, B.D. and Bach, J.F., Proceedings of the Second International Congress on Cyclosporine, *Transplant. Proc.*, 20(Suppl.):11131, 1988.
36. Colvin, M., The alkylating agents, in *Pharmacologic Principles of Cancer Treatment*, B.A. Chabner, Ed., Philadelphia: W.B. Saunders, 1982, 276–308.
37. Calabresi, P. and Parks, R., Antiproliferative agents and drugs used for immunosuppression, in *Goodman & Gilman's The Pharmacological Basis of Therapeutics*, 7th ed., A.G. Gilman et al., Eds., New York: Macmillan Publishing, 1985; 1247–1306.
38. Webb, D.R. and Winklestein, A., Immunosuppression, immunopotentiation and antiinflammatory drugs, in *Basic and Clinical Immunology* 4th ed., Stites, D.P. et al., Eds. Los Altos, CA: Lange Medical, 1982, 277–292.
39. Shand, F.L., Review/Commentary: The immunopharmacology of cyclophosphamide, *Int. J. Immunopharm.*, 1:165–171, 1979.
40. Luster, M.L. et al., Estrogen immunosuppression is regulated through estrogenic responses in the thymus, *J. Immunol.*, 133:110–116, 1984.
41. Pung, O.J. et al., Influence of steroidal and nonsteroidal sex hormones on host resistance in the mouse: increased susceptibility to *Listeria monocytogenes* following exposure to estrogenic hormones, *Infect. lmmun.*, 46:301–307, 1984.
42. Luster, M.L., Pfeifer, R.W., and Tucher, A.N., Influence of sex hormones on immunoregulation with specific reference to natural and synthetic estrogens, in *Endocrine Toxicology*, J.A. McLachlin, K. Korach, and J. Thomas, Eds., New York: Raven, 67–83.
43. Luster, M.L., Blank, J.A., and Dean, J.H., Molecular and cellular basis of chemically induced immunotoxicity, *Ann. Rev. Pharmacol. Toxicol.*, 27:23–49, 1987.
44. Haukaas, S.A., Hoisater, P.A., and Kalland, T., *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, 1982.
45. Ways, S.C. et al., Alterations in immune responsiveness in women exposed to diethylstilbestrol in utero, *Fertil. Steril.*, 48:193–197, 1987.
46. Pung, O.J. et al., 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, 1985.
47. Dean, J.H. et al., The effect of adult exposure to diethylstilbestrol in the mouse: alterations in tumor susceptibility and host resistance parameters, *J. Reticuloendothel. Soc.*, 28:571–583, 1980.
48. Luster, M.L. et al., The effect of adult exposure to diethylstilbestrol in the mouse. Alterations in immunological function, *J. Reticuloendothel. Soc.*, 28:561–569, 1980.
49. Lawrence, D.A., Immunotoxicity of heavy metals, in *Immunotoxicology and Immunopharmacology*, J.H. Dean et al., Eds., New York: Raven, 1985, pp. 341–353.
50. Bloom, J.C., Thiem, P.A., and Morgan, D.G., The role of conventional pathology and toxicology in evaluating the immunotoxic potential of xenobiotics, *Toxicol. Path.*, 15:283–293, 1987.
51. Bloom, J.C. et al., Cephalosporin-induced immune cytopenia in the dog — demonstration of cell-associated antibodiesl, *Blood*, 66:1232, 1985.

52. Caspritz, G. and Hadden, J., The immunopharmacology of immunotoxicology and immunorestoration, *Toxicol. Pathol.*, 15:320–322, 1987.
53. Bloom, J.C. et al., Gold-induced immune thrombocytopenia in the dog, *Vet. Pathol.*, 22:492–499, 1985.
54. Deldar, A. et al., Residual stem cell defects associated with cephalosporin therapy in dogs, *Blood,* 66:1202, 1985.
55. Coombs, R.R.A. and Gell, P.G.H., Classification of allergic reactions responsible for clinical hypersensitivity and disease, in *Clinical Aspects of Immunology,* P.G.H. Gell, R.R.A. Coombs, and D.J. Lachman, Eds., Oxford: Blackwell Scientific Publications, 1975, 761.
56. Idsøe, O. et al., Nature and extent of penicillin side-reactions, with particular reference to fatalities from anaphylactic shock, *Bull. WHO,* 38:159–188, 1968.
57. Garratty, G. and Petz, L.D., Drug-induced immune hemolytic anemia, *Am. J. Med.*, 58:398–407, 1975.
58. Descotes, G. and Mazue, G., Immunotoxicology, *Advances in Veterinary Science and Comparative Medicine,* 31:95–119, 1987.
59. Ueda, S. et al., Autologous immune complex nephritis in gold injected guinea pigs, *Nippon Jinzo Gakkai Shi,* 22:1221–1230, 1980.
60. Clark, W.R., *The Experimental Foundations of Modern Immunology,* 2nd ed. New York: John Wiley & Sons, 1983, 1-453.
61. Exon, J.H. et al., Immunotoxicology testing: an economical multiple assay approach, *Fund. Appl. Toxicol.*, 7:387–397, 1986.
62. Unkovic, J. et al., Poster, **Annual Meeting of the American Society of Dermatology,** Washington, D.C., 1984.
63. Russel, A.S., Drug-induced autoimmune disease, *Clin. Immun. Allergy,* 1:57, 1981.
64. Dighiero, G. et al., Murine hybridomas secreting natural monoclonal antibodies reacting with self antigens,. *J. Immunol.,* 135:2267–2271, 1983.
65. Bigazzi, P.E., Autoimmunity induced by chemicals, *Clin. Toxicol.,* 26:125–126, 1988.
66. Weigle, W.O., Analysis of autoimmunity through experimental models of thyroiditis and allergic encephalomyelitis, *Adv. Immunol.,* 30:159–275, 1980.
67. Talal, N., Autoimmune mechanisms in patients and animal models, *Toxicol. Pathol.*, 15:272–275, 1987.
68. Logue, G.L., Boyd, A.E., and Rosse, W.F., Chlorpropamide-induced immune hemolytic anemia, *New Engl. J. Med.*, 283:900–904, 1970.
69. Kleinman, S. et al., Positive direct antiglobulin tests and immune hemolytic anemia in patients receiving procainamide, *New Engl. J. Med.,* 311:809–812, 1984.
70. Mach, P.S., Brouilhet, H., and Smor, B., D-penicillamine: a modulator of anti-DNA antibody production, *Clin. Exp. lmmunol.,* 63:41418m, 1986.
71. Ntoso, K.A. et al., 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, 1986.
72. Druet, P. et al., Immunologically mediated glomerulonephritis induced by heavy metals, *Arch. Toxicol.,* 50:187–194, 1982.
73. Hassman, R.A. et al., The influence of lithium chloride on experimental autoimmune thyroid disease, *Clin. Exp. Immunol.,* 61:49–57, 1985.

74. Batchelor, F.R., Dewdney, J.M., and Cazzard, D., Penicillin allergy: the formation of penicilloyl determinant, *Nature* **(London)**, 206:362–364, 1965.
75. Perry, H. M., Tane, M., and Camody, S., Relationship of acetyl transferase activity to antinuclear antibodies and toxic symptoms in hypertensive patients treated with hydralazine, *J. Lab. Clin. Med.*, 76:114–125, 1970.
76. Spears, C.J. and Batchelor, J.R., Drug-induced autoimmune disease, *Adv. Nephrol.*, 16:219–230, 1987.
77. Kumagai, Y. et al., Clinical spectrum of connective tissue disease after cosmetic surgery. Observations of eighteen patients and a review of the Japanese literature, *Arthr. Rheum.*, 27:1–12, 1984.
78. Guillaume, J.C., Roujeau, J.C., and Touraine, R., Lupus systémique après protheses mammaires, *Ann. Derm. Verner.*, 111:703–704, 1984.
79. Bakke, O.M., Wardell, W.M., and Lasagna, L., Drug discontinuations in the United Kingdom and United States, 1964–1983: issues of safety, *Clin. Pharmacol. Therapy*, 35:559–567, 1984.
80. Golub, E.S. and Green, D.R., *Immunology: A Synthesis*, Sunderland, MA: Sinauer, 1991.
81. Gilman, A.G. et al., *The Pharmacological Basis of Therapeutics*, 8th ed. New York: Pergamon Press, 1990.

chapter 10

Integration of evaluations of safety pharmacology endpoints into existing study designs

As initially discussed in the first chapter of this volume, there are two separate approaches to evaluating the safety pharmacology of a potential new pharmaceutical. One is to design and conduct a series of free-standing, specialized studies each to address one of the International Conference on Harmonization (ICH) guideline requirements. The second is to integrate the required measurements into already existing safety evaluation designs to the maximum possible extent, performing extra, specialized studies and using additional animals only as absolutely necessary.

This dichotomy has actually existed for at least 20 years, first coming to the surface when the issue of neurotoxicity was raised in the late 1970s and again when immunotoxins became a concern. The relative advantages and disadvantages of each approach were discussed earlier and will not be repeated here. It should be stated that the separate study approach is favored by European regulatory authorities, and this is reflected in the wording of specific ICH guidances, though implicitly rather than explicitly. This desire persists even in the face of considerable clinical data, currently leading to such studies being done to support registration rather than clinical development. Rather than revisit the advantages and disadvantages, this chapter presents practical approaches to the integration of the core safety pharmacology measurements into commonly employed pre-investigational new drug (IND) safety assessment designs.

The core battery requirements for a new drug under ICH are stated as:

1. Cardiovascular:
   a. Hemodynamics (blood pressure, heart rate)
   b. Autonomic function (cardiovascular challenge)

   c. Electrophysiology (electrocardiogram [EKG] in dog)
2. Respiratory:
   a. Respiratory functions — measurement of rate and relative tidal volume in conscious animals
   b. Pulmonary Function — measurement of rate, tidal volume, and lung resistance and compliance in anaesthetized animals
3. Central nervous system (CNS):
   a. Irwin test — general assessment of effects on gross behavior and physiological state
   b. Locomotor activity — specific test for sedative, excitatory effect of compounds
   c. Neuromuscular function — assessment of grip strength
   d. Rotarod — test of motor coordination
   e. Anesthetic interaction — test for central interaction with barbiturates
   f. Anti/pro-convulsant activity — potentiation or inhibition of effects of pentylenetetrazole.
   g. Tail flick — tests for modulation of nociception (i.e., perception of pain, such as hot plate, tail pinch)
   h. Body temperature — measurement of effects on thermoregulation
   i. Autonomic function —- interaction with autonomic neurotransmitters *in vitro* or *in vivo*
   j. Drug dependency — test for physical dependence, tolerance, and substitution potential
   k. Learning and memory — measurement of learning ability and cognitive function in rats

These evaluations are required before a first-in-man (FIM) dosing — that is, effectively, to open an IND.

Correspondingly, there is also a generally required set of in vivo toxicology evaluations required to open an IND. These are acute and repeat dose studies in at least two species (one rodent and one nonrodent). Most commonly, this means at least four studies:

1. Acute dose — rat
2. Acute dose — nonrodent (dog or monkey)
3. 14/28 day — rat
4. 14/28 day — dog

We should start by considering the basic current design and conduct of such studies as currently performed in the pharmaceutical industry.[1,2]

Rodent acute toxicity

Many investigations into the sources of variability in acute toxicity testing have been conducted and these have been reviewed by Elsberry.[3] 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 and interlaboratory variation was found to be minimal. Thus, it is equally important that the details of the protocol be well described and followed. It is not appropriate to draw dosage-response conclusions 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 have been proposed by the interagency regulatory liaison group.[3,4] These do not differ markedly from those mandated by the Toxic Substance Control Act of 1986.

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 four to six dose levels 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 values of precise lethality data.[1] Retrospective analyses by DePass[5] and Olson et al.[6] have demonstrated that decreasing group size to two or three animals generally has little impact on overall study results. Thus, the number and size of dosage groups will depend, to an extent, on the methods of statistical analysis. Typically, five rodents per sex group are used to provide sufficient power to evaluate endpoints other than lethality.

In setting dosages for acute studies, a few common sense rules must be applied. First, the intrinsic biological and chemical activity of the drug must be considered. Zbinden and Flury-Roversi[7] have documented several cases where lethality was of no biological relevance. The oral lethality of tartic acid, for example, is due to the caustic action of a high concentration of acid in the gastrointestinal (GI) tract. A proper formulation for all preclinical testing is an important but often overlooked necessity. Additionally, it is uncommon that a completely unknown chemical will be tested. Factors such as known pharmacological profile, chemical or physical characteristics including molecular weight, particient coefficient, and the like, and the toxicity of related structures will commonly be known and offer some guidance.

A true complete toxicity screen will look at a broad range of endpoints, including clinical signs, body weight, food consumption, organ weights, and gross necropsy. It is also very common to perform an functional observation battery (FOB) as part of such studies in the rat.

Nonrodent acute toxicology

Because of animal costs and considerations of animal welfare, it is uncommon that acute testing in nonrodents will be conducted in an analogous manner as in rodent testing. Rather, some form of pyramiding or escalating dose study is more common.

Using this type of design, one can obtain information about lethality and limits of effects 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), to allow for clearance of drug from the bodies of test animals. The dosage at each session may be 1, 3, 10, 30, 100, 300, 1000, and 3000 mg/kg or 10, 20, 40, 80, 160, 320, 640, and 1280 mg/kg. One is literally stepping up, or pyramiding, the lethality-dosage curve. Dosing continues in this fashion until one or both animals die or until some practical upward limit is reached. For drugs, there is no longer a need to go higher than 1000 mg/kg for rodents and nonrodents. An alternative, but similar, design is the leapfrog study. This consists of two groups of two animals each. They are dosed on alternating days, but the dosages are increased each day. Extending the example of the pyramiding regimen, group one would receive 10, 60, and 120 mg/kg, while group two would be given 30, 100, and 120 mg/kg. This design is of 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 lethality in nonrodent species. An exploratory study typically uses an animal of each sex.

Three conclusions can be reached on the basis of data from a pyramiding dosage study. First, if none of the animals dies, then both the threshold or minimum lethal dosage (MLD) and LD_{50} are greater than the top or limit dosage. Second, if all animals die at the same dosage, then both the MLD and the LD_{50} are reported as being between the last two dosages given. This not uncommon finding is an indication that the lethality curve has a steep slope. Third, one animal may die or show an effect at one dosage and remaining equivalent effects 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 LD_{50} is reported as being between this later dosage and the dosage at which the last animal dies. A frequently employed variation with nonrodents, if severe toxicity is not observed, is to dose the animals five or seven consecutive days at the highest observed tolerated dose. The phase "B" or cumulative toxicity study portion serves to provide more confidence in selecting the top dose in subsequent repeat dose studies.

There are disadvantages to the pyramiding dose protocol. First, it cannot produce a lethality curve or provide for the calculation of an LD_{50}. Second, this method cannot identify delayed deaths. If an animal, for example, dies one hour after the second dosage, one has no any 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 animal for 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 by 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, as there is generally no need for a 1- to 2-week, postdosing observation or holding

period, the actual study may not take significantly more time than a test of more traditional design.

Detailed clinical signs, body weights, food consumption, gross necropsy observation organ weights, and toxicokinetic data are commonly recorded in all such studies. Evaluation of the meaning of such measure is not just 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. Thus, 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.

Drugs are seldom 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 given to nonrodent species by the intraperitoneal route. Test articles are normally given orally by capsule and by gavage to dogs and pigs, making the process 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 results. Another problem, which is most peculiar to the dog, is emesis. Oral administration of 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 repeated-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 when appropriate and necessary by 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 lesion 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.

Repeated dose (IND pivotal toxicity) studies

Repeat dose, or subchronic— shorter than chronic, studies areshotgun in nature; that is, they are designed to look at a very broad range of endpoints with the intention of screening as broadly as indications of toxicity. Meaningful findings are rarely limited to a single endpoint; 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 effects 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. Seeking to detect adverse functional (pharmacologic) effects are but a further extension of this approach.

Parameters of interest in the repeat-dose study can be considered as sets of measure, each with its own history, rationale, and requirements. 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.

Body weight

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 includes having each group vary insignificantly from one another in terms of body weight. Weights are measured prior to the initial dose, then typically 1, 3, 5, 7, 11, and 14 days thereafter. 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 is a very sensitive albeit nonspecific indicator of systemic toxicity.

Food 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 to accurately maintain dose levels. Additionally, the actual parameter 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 monthly basis. Water consumption, which is also sometimes measured, is similar in interpretation and use. Giving measured quantities of food as opposed to the traditional *ad libitum* feeding is recommended for studies of more than 90 days, as it precludes excessive body weight gain and the complications thereof.[8]

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 to be 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.[9]

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,[1] and evaluation 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 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 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 to evaluate adverse effects on the target organ system of concern.

In larger animals — dogs and primates — some of these measures such as electrocardiograms (EKGs) are made as a matter of course in all studies.

Clinical 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 pathology determinations as opposed to anatomical. The parameters measured cover clinical chemistry, hematology, and urinalysis using samples of blood and urine collected at predetermined intervals during the study. Conventionally, these intervals are evenly spaced at three points over the course of the study, with the first being 1 month after study initiation and the last being immediately prior to the 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. 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. The samples on which analysis is performed are from fixed points in time, which may miss transient changes — typically, increases — in some enzyme levels.

Pharmacokinetics and metabolism

Pharmaceutical subchronic toxicity studies are always accompanied by a parallel determination of the pharmacokinetics of the material of interest and 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 endpoints. 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.[10,11] 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.

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

Particularly in light of recent concerns with drug-induced arrhythmias, careful consideration must be given to incorporating adequate evaluation of drug induced alterations on cardiovascular function. This is usually achieved by measuring blood pressure, heart rate, and by an EKG 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. The Q to T interval should specifically be evaluated.

The FDA has current draft criteria[12] for endpoints to be incorporated in studies as a screen for neurotoxicity. In practice, a functional observation battery is employed at several endpoints — usually 1 and 3 months in to the study — to fill these requirements.

In response to concerns about potential effects of drugs on the immune system, the FDA[12] has proposed that a basic set of criteria be evaluated and considered in standard subchronic and chronic studies. Most of these endpoints are, it should be noted, already collected in traditional subchronic designs.

All subchronic and chronic toxicity studies now incorporate evaluation of the basic pharmacokinetics of a compound either in the study itself of in a parallel study.

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 during termination of all surviving animals from test and control groups. Tissues from all animals should 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.

The traditional design for a repeat-dose toxicity study is very straightforward. The appropriate number 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 purpose 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 staggerd 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 the same size as the experimental groups, to guard against 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 biologically insignificant. This is, however, an indefensible approach. Historical controls have value, but it is the concurrent control group in a study that is of concern.

Interval or satellite groups 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 the subject animals relative to other 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 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 to 80 dogs involved in a study, much less the 160 to 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 being 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 three% 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 to 5 days, then one sex should be initiated on dosing one 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 study is to stagger the start-up either 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 such as pharmacokinetics and special measures, but not others such as recovery and interval sacrifice.

Recording the surface electrocardiogram in dogs has been the basic gold standard for determining effects of pharmaceutical compounds on cardiac electrophysiology. In most cases, electrocardiograms are obtained as a component of the nonclinical evaluation studies, including safety pharmacology and toxicology assessments. The majority of evidence confirms the dog's role as the most predictive preclinical species with regard to human electrocardiography.[13] Detweiler provides a compelling discussion concerning the use of different species in electrocardiographic assessments. As previously mentioned, the rat should be considered anomalous with respect to the human EKG due to the overwhelming role of I_{to} in the repolarization process, although the use of nonhuman primates is increasing due to a variety of development factors — limited compound requirements, canine toxicity, biotechnology products. Detweiler describes the nonhuman primate, (namely the cynomolgous monkey, as the most misleading species. The reasons are varied, but primarily focus on the inherent characteristics of the species in addition to the technical challenges presented with the model. Clearly, cost and ethical concerns also play a key role in the decision process. Most investigators continue to rely on the dog, although a few also examine effects on EKG in other species. Recognizing the expert opinion on the utility of the dog as a preclinical predictor of human electrocardiography, the value of these additional examinations remains dubious for most traditional NCEs.

For standard toxicology studies, electrocardiograms are usually recorded in either a single Lead II or multilead configuration. Animals should be conscious, although some investigators also rely on data from anesthetized subjects. If anesthetized animals are used, consideration should be given to the ancillary effects of the anesthetic on cardiovascular parameters, especially the QT interval. Most anesthetic agents affect hemodynamic and/or electrophysiology to the extent that their use should be restricted to special circumstances.[14]

The dog that has been instrumented with telemetry technology is currently considered by many as the gold standard model for evaluation of a compound's potential for QT interval prolongation. Conscious, restrained animals of all species have artificially raised heart rates and sympathetic drive that can easily confound interpretation of electrocardiographic effect. In the beagle, heart rates of approximately 50 to 100 beats per minute (bpm) have been recorded in the unrestratined telemeterized subject, while those manually restrained present with rates of 90 to 130 bpm. Numerous correction formulae have been developed to adjust for the inverse relationship between heart rate and QT interval, although the shortcomings of these

mathematical interpolations are many. Many methods of correction, and indeed noncorrection, have been proposed.[15–18] Although the use of conscious, sling-trained dogs can reduce the variability observed in the restrained animal, it is still considered preferable to utilize telemetry implants that derive cardiovascular parameters from the conscious, free-roaming animal. In certain cases, Holter technology can be used with the dog in lieu of implanted telemetry. Utilizing this route of examination, animals are not necessarily subjected to surgery and recovery periods and their potential sequella as is the case with most telemetry collection systems. With either route, the experimental subject is allowed to remain undisturbed in its home cage while the measurements are collected and stored by data acquisition systems. In theory, this allows for the maximum amount of relevant information to be captured with a minimum of potentially confounding interactions.

Indeed, a recent survey on the correlation between nonclinical safety assessment paradigms and the toxicity of human cardiovascular toxicities was among the most highly correlated animal species studies.[19] Roughly 85% of human cardiovascular toxicities were predicted by studying the dog, a testimony to its historical usage and physiological similarity.

If electrocardiographic examinations result in changes with QT interval or other EKG intervals, it is possible to perform detailed *in vivo* electrophysical studies. These experimental studies include monophasic action potential duration measurements via the use of special electrodes and proarrhythmia models. These time-consuming and technically challenging experiments are usually best performed by individuals possessing extensive experience with *in vivo* preparations and the subtleties associated with these models. Examination of the monophasic action potential *in vivo* is most often performed using the anesthetized dog. Briefly, MAP electrodes can be placed on the epicardial and endocardial surface of the myocardium and action potentials can be recorded prior to and following test article administration. Thus, apparent changes in the QT interval can be examined at the source of this parameter, the ventricular monophasic action potential.

An anesthetized canine model utilizing animals with chronic idioventricular bradycardia after complete artrial ventricular (AV) block is also available to examine proarrhythmic potential.[20] Programmed electrical stimulation can induce early afterdepolarization-altering agents with this preparation. Major caveats with proarrhythmia assays include their technical complexity[21,22] and minimal evidence suggesting predictivity for humans, although the latter characteristic is primarily due to the very low frequency of performance.

Integration

It should be clear at this point that evaluations equivalent to all but the core safety pharmacology battery requirements, except for a few of the specialized

pharmacologic challenge requirements, are already performed in standard pre-IND systemic toxicity studies. These match up as follows:

Cardiovascular:
- Hemodynamics (blood pressure, heart rate) **B**
- Autonomic function (cardiovascular challenge) ***a**
- Electrophysiology (EKG in dog) **B**

Respiratory:
- 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 anaesthetized animals ***b**

CNS:
- Irwin test — general assessment of effects on gross behavior and physiological state. **A**
- Locomotor activity — specific test for sedative, excitatory effect of compounds
- Neuromuscular function — assessment of grip strength **A**
- Rotarod — test of motor coordination **A**
- Anesthetic interaction — test for central interaction with barbiturates *c
- Anti/pro-convulsant activity — potentiation or inhibition of effects of pentylenetetrazole ***d**
- Tailflick — tests for modulation of nociception (i.e., pain, such as tail pinch) **A**
- Body temperature — measurement of effects on thermoregulation **A**
- Autonomic function — interaction with autonomic neurotransmitters *in vitro* or *in vivo* ***e**
- Drug dependency — test for physical dependence, tolerance, and substitution potential ***f**
- Learning and memory — measurement of learning ability andcognitive function in rats ***g**

Note: A = functional observation battery (FOB) rat acute or 28-day study; B = cardiovascular measurements in dog 28-day study, and * is not currently incorporated.

Those requirements currently not addressed (denoted by an asterisk *) can largely be addressed as follows:

a. Function could be assessed within the existing dog cardiovascular assessment framework.
b. Pulmonary function assessment would require either a separate test or a significant change to one of the current rat study designs.
c, d, e, and f. These are pharmacologic challenge tests and need to be independently conducted.
g. These need to be satisfied by a separate maze study.

Otherwise, simple careful consideration of dose selection and timing of measurements should serve to make the addressal of desired endpoints more than adequate to address any concerns. Such data has the advantage of being accompanied by all the other information collected at the same time including toxicokinetics thus to allow much improved evaluation as to meaning and relevance while also limiting the increase in the number of animals required to evaluate the safety of a potential new drug.

References

1. Gad, S.C. and Chengelis, C.P., *Acute Toxicology Testing*, 2nd ed., San Diego: Academic Press, 1998.
2. Gad, S.C., *Drug Safety Evaluation*, New York: Wiley-Interscience, 2002.
3. Elsberry, D., Screening approaches for acute and subacute toxicity studies, in *Safety Evaluations of Drugs and Chemicals*, W. Lloyd, Ed., Washington, D.C.: Hemisphere Publishing, 1986, 145–151.
4. Interagency Regulatory Liason Group (IRLG), *Testing Standards and Guidelines Work Group (HFE-88)*, Washington, D.C.: IRLG1981.
5. Depass, L., Alternative approaches in median lethality (LD_{50}) and acute toxicity testing, *Tox. Lett.*, 49:159–170, 1989.
6. Olson, H. et al., Reduction in animals used for acute toxicity testing based on retrospective analysis, *Toxicologist*, 141(abstract), 1990.
7. Zbinden, G. and Flury-Roversi, M., Significance of the LD50 Test for the toxicological evaluation of chemical substances, *Arch. Toxicol.*, 47:77–99, 1981.
8. Allaben, W.J. and Hart, R.W., Nutrition and toxicity modulation: The impact of animal body weight on study outcome, *Int. J. Toxicol.*, 17(Suppl.) 2:1–3, 1998.
9. Arnold, D.L., Grice, H.C., and Krawski, D.R., *Handbook of In Vivo Toxicity Testing*, San Diego: Academic Press, 1990.
10. Yacobi, A., Skelly, J.P., and Batra, V.K., *Toxicokinetics and New Drug Development*, New York: Pergamon Press, 1989.
11. Tse, F.L.S. and Jaffe, J.M., *Preclinical Drug Disposition,* New York: Marcel Dekker, 1991.
12. FDA, Toxicological principals for the safety of food ingredients, *Redbook 2000*, Washington, D.C.: Center for Food Safety and Applied Nutrition, FDA, 2000.
13. Detweiler, D.K., Electrocardiography in toxicological studies, in *Comprehensive Toxicology*, Spies, I.G., McQueen, C.A., and Gandolfi, A.J., Eds., New York: Pergamon Press, 1997, 95–114.
14. Gross, D.A., Cardiovascular effects of inhalant anesthetic agents, in *Animal Models in Cardiovascular Research*, Dordrecht, Netherlands: Kluwer Academic Publishers, 1994, 253–301.
15. Batchvarov, V., and Malik, M., Individual patterns of QT/RR relationship, *Card. Electrophysiol. Rev.*, 6(3):282–288, 2002.
16. Davey, P., A new physiological method for heart rate correction of the QT interval, *Heart,* 82(2):183–186, 1999.
17. Spence, S. et al., The heart rate corrected QT interval of conscious beagle dogs: a formula based on analysis of covariance, *Toxicol. Sci.*, 45(2):247–258, 1998.
18. Van de Water, A. et al., An improved method to correct the QT interval of the electrocardiogram, *J. Pharmacol. Meth.*, 22(3):207–217, 1989.

19. Olson, H. et al., Concordance of the toxicity of pharmaceuticals in humans and in animals, *Regul. Toxicol. Pharmacol.*, 32:56–67, 2000.
20. Vos, M.A. et al., Reproducible induction of early afterdepolarizations and torsade de pointes arrhythmias by D-sotalol and pacing dogs with chronic atrioventricular block, *Circulation*, 3:864-872, 1990–1991.
21. Carlsson, L., Almgren, O., and Duker, G., QTU-prolongation and torsades de pointes induced by putative class III antiarrhythmic agent in the rabbit: etiology and interventions, *J. Cardiovasc. Pharmacol.*, 16(2): 276–285, 1990.
22. Weissenburger, J., Davy, J.M., and Chezalviel, F., Experimental models of torsades de pointes, *Fundam. Clin. Pharmacol.*, 1993, 7(1):29–38.

Selected regulatory and toxicological acronyms

510(k) Premarket notification for change in a device
AALAS American Association of Laboratory Animal Science
AAMI Association for the Advancement of Medical Instrumentation
ABT American Board of Toxicology
ACGIH American Conference of Governmental Industrial Hygienists
ACT American College of Toxicology
ADE Adverse Drug Event (of drug substances)
ADI Allowable Daily Intake
AIDS Acquired Immune Deficiency Syndrome
AIMD Active Implantable Medical Devices
ANSI American National Standards Institute
APHIS Animal and Plant Health Inspection Service
ASTM American Society for Testing and Materials
CAS Chemical Abstract Service
CBER Center for Biologic Evaluation and Research (FDA)
CDER Center for Drug Evaluation and Research (FDA)
CDRH Center for Devices and Radiological Health (FDA)
CFAN Center for Food and Nutrition (FDA)
CFR *Code of Federal Regulations*
CIIT Chemical Industries Institute of Toxicology
CPMP Committee on Proprietary Medicinal Products (U.K.)
CSA Control Standard Endotoxin
CSM Committee on Safety of Medicines (U.K.)
CTC Clinical Trial Certificate (U.K.)
CTX Clinical Trial Certificate Exemption (U.K.)
CVM Center for Veterinary Medicine (FDA)
DART Development and Reproduction Toxicology
DHHS Department of Health and Human Services
DIA Drug Information Associates
DMF Drug (or Device) Master File

DSHEA Dietary Supplement Health and Education Act
EEC European Economic Community
EFPIA European Federation of Pharmaceutical Industries Association
EM Electron Microscopy
EPA Environmental Protection Agency
EU European Union
FCA Freund's Complete Adjuvant
FDA Food and Drug Administration
FDCA Food, Drug and Cosmetic Act
FDLI Food and Drug Law Institute
FIFRA Federal Insecticides, Fungicides and Rodenticides Act
FIM First in Man
GCP Good Clinical Practices
GLP Good Laboratory Practices
GMP Good Manufacturing Practices
GPMT Guinea Pig Maximization Test
HEW Department of Health, Education and Welfare (no longer exists)
HIMA Health Industry Manufacturer's Association
HSDB Hazardous Substances Data Bank
IARC International Agency for Research on Cancer
ICH International Conference on Harmonization
ID Intradermal
IDE Investigational Device Exemption
IND(A) Investigational New Drug Application
INN International Nonproprietary Names
IP Intraperitoneal
IRAG Interagency Regulatory Alternatives Group
IRB Institutional Review Board
IRLG Interagency Regulatory Liaison Group
ISO International Standards Organization
IUD Intrauterine Device
IV Intravenous
JECFA Joint Expert Committee for Food Additives
JMAFF Japanese Ministry of Agriculture, Forestry, and Fishery
JPMA Japanese Pharmaceutical Manufacturers Association
LA Licensing Authority (U.K.)
LAL *Limulus* amebocyte lysate
LD_{50} Lethal Dose 50: The dose calculated to kill 50% of a subject population, median lethal dose
LOEL Lowest observed effect level
MAA Marketing Authorization Application (EEC)
MCA Medicines Control Agency
MD Medical device
MedDRA Medical Dictionary for Regulatory Activities
MHW Ministry of Health and Welfare (Japan)
MID Maximum implantable dose

MOW Margin of Exposure
MOU Memorandum of Understanding
MRL Maximum Residue Limits
MSDS Material Safety Data Sheet
MTD maximum tolerated dose
NAS National Academy of Science
NCTR National Center for Toxicological Research
NDA new drug application
NIH National Institutes of Health
NIOSH National Institute Occupational Safety and Health
NK natural killer
NLM National Library of Medicine
NOEL no-observable-effect level
NTP National Toxicology Program
ODE Office of Device Evaluation
OECD Organization for Economic Cooperation and Development
PDI Primary Dermal Irritancy
PDN Product Development Notification
PEL Permissible Exposure Limit
PhRMA Pharmaceutical Research and Manufacturers Association
PL Produce License (U.K.)
PLA Produce License Application
PMA Premarket Approval Applications
po per os (orally)
PTC Points to Consider
QAU Quality Assurance Unit
RAC Recombinant DNA Advisory Committee
RCRA Resources Conservation and Recovery Act
sc subcutaneous
SCE sister chromatic exchange
SNUR Significant New Use Regulations
SOP Standard Operating Procedure
SOT Society of Toxicology
SRM Standard Reference Materials (Japan)
STEL Short-Term Exposure Limit
TLV threshold limit value
USAN United States Adopted Name Council
USDA United States Department of Agriculture
USEPA United States Environmental Protection Agency
USP United States Pharmacopoeia
VAERS Vaccine Adverse Event Reporting System
VSD Vaccine Safety Data Link
WHO World Health Organization

Safety pharmacology labs

Lab	Location	Phone no.	Web site	Additional services	Rat	Rabbits	Dog	Carcinogenicity	DART	Inhalation	Primate	Pig	IV Infusion	Genotoxic	Devices	Metabolism	Analytical	Special Studies
Austrian Research Center	Seibersdorf, Austria	+43 (0) 50550-0	www.arcs.ac.at/	Yes	X	X			X									SP
Calvert	Olyphant, PA	(570) 586-2411	www.calvertpreclin.com	Yes	X		X	X	X		X	X	X		X	X	X	SP
Central Toxicology Labs. — Sygenta	Alderley Park, Macclesfield SK10, 4TJ, UK	+44 16255 15852 (p) +44 16255 17314 (f)		Yes	X	X	X	X	X	X	X			X		X	X	SP
Charles River			www.criver.com	Yes														
	Worchester, MA;	(978) 658-6000				X	X			X		X	X			X	X	SP
	Redfield, AR;	(877) 274-8371			X	X	X	X			X		X					
	Sparks, NV:								X		X							
	Horsham, PA:								X									
	Spencerville, OH	(419) 647-4196			X	X	X	X	X			X						
CIT – Centre Internat. De Toxicologie	B.P. 563, Evreux, France	+33 2 32 292626 (p) +33 2 32 678705 (f)	www.citox.com	Yes	X	X	X	X	X		X	X	X	X		X	X	SP
Covance	Vienna, VA Madison, WI Harrogete, U.K. Muenster, FRG	(888) COVANCE (888) 541-LABS	www.covance.com	Yes, clinical, consulting +	X	X	X	X	X		X	X	X	X	X	X	X	SP
ILS	RTP, NC	(919) 544-5857	www.ils-limited.co.uk	Yes, data analysis, ecotoxicology, +	X			X	X					X		X	X	SP
Inveresk	Tranent, Scotland Montreal, QU	44 (0) 1875 614555 (514) 630-8200	www.inveresk.com	clinical	X	X	X	X	X	X	X	X	X	X		X	X	SP
MDS Pharma Services	L'Arbresle, France Geneva, Switzerland	Details to be delivered	www.mdsps.com	Yes, discovery, clinical, +	X	X	X	X	X				X	X		X	X	SP
MPI	Mattawan, MI	(269) 668-3336	www.mpi.com	Yes	X	X	X	X	X	X	X	X	X			X	X	SP
Quintiles	Ledbury, UK & US	44 (0) 131 451 2560 (877) 988-2100	www.quintiles.com	Yes, clinical, development, +	X	X	X	X	X			X	X	X		X	X	SP
Rallis Research Center	Bangalore, India		www.rallis.co.in		X			X										
RCC Ltd.	Baltimore, MD Zelgliwe.g., 1 CH-4452 Itegen Switzerland	(410) 385-1666 +41 61975 1111 (p)	www.rcc.ch	Yes, consulting, regulatory +	X	X	X	X	X	X		X		X	X	X	X	SP
SafePharm	Derby, UK	00 44 (0) 1332 792896	www.safepharm.com	Yes, fish, +	X			X	X	X				X	X	X	X	SP
SGS	Fairfield, NJ	(800) 777-8378	www.ustesting.sgsna.com	Yes, combustion, fish, +		X								X	X		X	SP
Southern RI	Birmingham, AL	(888) 322-1166	www.southernreaserch.com	Yes, formulation, development, +	X	X	X	X	X		X	X	X	X	X	X	X	SP
TherImmune	Gaithersburg, MD	(301) 330-3733	www.therimmune	Yes	X	X	X	X	X		X	X	X			X	X	SP
TNO	Netherlands	+31 (0)30 694 41 44	www.voedig.tno.nl/	Yes	X											X		SP, WL

TNO Pharma	Utrechtsewe.g., 48 NL-3700AJ Zeist 3704 HE The Netherlands	+31 30 694 4806 (p) +31 30 694 4845 (f)	www.pharma.tno.nl	Yes		X	X	X	X	X	?	X		X		X	X	SP, WL
WIL Research Labs, Inc.	1407 George Rd. Ashland, OH 44805	(419) 2898700 (p) (419) 2893650 (f)	www.wilresearch.com	Yes		X	X	X	X	X	X	X	X	X		?	X	X	SP, WL

Note: A "+" listed in the additional services column indicates that there are more services available than just those listed; SP in the special studies column indicates safety pharmacology, WL in this column indicates wildlife testing is available.