
1

STRATEGY AND PHASING FOR DRUG SAFETY EVALUATION IN THE DISCOVERY AND DEVELOPMENT OF PHARMACEUTICALS

1.1. INTRODUCTION

The preclinical assessment of the safety of potential new pharmaceuticals represents a special case of the general practice of toxicology (Gad, 1996, 2000; Meyer, 1989), possessing its own peculiarities and special considerations, and differing in several ways from the practice of toxicology in other fields—for some significant reasons. Because of the economics involved and the essential close interactions with other activities, (e.g., clinical trials, chemical process optimization, formulation development, regulatory reviews, etc.), the development and execution of a crisp, timely and flexible, yet scientifically sound, program is a prerequisite for success. The ultimate aim of preclinical assessment also makes it different. A good pharmaceutical safety assessment program seeks to efficiently and effectively move safe, potential therapeutic agents into, and support them through, the clinical evaluation, then to registration, and, finally, to market. This requires the quick identification of those agents that are not safe. At the same time, the very biological activity which makes a drug efficacious also acts to complicate the design and interpretation of safety studies.

Pharmaceuticals, unlike industrial chemicals, agricultural chemicals, and environmental agents, are intended to have human exposure and biological activity. And, unlike these materials and food additives, pharmaceuticals are intended to have biological effects on the people that receive them. Frequently, the interpretation of results and the formulation of decisions about the continued development and

eventual use of a drug are based on an understanding of both the potential adverse effects of the agent (its safety) and its likely benefits, as well as the dose separation between these two (the “therapeutic index”). This makes a clear understanding of dose-response relationships critical, so that the actual risk/benefit ratio can be identified. It is also essential that the pharmacokinetics be understood and that “doses” (plasma tissue levels) at target organ sites be known (Scheuplein et al., 1990). Integral evaluation of pharmacokinetics are essential to any effective safety evaluation program.

The development and safety evaluation of pharmaceuticals have many aspects specified by regulatory agencies, and this has also tended to make the process more complex [until recently, as ICH (International Conference on Harmonization) has tended to take hold] as markets have truly become global. An extensive set of safety evaluations is absolutely required before a product is ever approved for market. Regulatory agencies have increasingly come to require not only the establishment of a “clean dose” in two species with adequate safety factors to cover potential differences between species, but also an elucidation of the mechanisms underlying such adverse effects as are seen at higher doses and are not well understood. These regulatory requirements are compelling for the pharmaceutical toxicologist (Traina, 1983; Smith, 1992). There is not, however, a set menu of what must be done. Rather, much (particularly in terms of the timing of testing) is open to professional judgment and is tailored for the specific agent involved and its therapeutic claim.

The discovery, development, and registration of a pharmaceutical is an immensely expensive operation, and represents a rather unique challenge (Zbinden, 1992). For every 9000 to 10,000 compounds specifically synthesized or isolated as potential therapeutics, one (on average) will actually reach the market. This process is illustrated diagrammatically in Figure 1.1. Each successive stage in the process is more expensive, making it of great interest to identify as early as possible those agents that are not likely to go the entire distance, allowing a concentration of effort on the compounds that have the highest probability of reaching the market. Compounds “drop out” of the process primarily for three reasons:

1. Toxicity or (lack of) tolerance.
2. (lack of) efficacy.
3. (lack of) bioavailability of the active moiety in man.

Early identification of poor or noncompetitive candidates in each of these three categories is thus extremely important (Fishlock, 1990), forming the basis for the use of screening in pharmaceutical discovery and development. How much and which resources to invest in screening, and each successive step in support of the development of a potential drug, are matters of strategy and phasing that are detailed in a later section of this chapter. *In vitro* methods are increasingly providing new tools for use in both early screening and the understanding of mechanisms of observed toxicity in preclinical and clinical studies (Gad, 1989b, 2001), particularly with the growing capabilities and influence of genomic and proteomic technologies.

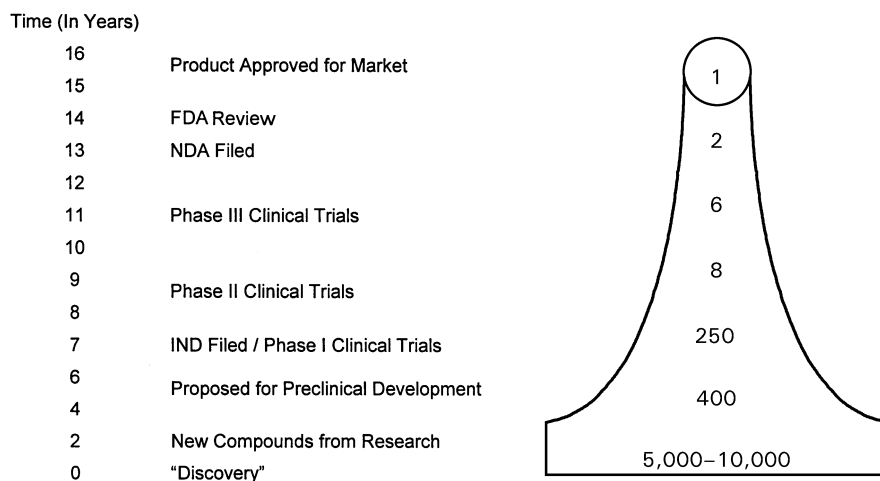


FIGURE 1.1. Attrition during the development of new molecules with a promise of therapeutic potential. Over the course of taking a new molecular entity through scale-up, safety and efficacy testing, and, finally, to market, typically only one out of every 9000 to 10,000 will go to the marketplace.

This is increasingly important as the societal concern over drug prices has grown (Littlehales, 1999). Additionally, the marketplace for new drugs is exceedingly competitive. The rewards for being either early (first or second) into the marketplace or achieving a significant therapeutic advantage are enormous in terms of eventual market share. Additionally, the first drug approved sets agency expectations for those drugs which follow. In mid-2001, there are 182 pharmaceutical products awaiting approval (41 of these are biotech products), the “oldest” having been in review seven years and some 1700 additional agents in the IND stage (Bryostowski, 2001). Not all of these (particularly the oldest) will be economically successful.

The successful operation of a safety assessment program in the pharmaceutical industry requires that four different phases of the product-related operation be simultaneously supported. These four phases of pharmaceutical product support [discovery support, investigation new drug (IND) support, clinical and registration support, and product support] constitute the vast majority of what is done by the safety assessment groups in the pharmaceutical industry. The constant adjustment of balance of resources between these four areas is the greatest management challenge in pharmaceutical safety assessment. An additional area, occupational toxicology, is conducted in a manner similar to that for industrial environments and is the subject of Chapter 14 of this volume. In most companies, occupational toxicology is the responsibility of a separate group.

The usual way in which transition (or “flow”) between the different phases is handled in safety assessment is to use a tiered testing approach. Each tier generates more specific data (and costs more to do so) and draws on the information generated in earlier tiers to refine the design of new studies. Different tiers are keyed to the

support of successive decision points (go/no-go points) in the development process, with the intent of reducing risks as early as possible.

The first real critical decisions concerning the potential use of a compound in humans are the most difficult. They require an understanding of how well particular animal models work in predicting adverse effects in humans (usually very well, but there are notable lapses; for example, giving false positives and false negatives), and an understanding of what initial clinical trials are intended to do. Though an approved IND grants one entry into limited evaluations of drug effects in man, flexibility in the execution and analysis of these studies offers a significant opportunity to also investigate efficacy (O'Grady and Linet, 1990).

Once past the discovery and initial development stages, the safety assessment aspects of the process become extremely tightly connected with the other aspects of the development of a compound, particularly the clinical aspects. These interconnections are coordinated by project management systems. At many times during the early years of the development process, safety assessment constitutes the rate-limiting step; it is, in the language of project management, on the critical path.

Another way in which pharmaceutical safety assessment varies from toxicology as practiced in other industries is that it is a much more multidisciplinary and integrated process. This particularly stands out in the incorporation of the evaluation of ADME (absorption, distribution, metabolism and excretion) aspects in the safety evaluation process. These pharmacokinetic–metabolism (PKM) aspects are evaluated for each of the animal model species (most commonly the rat and dog or primate) utilized to evaluate the preclinical systemic toxicity of a potential drug prior to evaluation in man. Frequently, *in vitro* characterizations of metabolism for model (or potential model) species and man are performed to allow optimal model selection and understanding of findings. This allows for an early appreciation of both the potential bioavailability of active drug moieties and the relative predictive values of the various animal models. Such data early on are also very useful (in fact, sometimes essential) in setting dose levels for later animal studies and in projecting safe dose levels for clinical use. Unlike most other areas of industrial toxicology, one is not limited to extrapolating the relationships between administered dose and systemic effects. Rather, one has significant information on systemic levels of the therapeutic moiety; typically, total area under the curve (AUC), peak plasma levels (C_{\max}), and plasma half-lives, at a minimum. Chapter 18 looks at these aspects in detail.

The state of the art for preclinical safety assessment has now developed to the point where the resulting products of the effort (reports, IND/NDA summaries, and the overall professional assessment of them) are expected to reflect and integrate the best effort of all the available scientific disciplines. Actual data and discussion should thus come from toxicology, pharmacology, pathology, and metabolism, at a minimum. The success of current premarket efforts to develop and ensure that only safe drugs make it to market are generally good, but clearly not perfect. This is reflected in popular (Arnst, 1998; Raeburn, 1999) and professional (Moore, et al., 1998; Lazarou et al., 1998) articles looking at both the number of recent marketed drug withdrawals for safety (summarized in Table 1.1) and at rates of drug-related

TABLE 1.1. Post-approval Adverse Side Effects and Related Drug Withdrawals Since 1990

- 51% of approval Drugs had serious post-approval identified side effects
- FDAMA passed in 1997

Year	Drug	Indication/Class	Causative Side Effect
1991	Enkaid (4 years on market)	Antiarrhythmic	Cardiovascular (sudden cardiac death)
1992	Temafloxacin	Antibiotic	Blood & Kidney
1997	Fenfuramine*/Dexafuramine (Combo used since 1984) (*24 years on market)	Diet pill	Heart Valve Abnormalities
1998	Posicor (Midefradil) (1 year on market)	Ca ⁺⁺ Channel Blocker	Lethal Drug Interactions (Inhibited Liver Enzymes)
1999	Duract (Bronfemic Sodium) (Early preapproval warnings of liver enzymes)	Pain Relief	Liver Damage
	Tronan (use severely restricted)	Antibiotic	Liver/Kidney Damage
	Raxar	Quinolone antibiotic	QT internal prolongation/ ventricular arrhythmias (deaths)
	Hismanal	Antihistamine	Drug-drug interactions
2000	Rotashield Renzulin (Approved Dec 1996)	Rotavirus Vaccine Type II Diabetes	Bowel Obstruction Liver Damage
2001	Propulsid Lotonox Phenypropanolamine (PPA) Baychlor	Heartburn Irritable Bowel Syndrome OTC ingredient Cholesterol reducing (statin)	Cardiovascular Irregularities/Deaths Ischemic colitis/death Hemorrhagic stroke Rhabdomyolysis (muscle- weakening) (deaths)

AALAC certified laboratory. In-housing testing included acute, subacute, and subchronic oral, dermal and inhalation studies and specialty reproductive, behavioural, haematological and renal function toxicity studies. Preparation of risk assessment, submissions and presentations to regulatory agencies and trade association.

adverse drug events and deaths in hospital patients. It is hoped that this system can be improved, and there are a lot of efforts to improve or optimize drug candidate selection and development (Lesko, et al., 2000).

1.2. REGULATORY REQUIREMENTS

Minimum standards and requirements for safety assessment of new pharmaceuticals are established by the need to meet regulatory requirements for developing, and eventually gaining approval to market, the agent. Determining what these requirements are is complicated by (1) the need to compete in a global market, which means gaining regulatory approval in multiple countries that do not have the same standards or requirements, and (2) the fact that the requirements are documented as guidelines, the interpretation of which is subject to change as experience alters judgments. The ICH process has much improved this situation, as detailed in Chapter 2.

Standards for the performance of studies (which is one part of regulatory requirements) have as their most important component good laboratory practices (GLPs). Good laboratory practices largely dictate the logistics of safety assessment: training, adherence to other regulations (such as those governing the requirements for animal care), and (most of all) the documentation and record-keeping that are involved in the process. There are multiple sets of GLP regulations (in the United States alone, agencies such as the FDA and EPA each have their own) that are not identical; however, adherence to U.S. Food and Drug Administration GLPs (FDA, 1987a) will rarely lead one astray.

Not all studies that are done to assess the preclinical safety of a new pharmaceutical need be done in strict adherence to GLPs. Those studies that are “meant to support the safety of a new agent” (i.e., are *required* by regulatory guidelines) must be so conducted or run a significant risk of rejection. However, there are also many other studies of an exploratory nature (such as range finders and studies done to understand the mechanisms of toxicity) that are not required by the FDA, and which may be done without strict adherence to GLPs. A common example are those studies performed early on to support research in selecting candidate agents. Such studies do not meet the requirements for having a validated analytical method to verify the identity, composition, and stability of materials being assayed, yet they are essential to the processes of discovery and development of new drugs. All such studies must eventually be reported to the FDA if an IND application is filed, but the FDA does not in practice “reject” such studies (and therefore the IND) because they are “non-GLP.”

There is a second set of “standards” of study conduct that are less well defined. These are “generally accepted practice,” and though not written down in regulation, are just as important as GLPs for studies to be accepted by the FDA and the scientific community. These standards, which are set by what is generally accepted as good science by the scientific community, include techniques, instruments utilized, and interpretation of results. Most of the chapters in this book will reflect these generally accepted practices in one form or another.

Guidelines establish which studies must be done for each step in the process of development. Though guidelines supposedly are suggestion (and not requirements), they are in fact generally treated as minimums by the promulgating agency. The exceptions to this are special cases where a drug is to meet some significant need (a life-threatening disease such as AIDS) or where there are real technological limitations as to what can be done (as with many of the new biologically derived [or biotechnology] agents, where limitations on compound availability and biological responses make traditional approaches inappropriate).

There are some significant differences in guideline requirements between the major countries [see Alder and Zbinden (1988) for an excellent country-by-country review of requirements], though this source is now becoming dated. The core of what studies are generally done are those studies conducted to meet U.S. FDA requirements. These are presented in Table 1.2. As will be discussed in Chapter 2, these guidelines are giving way to the ICH guidelines. However, while the length and details of studies have changed, the nature and order of studies remain the same.

The major variations in requirements for other countries still tend to be in the area of special studies. The United States does not formally require any genotoxicity studies, but common practice for U.S. drug registration is to perform at least a bacterial gene mutation assay (Ames test), a mammalian cell mutation assay and a clastogenicity assay, while Japan requires specific tests, including a gene mutation assay in *Escherichia coli*. Likewise, the European Economic Community (EEC) has a specified set of requirements, while individual countries have additional special requirements (Italy, for example, requires a mutagenicity assay in yeast). As detailed in Chapter 6, the new ICH genotoxicity guidelines have come to meet multinational requirements. Japan maintains a special requirement for an antigenicity assay in guinea pigs. The new safety pharmacology requirements are likely to be adopted over a period of time by different adherents.

It is possible to interact with the various regulatory agencies (particularly the FDA) when peculiarities of science or technology leave one with an unclear understanding of what testing is required. It is best if such discussions directly involve the scientists who understand the problems, and it is essential that the scientists at the FDA be approached with a course of action (along with its rationale) that has been proposed to the agency in advance.

The actual submissions to a regulatory agency that request permission either to initiate (or advance) clinical trials of a drug, or to market a drug, are not just bundles of reports on studies. Rather, they take the form of summaries that meet mandated requirements for format, accompanied by the reports discussed in these summaries (Guarino, 1987). In the United States, these summaries are the appropriate section of the IND and the New Drug Application (NDA). The formats for these documents have recently been revised (FDA, 1987b). The EEC equivalent is the expert report, as presented in EEC Directive 75/319. Similar approaches are required by other countries. In each of these cases, textual summaries are accompanied by tables that also serve to summarize significant points of study design and of study findings.

All of these approaches have in common that they are to present integrated evaluations of the preclinical safety data that are available on a potential new drug.

TABLE 1.2. Synopsis of General Guidelines for Animal Toxicity Studies (U.S. FDA)

Category	Duration of human administration ^a	Phase ^b	Subacute or chronic toxicity ^c	Special studies
Oral or parenteral	Several days (up to 3)	I, II, III, NDA	2 species: 2 weeks	For parenterally administered drugs; compatibility with blood and local tolerance at injection site where applicable.
	Up to 2 weeks	I	2 species: 4 weeks	
		II	2 species: up to 4 weeks	
		III, NDA	2 species: up to 3 months	
	Up to 3 months	I, II III NDA	2 species: 4 weeks 2 species: 3 months 2 species: up to 6 months	
	6 months to unlimited	I, II III NDA	2 species: 3 months 2 species: 6 months or longer 2 species: 12 months in rodents, 9 months in nonrodents + 2 rodent species for CA; 18 months (mouse)— may be met by use of a transgenic model 24 months (rat)	

Inhalation (general anesthetics)	Single administration	I, II, III, NDA	4 species: 5 days (3 h/day)	
Dermal	Single application	I	1 species: single 24-h exposure followed by 2-week observation	Sensitization
	Single or short-term application	II	1 species: 20-day repeated exposure (intact and abraded skin)	
	Short-term application Unlimited application	III NDA	As above As above, but intact skin study extended up to 6 months	
Ophthalmic	Single application Multiple application	I I, II, III	1 species: 3 weeks, daily applications, as in clinical use	Eye irritation tests with graded doses
		NDA	1 species: duration commensurate with period of drug administration	
Vaginal or rectal	Single application Multiple application	I I, II, III, NDA	2 species: duration and number of applications determined by proposed use	Local and systematic toxicity after vaginal or rectal application in 2 species

(continued)

TABLE 1.2. (continued)

Category	Duration of human administration ^a	Phase ^b	Subacute or chronic toxicity ^c	Special studies
Drug combinations ^d	I II, III, NDA	2 species: up to 3 months		Lethality by appropriate route, compared to components run concurrently in 1 species

^a Phase I dosing of females if childbearing potential requires a Segment II study in at least one species; Phase III dosing of this population requires a Segment I study and both Segment II studies.

^b Phase I, II, and III are defined in Section 130.0 of the New Drug Regulations.

^c Acute toxicity should be determined in three species; subacute or chronic studies should be by the route to be used clinically. Suitable mutagenicity studies should also be performed.

Observations:

Body weights

Metabolic studies

Gross and microscopic examination

Coagulation tests

^d Where toxicity data are available on each drug individually

Food consumption

Ophthalmologic examination

Hemogram

Others as appropriate

Where toxicity data are available on each drug individually

Behavior

Fasting blood sugar

Liver and kidney function tests

The individual studies and reports are to be tied together to present a single, cohesive overview of what is known about the safety of a drug.

Lever (1987) presents an excellent overview of the regulatory process involved in FDA oversight of drug development, and gives the historical perspective for the evolution of the conservative process that is designed to ensure that any new pharmaceutical is both safe and efficacious.

There are other regulatory, legal and ethical safety assessment requirements beyond those involved in the selection and marketing of a drug as a product entity. The actual drug product must be manufactured and transported in a safe manner, and any waste associated with this manufacture disposed of properly. Chapter 14 of this volume specifically addresses this often overlooked aspect of safety assessment programs.

1.3. ESSENTIAL ELEMENTS OF PROJECT MANAGEMENT

It is important to keep in mind that safety assessment is only one of many components involved in the discovery and development of new pharmaceuticals. The entire process has become enormously expensive, and completing the transit of a new drug from discovery to market has to be as efficient and expeditious a process as possible. Even the narrow part of this process (safety assessment) is dependent on many separate efforts. Compounds must be made, analytical and bioanalytical methods developed, and dosage formulations developed, to name a few. One needs only to refer to Beyer (1978), Hamner (1982), Matoren (1984), Sneader (1986) (a good short overview), Zbinden, (1992) or Spilker (1994) for more details on this entire process and all of its components.

The coordination of this entire complex process is the province of project management, the objective of which is to ensure that all the necessary parts and components of a project mate up. This discipline in its modern form was first developed for the Polaris missile project in the 1960s. Its major tool that is familiar to pharmaceutical scientists is the “network” or PERT (Program Evaluation Review Technique) chart, as illustrated in Figure 1.2. This chart is a tool that allows one to see and coordinate the relationships between the different components of a project. One outcome of the development of such a network is identification of the rate-limiting steps, which, in aggregate, comprise the critical path (see Table 1.3 for a lexicon of the terms used in project management).

A second graphic tool from project management is the Gantt chart, as illustrated in Figure 1.3. This chart allows one to visualize the efforts underway in any one area, such as safety assessment, for all projects that are currently being worked on.

Figure 1.4 is a hybrid from of the PERT and Gantt charts, designed to allow one to visualize all the resources involved in any one project.

An understanding of the key concepts of project management and their implications are critical for strategic planning and thinking for safety assessment. Kliem (1986) and Knutson (1980) offer excellent further reading in the area of project management.

TABLE 1.3. Glossary of Project Management Terms

Activity	The work or effort needed to complete a particular event. It consumes time and resources.
Average daily resource requirement	The likely amount of resources required to complete an activity or several activities on any workday during a project. The average daily labor requirement is one example.
CPM	Acronym for Critical Path Method. A network diagramming technique that places emphasis on time, cost, and the completion of events.
Critical path	The longest route through a network that contains activities absolutely crucial to the completion of the project.
Dummy arrow	A dashed line indicating an activity that uses no time or resources.
Duration	The time it takes to complete an activity.
Earliest finish	The earliest time an activity can be completed.
Earliest start	The earliest an activity can begin if all activities before it are finished. It is the earliest time that an activity leaves its initiation node.
Event	A synonym for node. A point in time that indicates the accomplishment of a milestone. It consumes neither time nor resources and is indicated whenever two or more arrows intersect.
Free float	The amount of time that an activity can be delayed without affecting succeeding activities.
Gantt chart	A bar chart indicating the time interval for each of the major phases of a project.
Histogram	A synonym for bar chart.
Latest finish	The latest time an activity can be completed without extending the length of a project.
Latest start	The latest time an activity can begin without lengthening a project.
Leveling	The process of “smoothing” out labor, material and equipment requirements to facilitate resource allocation. The project manager accomplishes this by “rescheduling” noncritical activities so that the total resource requirements for a particular day match the average daily resource requirements.
Most likely time	Used in PERT diagramming. The most realistic time estimate for completing an activity or project under normal conditions.
Node	A synonym for event.
Optimistic time	Used in PERT diagramming. The time the firm can complete an activity or project under the most ideal conditions.
PERT	Acronym for Program Evaluation and Review Technique. A network diagramming technique that places emphasis on the completion of events rather than cost or time.
Pessimistic time	Used in PERT diagramming. The time the firm can complete an activity or project under the worst conditions.
Project	The overall work or effort being planned. It has only one beginning node and ending node. Between those nodes are countless activities and their respective nodes.
Project phase	A major component, or segment, of a project. It is determined by the process known as project breakdown structuring.
Total float	The total amount of flexibility in scheduling activities on a noncritical path. Hence, it provides the time an activity could be prolonged without extending a project’s final completion date.

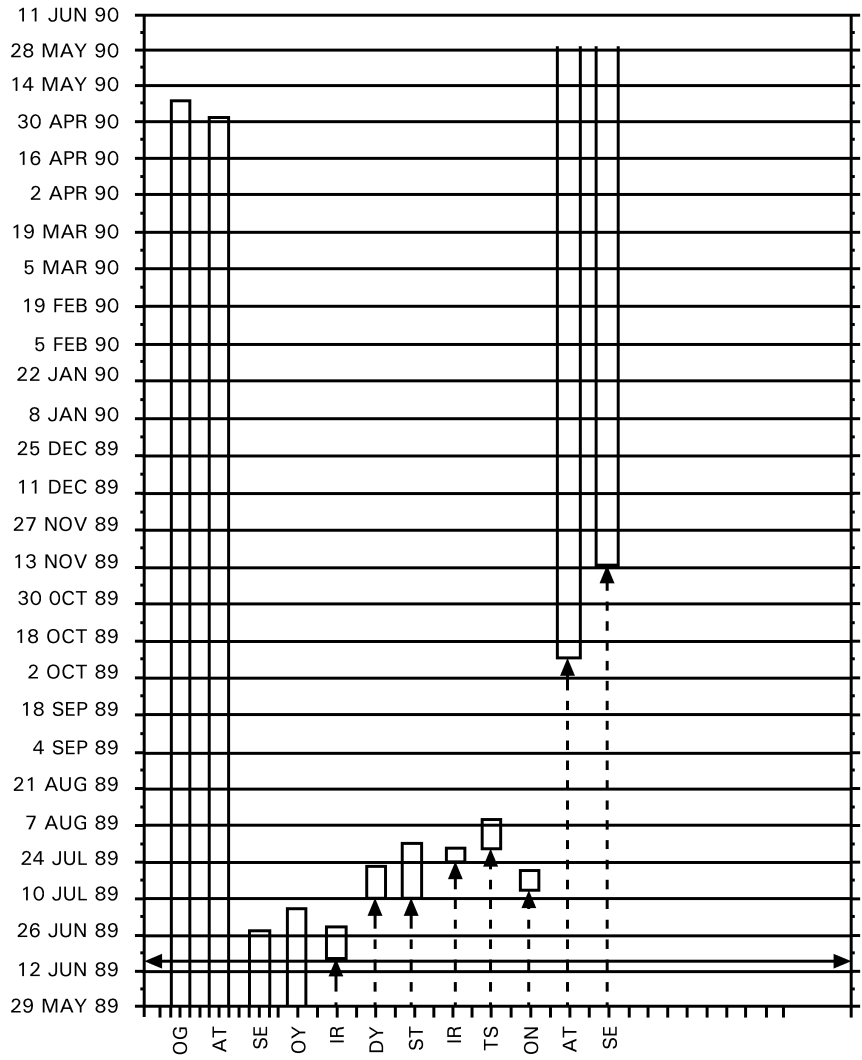


FIGURE 1.3. A GANTT or bar chart showing scheduling of the major safety assessment activities (studies) involved in a pharmaceutical development project.

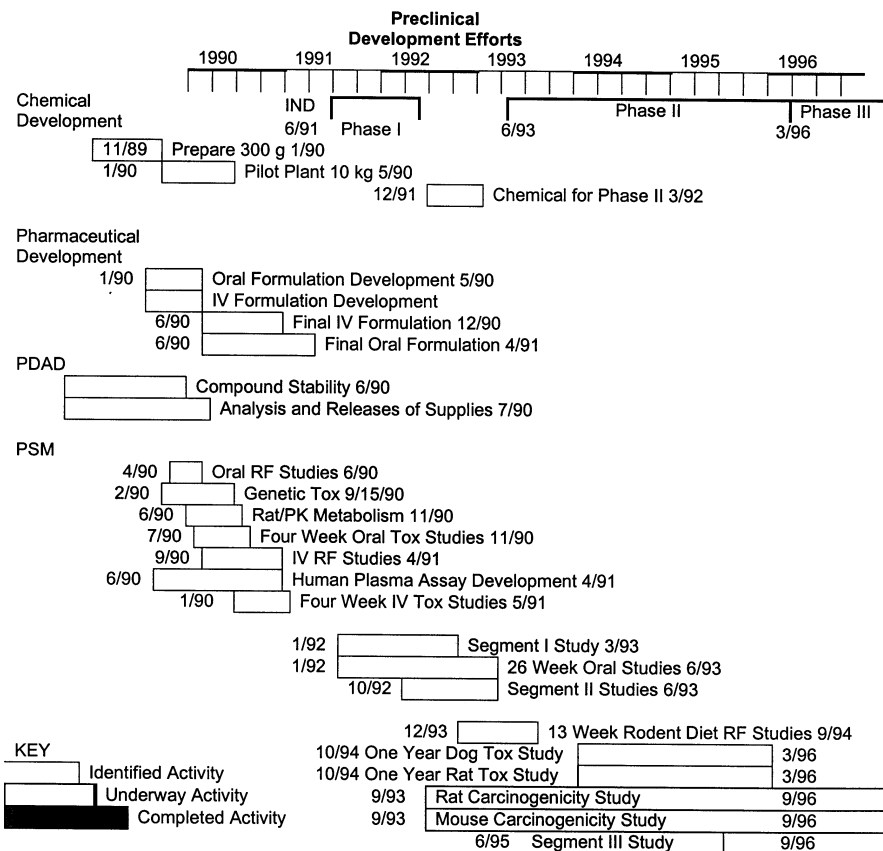


FIGURE 1.4. A hybrid project GANTT chart, which identifies the work of each of the development functions ("line operations") in the development of a new compound and how it matches the phase of development.

1.4. SCREENS: THEIR USE AND INTERPRETATION IN SAFETY ASSESSMENT

Much (perhaps even most) of what is performed in safety assessment can be considered screening, trying to determine if some effect is or is not (to an acceptable level of confidence) present (Zbinden et al., 1984). The general concepts of such screens are familiar to toxicologists in the pharmaceutical industry because the approach is a major part of the activities of the pharmacologists involved in the discovery of new compounds. But the principles underlying screening are not generally well recognized or understood. And such understanding is essential to the proper use, design, and analysis of screens (Gad, 1988a, 1989a). Screens are the biological equivalent of exploratory data analysis, or EDA (Tukey, 1977).

Each test or assay has an associated activity criterion, that is, a level above which the activity of interest is judged to be present. If the result for a particular test compound meets this criterion, the compound may pass to the next stage. This criterion could be based on statistical significance (e.g., all compounds with observed activities significantly greater than the control at the 5% level could be tagged). However, for early screens, such a formal criterion may be too strict, resulting in few compounds being identified as “active.”

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

Statistical modeling of the assay system may lead to the improvement of the design of the system by reducing the interval between discoveries of active compounds. The objectives behind a screen and considerations of (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 per group, and increase the range and number of doses. The result will be less information on more structure, but there will be an overall increase in the frequency of discovery of active compounds (assuming that truly active compounds are entering the system at a steady rate).

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

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

1.4.1. Characteristics of Screens

An excellent introduction to the characteristics of screens is Redman's (1981) interesting approach, which identifies four characteristics of an assay. Redman assumes that a compound is either active or inactive and that the proportion of

activities in a compound can be estimated from past experience. After testing, a compound will be classified as positive or negative (i.e., possessing or lacking activity). It is then possible to design the assay so as to optimize the following characteristics.

1. Sensitivity: the ratio of true positives to total activities;
2. Specificity: the ratio of true negatives to total inactives;
3. Positive accuracy: the ratio of true to observed positives;
4. Negative accuracy: the ratio of true to observed negatives;
5. Capacity: the number of compounds that can be evaluated;
6. Reproducibility: the probability that a screen will produce the same result at another time (and, perhaps, in some other lab).

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

To simplify this presentation, datasets drawn only from neuromuscular screening activity were used. However, the evaluation and approaches should be valid for all similar screening datasets, regardless of source. The methods are not sensitive to the biases introduced by the degree of interdependence found in many screening batteries that use multiple measures (such as the neurobehavioral screen).

1. Screens almost always focus on detecting a single endpoint of effect (such as mutagenicity, lethality, neurotoxicity, or development toxicity), and have a particular set of operating characteristics in common.
2. A large number of compounds are evaluated, so ease and speed of performance (which may also be considered efficiency) are very desirable characteristics.
3. The screen must be very sensitive in its detection of potential effective agents. An absolute minimum of active agents should escape detection; that is, there should be very few false negatives (in other words, the type II error rate or beta level should be low). Stated yet another way, the signal gain should be way up.
4. It is desirable that the number of false positives be small (i.e., there should be a low type I error rate or alpha level).
5. Items (2)–(4), which are all to some degree contradictory, require the involved researchers to agree on a set of compromises, starting with the acceptance of a relatively high alpha level (0.10 or more), that is, an increased noise level.

6. In an effort to better serve item (2), safety assessment screens are frequently performed in batteries so that multiple 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 (3).
7. This screen should use small amounts of compound to make item (1) possible and should allow evaluation of materials that have limited availability (such as novel compounds) early on in development.
8. Any screening system should be validated initially using a set of blind (positive and negative) controls. These blind controls should also be evaluated in the screening system on a regular basis to ensure continuing proper operation of the screen. As such, the analysis techniques used here can then be used to ensure the quality or modify 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.
10. Sample (group) sizes are generally small.
11. The data tend to be imprecisely gathered (often because researchers are unsure of what they are looking for), and therefore possess extreme within-group variability. Control and historical data are not used to adjust for variability or modify test performance.
12. Proper dose selection is essential for effective and efficient screen design and conduct. If insufficient data are available, a suitably broad range of doses must be evaluated (however, this technique is undesirable on multiple grounds, as has already been pointed out).

The design, use and analysis of screens is covered in detail in Chapter 4 of this volume.

1.5. STRATEGY AND PHASING

Regulatory requirements and our understanding of the pharmacology, marketing, and clinical objectives for a potential product provide a framework of requirements for the safety assessment of potential new pharmaceuticals. How one meets these requirements is not fixed, however. Rather, exactly what is done and when activities are performed are reflections of the philosophy and managerial climate of the organization that is doing the discovery and development. It should be kept in mind that establishing and maintaining an excellent information base on the biological basis for a compound's expected therapeutic activity and safety is essential but often left undone. This subject is addressed in Chapter 2 of this volume.

There are multiple phases involved in the safety assessment portion of the discovery, development and marketing process. The actual conduct of the studies in each phase forms the basis of the bulk of the chapters in this book. However, unless

the pieces are coordinated well and utilized effectively (and completed at the right times), success of the safety assessment program is unlikely or very expensive.

First, support needs to be given to basic research (also called discovery, biology, or pharmacology in different organizations) so that it can efficiently produce a stream of potential new product compounds with as few overt toxicity concerns as possible. This means that there must be early and regular interaction between the individuals involved, and that safety assessment must provide screening services to rank the specific safety concerns of the compounds. These screens may be *in vitro* (both for genetic and nongenetic endpoints) or *in vivo* (designed on purpose for a single endpoint, such as effects on reproductive performance, promotion activity, etc.). There must also be ongoing work to elucidate the mechanisms and structure-activity relationships behind those toxicities that are identified (Gad, 1989b).

Second is the traditional core of safety assessment that is viewed as development. Development includes providing the studies to support compounds getting into the clinic (an IND application being filed and accepted), evaluating a compound to the point at which it is considered safe, able to be absorbed, and effective (clinical phase II), and, finally, registration (filing an NDA and having it approved). Various organizations break this process up differently. Judgements are generally made on the likelihood of compounds failing (“dying”) at different stages in the clinical development process, and the phasing of preclinical support is selected and or adjusted accordingly. If an organization has a history of many compounds failing early in the clinic (such as in the initial phase I tolerance trials, where there may be only three to ten days of human dosing), then initial “pivotal” preclinical studies are likely to be only four-week-long studies. If compounds tend to fail only in longer efficacy trials, then it is more efficient to run longer initial preclinical trials. Figure 1.5 shows several variations on these approaches. Additionally, the degree of risk involved in study design (particularly in dose selection) is also an organizational characteristic. Pivotal studies can fail on two counts associated with dose selection. Either they cannot identify a “safe” (no-effect) dose or they can neglect to find a dose that demonstrates a toxic effect (and therefore allows identification of potential target organs). Therefore, picking the doses for such studies is an art that has been risky because, traditionally, only three different dose groups have been used, and before clinical trials are conducted there is at best a guess as to what clinical dose will need to be cleared. The use of four (or five) dose groups only marginally increases study cost, and, in those cases where the uncertainty around dose selection is great, provides a low-cost alternative to repeating the study.

Pivotal studies can also be called shotgun tests, because it is unknown in advance what endpoints are being aimed at. Rather, the purpose of the study is to identify and quantitate all potential systemic effects resulting from a single exposure to a compound. Once known, specific target organ effects can then be studied in detail if so desired. Accordingly, the generalized design of these studies is to expose groups of animals to controlled amounts or concentrations of the material of interest, and then to observe for and measure as many parameters as practical over a period past or during the exposure. Further classification of tests within this category would be the route by which test animals are exposed/dosed or by the length of dosing.

PLAN 1: Clinical Decision Point** Is Short Term Tolerance Or Human Pharmacokinetics

<u>Range Finder</u>	<u>Pivotal Study</u>	<u>Phase I</u>	<u>Phase II</u>
"Pyramid" or "rolling acute" (see Chapter 4)	→ Two or four weeks in two species by the intended route	→ Tolerance and pharmacokinetics (PK) with up to 3 to 14 days human dosing	→ DELAY*** →

PLAN 2: Clinical Decision Point Is An Indication Of Efficacy In Man

<u>Range Finder</u>	<u>Pivotal Study</u>	<u>Phase I/II</u>	<u>Phase III/III</u>
"Pyramid" or "rolling acute" and/or two or four week study	→ Thirteen weeks in two species	→ Tolerance, PK and efficacy with human dosing up to one month in length	→ DELAY*** →

PLAN 3: Plan For Success Or Resources Are Not A Constraint

<u>Range Finder</u>	<u>Pivotal Study</u>	<u>Phase I/II</u>	<u>Phase III</u>
"Pyramid" or "rolling acute: and/or two week study	Four weeks in two species	<u>Preclinical</u> Thirteen weeks in two species	<u>Preclinical</u> One year in two species Carcinogenicity in two rodents (if required)

FIGURE 1.5. Three different approaches to matching preclinical safety efforts to support the clinical development of a new drug. Which is the best one for any specific case depends on considerations of resource availability and organizational tolerance of "risk." In Plan 1, little effort will be "wasted" on projects that fail during early (phase I) clinical trials—but if phase I trials are successful, there will be major delays. In Plan 3, clinical development will never be held up waiting for more safety work, but a lot of effort will go into projects that never get past Phase I. Plan 2 is a compromise. Delays are to allow additional preclinical (animal safety) studies to support longer clinical trials in accordance with FDA or other applicable guidelines.

"Acute," for example, implies a single exposure interval (of 24 h or less) or dose of test material. Using the second scheme (length of dosing), the objectives of the successive sets of pivotal studies could be defined as follows:

Acute studies:

1. Set doses for next studies.
2. Identify very or unusually toxic agents.
3. Estimate lethality potential.
4. Identify organ system affected.

Two-week studies:

1. Set doses for next studies.
2. Identify organ toxicity.

3. Identify very or unusually toxic agents.
4. Estimate lethality potential.
5. Evaluate potential for accumulation of effects.
6. Get estimate of kinetic properties (blood sampling/urine sampling).

Four-week studies:

1. Set doses for next studies.
2. Identify organ toxicity.
3. Identify very or unusually toxic agents.
4. Estimate lethality potential.
5. Evaluate potential for accumulation of effects.
6. Get estimate of kinetic properties (blood sampling/urine sampling).
7. Elucidate nature of specific types of target organ toxicities induced by repeated exposure.

Thirteen-week studies:

1. Set doses for next studies.
2. Identify organ toxicity.
3. Identify very or unusually toxic agents.
4. Evaluate potential for accumulation of effects.
5. Evaluate pharmacokinetic properties.
6. Elucidate nature of specific types of target organ toxicities induced by repeated exposure.
7. Evaluate reversibility of toxic effects.

Chronic studies:

1. Elucidate nature of specific types of target organ toxicities induced by prolonged repeated exposure.
2. Identify potential carcinogens.

The problems of scheduling and sequencing toxicology studies and entire testing programs have been minimally addressed in print. Though there are several books and many articles available that address the question of scheduling multiple tasks in a service organization (French, 1982), and an extremely large literature on project management (as briefly overviewed earlier in this chapter), no literature specific to a research testing organization exists.

For all the literature on project management, however, a review will quickly establish that it does not address the rather numerous details that affect study/program scheduling and management. There is, in fact, to my knowledge, only a single article (Levy et al., 1977) in the literature that addresses scheduling, and it describes a computerized scheduling system for single studies.

There are commercial computer packages available for handling the network construction, interactions, and calculations involved in what, as will be shown below, is a complicated process. These packages are available for use on both mainframe and microcomputer systems.

Scheduling for the single study case is relatively simple. One should begin with the length of the actual study and then factor in the time needed before the study is started to secure the following resources:

- Animals must be on hand and properly acclimated (usually for at least two weeks prior to the start of the study).
- Vivarium space, caging, and animal care support must be available.
- Technical support for any special measurements such as necropsy, hematology, urinalysis, and clinical chemistry must be available on the dates specified in the protocol.
- Necessary and sufficient test material must be on hand.
- A formal written protocol suitable to fill regulatory requirements must be on hand and signed.

The actual study (from first dosing or exposure of animals to the last observation and termination of the animals) is called the in-life phase, and many people assume the length of the in-life phase defines the length of a study. Rather, a study is not truly completed until any samples (blood, urine, and tissue) are analyzed, slides are prepared and microscopically evaluated, data are statistically analyzed, and a report is written, proofed, and signed off. Roll all of this together, and if you are conducting a single study under contract in an outside laboratory, an estimate of the least time involved in its completion should be equal to (other than in the case of an acute or single and point study) no more than

$$L + 6 \text{ weeks} + \frac{1}{2}L,$$

where L is the length of the study. If the study is a single endpoint study and does not involve pathology, then the least time can be shortened to $L + 6$ weeks. In general, the best that can be done is $L + 10$ weeks.

When one is scheduling out an entire testing program on contract, it should be noted that, if multiple tiers of tests are to be performed (such as acute, two-week, thirteen-week, and lifetime studies), then these must be conducted sequentially, as the answer from each study in the series defines the design and sets the doses for the subsequent study.

If, instead of contracting out, one is concerned with managing a testing laboratory, then the situation is considerably more complex. The factors and activities involved are outlined below. Within these steps are rate-limiting factors that are invariably due to some critical point or pathway. Identification of such critical factors is one of the first steps for a manager to take to establish effective control over either a facility or program.

Before any study is actually initiated, a number of prestudy activities must occur (and, therefore, these activities are currently underway, to one extent or another, for the studies not yet underway but already authorized or planned for this year for any laboratory).

- Test material procurement and characterization.
- Development of formulation and dosage forms for study.
- If inhalation study, development of generation and analysis methodology, chamber trials, and verification of proper chamber distribution.
- Development and implementation of necessary safety steps to protect involved laboratory personnel.
- Arrangement for waste disposal.
- Scheduling to assure availability of animal rooms, manpower, equipment, and support services (pathology and clinical).
- Preparation of protocols.
- Animal procurement, health surveillance, and quarantine.
- Preparation of data forms and books.
- Conduct of prestudy measurements on study animals to set baseline rates of body weight gain and clinical chemistry values.

After completion of the in-life phase (i.e., the period during which live animals are used) of any study, significant additional effort is still required to complete the research. This effort includes the following.

- Preparation of data forms and books. Preparation of tissue slides and microscopic evaluation of these slides;
- Preparation of data tables;
- Statistical analysis of data;
- Preparation of reports.

There are a number of devices available to a manager to help improve the performance of a laboratory involved in these activities. One such device (cross-training) is generally applicable enough to be particularly attractive.

Identification of rate-limiting steps in a toxicology laboratory over a period of time usually reveals that at least some of these are variable (almost with the season). At times, there is too much work of one kind (say, inhalation studies) and too little of another (say, dietary studies). The available staff for inhalation studies cannot handle

this peak load and since the skills of these two groups are somewhat different, the dietary staff (which is now not fully occupied) cannot simply relocate down the hall and help out. However, if, early on, one identifies low- and medium-skill aspects of the work involved in inhalation studies, one could cross-train the dietary staff at a convenient time so that it could be redeployed to meet peak loads.

It should be kept in mind that there are a number of common mistakes (in both the design and conduct of studies and in how information from studies is used) that have led to unfortunate results, ranging from losses in time and money and the discarding of perfectly good potential products to serious threats to people's health. Such outcomes are indeed the great disasters in product safety assessment, especially since many of them are avoidable if attention is paid to a few basic principles.

It is quite possible to design a study for failure. Common shortfalls include

1. Using the wrong animal model.
2. Using the wrong route or dosing regimen.
3. Using the wrong vehicle or formulation of test material.
4. Using the wrong dose level. In studies where several dose levels are studied, the worst outcome is to have an effect at the lowest dose level tested (i.e., the safe dosage in animals remains unknown). The next worst outcome is to have no effect at the highest dose tested (generally meaning that the signs of toxicity remain unknown, invalidating the study in the eyes of many regulatory agencies).
5. Making leaps of faith. An example is to set dosage levels based on others' data and to then dose all test animals. At the end of the day, all animals in all dose levels are dead. The study is over; the problem remains.
6. Using the wrong concentration of test materials in a study. Many effects (including both dermal and gastrointestinal irritation, for example) are very concentration dependent.
7. Failing to include a recovery (or rebound) group. If one finds an effect in a 90-day study (say, gastric hyperplasia), how does one interpret it? How does one respond to the regulatory question, "Will it progress to cancer?" If an additional group of animals were included in dosing, then were maintained for a month after dosing had been completed, recovery (reversibility) could be both evaluated and (if present) demonstrated.

Additionally, there are specialized studies designed to address endpoints of concern for almost all drugs (carcinogenicity, reproductive or developmental toxicity) or concerns specific to a compound or family of compounds (local irritation, neurotoxicity, or immunotoxicity, for example). When these are done, timing also requires careful consideration. It must always be kept in mind that the intention is to ensure the safety of people in whom the drug is to be evaluated (clinical trials) or used therapeutically. An understanding of special concerns for both populations should be considered essential.

Safety evaluation does not cease being an essential element in the success of the pharmaceutical industry once a product is on the market. It is also essential to support marketed products and ensure that their use is not only effective but also safe and unclouded by unfounded perceptions of safety problems. This requires not only that clinical trials be monitored during development (Spector et al., 1988), but also that experience in the marketplace be monitored.

The design and conduct of safety assessment studies and programs also require an understanding of some basic concepts:

1. The studies are performed to establish or deny the safety of a compound, rather than to characterize the toxicity of a compound.
2. Because pharmaceuticals are intended to affect the functioning of biological systems, and safety assessment characterizes the effects of higher-than-therapeutic doses of compounds, it is essential that one be able to differentiate between hyperpharmacology and true (undesirable) adverse effects.
3. Focus of the development process for a new pharmaceutical is an essential aspect of success, but is also difficult to maintain. Clinical research units generally desire to pursue as many or as broad claims as possible for a new agent, and frequently also apply pressure for the development of multiple forms for administration by different routes. These forces must be resisted because they vastly increase the work involved in safety assessment, and they may also produce results (in one route) that cloud evaluation [and impede Institutional Review Board (IRB) and regulatory approval] of the route of main interest.

1.6. CRITICAL CONSIDERATIONS

In general, what the management of a pharmaceutical development enterprise wants to know at the beginning of a project are three things: what are the risks (and how big are they), how long will it take, and how much (money and test compound) will it take?

The risks question is beyond the scope of this volume. The time question was addressed earlier in this chapter. How much money is also beyond the scope of this volume. But calculating projected compound needs for studies is a fine challenge in the design and conduct of a safety evaluation program. The basic calculation is simple. The amount needed for a study is equal to

$$N W I L D,$$

where

N = the number of animals per group.

W = the mean weight per animal during the course of the study (in kg).

I = the total number of doses to be delivered (such as in a 28-day study, 28 consecutive doses).

L = a loss or efficiency factor (to allow for losses in formulation and dose delivery, a 10% factor is commonly employed, meaning a value of 1.1 is utilized).

D = the total dose factor. This is the sum of all the dose levels. For example, if the groups are to receive 1000, 300, 100, 30 and 30 mg/kg, then the total dose factor is $1000 + 300 + 100 + 30$ or 1430 mg/kg.

As an example, let's take a 28-day study in rats where there are 10 males and 10 females per group and the dose levels employed at 1000, 3000, 100 and 30 mg/kg. Over the course of the 28 days the average weight of the rats is likely to be 300 g (or 0.3 kg). This means our values are

$$N = 20$$

$$W = 0.3 \text{ kg}$$

$$I = 28$$

$$L = 1.1$$

$$D = 1430 \text{ mg/kg}$$

and therefore our total compound needs will be $(20)(0.3)(28)(1.1)(1430 \text{ mg}) = 2642.64 \text{ mg}$ or 0.264 kg. This is the simplest case, but shows the principles.

A governing principle of pharmaceutical safety assessment is the determination of safety factors: the ratio between the therapeutic dose (that which achieves the desired therapeutic effect) and the highest dose which evokes no toxicity. This grows yet more complex (but has less uncertainty) if one bases these ratios on plasma levels rather than administered doses. Traditionally based on beliefs as to differences of species sensitivity, it has been held that a minimum of a five-fold (5X) safety factor should be observed based on toxicity findings in nonrodents and a ten-fold (10X) based on rodents.

The desire to achieve at least such minimal therapeutic indices and to also identify levels associated with toxicity (and the associated toxic effects) form the basis of dose selection for systemic (and most other *in vivo*) toxicity studies.

1.7. SPECIAL CASES IN SAFETY ASSESSMENT

It may seem that the course of preclinical safety assessment (and of other aspects of development) of a pharmaceutical is a relatively linear and well-marked route, within some limits. This is generally the case, but not always. There are a number of special cases where the pattern and phasing of development (and of what is required for safety assessment) do not fit the usual pattern. Four of these cases are

1. When the drug is intended to treat a life-threatening disease, such as acquired immunodeficiency syndrome (AIDS).
2. When the drug is actually a combination of two previously existing drug entities.
3. When the drug actually consists of two or more isomers.
4. When the drug is a peptide produced by a biotechnology process.

Drugs intended to treat a life-threatening disease for which there is no effective treatment are generally evaluated against less rigorous standards of safety when making decisions about advancing them into and through clinical testing. This acceptance of increased risk (moderated by the fact that the individuals involved will die if not treated at all) is balanced against the potential benefit. These changes in standards usually mean that the phasing of testing is shifted: animal safety studies may be done in parallel or (in the case of chronic and carcinogenicity studies) after clinical trials and commercialization. But the same work must still be performed eventually.

Combination drugs, at least in terms of safety studies up to carcinogenicity studies, are considered by regulatory agencies as new drug entities and must be so evaluated. The accordingly required safety tests must be performed on a mixture with the same ratio of components as is to be a product. Any significant change in ratios of active components means one is again evaluating, in regulatory eyes, a new drug entity.

Now that it is possible to produce drugs that have multiple isomers in the form of single isomers (as opposed to racemic mixtures), for good historical reasons, regulatory agencies are requiring at least some data to support any decision to develop the mixture as opposed to a single isomer. One must, at a minimum, establish that the isomers are of generally equivalent therapeutic activity, and, if there is therapeutic equivalence, that any undesirable biological activity is not present to a greater degree in one isomer or another.

1.8. SUMMARY

It is the belief of this author that the entire safety assessment process that supports pharmaceutical research and development is a multistage process of which no single element is overwhelmingly complex. These elements must be coordinated and their timing and employment carefully considered on a repeated basis. Focus on the objectives of the process, including a clear definition of the questions being addressed by each study, is essential, as is the full integration of the technical talents of each of the many disciplines involved. A firm understanding of the planned clinical development of the drug is essential. To stay competitive requires that new technologies be identified and incorporated effectively into safety assessment programs as they become available. It is hoped that this volume will provide the essential knowledge of the key elements to allow these goals to be realized.

REFERENCES

- Alder, S. and Zbinden, G. (1988). *National and International Drug Safety Guidelines*. MTC Verlag, Zollikon, Switzerland.
- Arnst, C. (1998). Danger: Read the label. *Business Week*, May 11, p. 100.

- Beyer, K. (1978). Discovery, Development and Delivery of New Drugs. *Monographs in Pharmacology and Physiology*, No. 12. Spectrum, New York.
- Bryostowski, M. (2001). On the horizon. *R&D Directions*, May 2001, pp. 30–44.
- FDA (Food and Drug Administration). (1987a). *Good Laboratory Practice Regulations: Final Rule*. 21 CFR Part 58, *Federal Register*, September 4, 1987.
- FDA (Food and Drug Administration). (1987b). New drug, antibiotic, and biologic drug produce regulations. 21 CFR Parts 312, 314, 511, and 514, *Federal Register*, 52(53) 8798–8857.
- Fishlock, D. (1990, April 24). Survival of the fittest drugs. *Financial Times*, pp. 16–17.
- French, S. (1982). *Sequencing and Scheduling*. Halsted Press, New York.
- Gad, S.C. (1988a). An approach to the design and analysis of screening studies in toxicology, *J. Am. Coll. Toxicol.* 7(2): 127–138.
- Gad, S.C. (1989a). Principles of screening in toxicology: with special emphasis on applications to neurotoxicology. *J. Am. Coll. Toxicol.* 8(1): 21–27.
- Gad, S.C. (1989b). A tier testing strategy incorporating in vitro testing methods for pharmaceutical safety assessment. *Humane Innovations and Alternatives in Animal Experimentation* 3: 75–79.
- Gad, S.C. (1996). Preclinical toxicity testing in the development of new therapeutic agents, *Scand. J. Lab Anim. Sci.* 23: 299–314.
- Gad, S.C. (2000). *Product Safety Evaluation Handbook*, 2nd ed. Marcel Dekker, New York.
- Gad, S.C. (2001). *In Vitro Toxicology*, 2nd ed. Taylor and Francis, Philadelphia, PA.
- Guarino, R.A. (1987). *New Drug Approval Process*. Marcel Dekker, New York.
- Hamner, C.E. (1982). *Drug Development*. CRC Press, Boca Raton, FL, pp. 53–80.
- Kliem, R.L. (1986). *The Secrets of Successful Project Management*. Wiley, New York.
- Knutson, J.R. (1980). *How to Be a Successful Project Manager*. American Management Associations, New York.
- Lazarou, J., Parmeranz, B.H. and Corey, P.N. (1998). Incidence of Adverse Drug Reactions in Hospitalized Patients. *JAMA*, 279: 1200–1209.
- Leber, P. (1987). FDA: The federal regulations of drug development. In: *Psychopharmacology: The Third Generation of Progress* (Meltzer, H.Y., ed.). Raven Press, New York, pp. 1675–1683.
- Lesko, L.J., Rowland, M., Peck, C.C., and Blaschke, T.F. (2000). Optimizing the science of drug development: opportunities for better candidate selection and accelerated evaluation in humans, *Pharmaceutical Research* 17: 1335–1344.
- Levy, A.E., Simon, R.C., Beerman, T.H., and Fold, R.M. (1977). Scheduling of toxicology protocol studies. *Comput. Biomed. Res.* 10: 139–151.
- Littlehales, C. (1999). The price of a pill. *Modern Drug Discovery*, Jan/Feb 1999, pp. 21–30.
- Matoren, G.M. (1984). *The Clinical Process in the Pharmaceutical Industry*. Marcel Dekker, New York, pp. 273–284.
- Meyer, D.S. (1989). Safety evaluation of new drugs. In: *Modern Drug Research* (Martin, Y.C., Kutter, E., and Austel, V., Eds). Marcel Dekker, New York, pp. 355–399.
- Moore, T.J., Patsy, D. and Furnberg, J. (1998). Time to act on drug safety, *JAMA* 279: 1971–1976.

- O'Grady, J. and Linet, O.I. (1990). *Early Phase Drug Evaluation in Man*. CRC Press, Boca Raton, FL.
- Raeburn, P. (1999). Drug safety needs a second opinion. *Business Week*, Sept. 20, pp. 72–74.
- Redman, C. (1981). Screening compounds for clinically active drugs. In: *Statistics in the Pharmaceutical Industry*. (C.R. Buncher and J. Tsay, Eds.). Marcel Dekker, New York, pp. 19–42.
- Scheuplein, R.J., Shoal, S.E., and Brown, R.N. (1990). Role of pharmacokinetics in safety evaluation and regulatory considerations. *Ann. Rev. Pharmacol. Toxicol.* 30: 197–218.
- Smith, Charles G. (1992). *The Process of New Drug Discovery and Development*. CRC Press, Boca Raton, FL.
- Sneader, W. (1986). *Drug Development: From Laboratory to Clinic*. Wiley, New York.
- Spector, R., Park, G.D., Johnson, G.F., and Vessell, E.S. (1988). Therapeutic drug monitoring. *Clin. Pharmacol. Therapeu.* 43: 345–353.
- Spilker, B. (1994). *Multinational Drug Companies*, 2nd ed. Raven Press, New York.
- Traina, V.M. (1983). The role of toxicology in drug research and development. *Med. Res. Rev.* 3: 43–72.
- Tukey, J.W. (1977). *Exploratory Data Analysis*. Addison-Wesley, Reading, MA.
- Zbinden, G. (1992). *The Source of the River Po*. Haag and Herschen, Frankfurt, Germany.
- Zbinden, G., Elsner, J., and Boelsterli, U.A. (1984). Toxicological screening. *Reg. Toxicol. Pharmacol.* 4: 275–286.