3
Dose-Finding Studies in Phase I and Estimation of Maximally Tolerated Dose

MARLENE MODI

3.1 Introduction

Historically, drugs have been marketed at excessive doses (i.e., doses well onto the plateau of the efficacy dose–response relationship) with some patients experiencing adverse events (AEs) unnecessarily (Herxheimer, 1991; ICH-E4, 1994). Over the last 5 years, a greater effort has been made to ensure that the best benefit to risk assessment is obtained for each new drug (Andrews and Dombeck, 2004; Bush et al., 2005). The benefit to risk assessment of marketed drugs has been improved, in some cases, by postmarketing label changes, which aim to optimize the dosage regimen for the indicated populations (Cross et al., 2002). These postmarketing changes in the label may reflect the quality of drug development, regulatory review and postmarketing surveillance.

Information obtained in early clinical development about the average dose–response relationship in the intended patient population for a drug's desirable and undesirable effects is extraordinarily valuable, in that it lays the foundation for future dose–response studies (ICH-E4, 1994). Greater emphasis is being placed on the integration of information and ensuring effective decision-making during drug development. The pharmaceutical industrial sponsor of a compound is encouraged to discuss with health authorities as early as possible the type and number of clinical pharmacology studies that are needed to support labeling and approval. Also, the sponsor reviews with health authorities the use of preclinical and early clinical exposure–response information to guide the design of future dose–response, pharmacokinetic–pharmacodynamic (PK–PD) and clinical efficacy studies (FDA, 2003a,b). In this atmosphere of vigilance and information management, the selection of dose is considered a critical element of the benefit to risk assessment.

3.2 Basic Concepts

Initially, the development of a new chemical entity is influenced by its anticipated pharmacological actions in patients as suggested by its effects in animal models as well as its toxicology and PK profile in animals (Lesko et al., 2000; Peck

et al., 1994). The severity of the disease state and the availability of effective and safe alternative treatments are also key factors in formulating a development plan. Each new chemical entity is evaluated against key parameters of a target profile. These key elements of the target profile are essentially the components of a draft label, which are updated as the compound proceeds through development.

In terms of facilitating drug development and increasing the likelihood of marketing a drug successfully, the inherent properties of an ideal drug are often contrasted with those of the new chemical entity. An ideal drug is effective in controlling or reversing the pathophysiology of the clinical condition for which the drug is intended. It does not adversely affect other disease processes or result in adverse interactions with other drugs. It can be administered over a broad range of doses with minimal toxicity. The ideal drug is uniformly metabolized or eliminated by other mechanisms in a predictable manner that is not altered by organ impairment and is not influenced by age, race or gender. Few, if any, drugs possess all of these characteristics.

Information collected during drug development accumulates with each new phase leading to an understanding of the drug's inherent properties that are consistently shown throughout all phases of development (Figure 3.1). A brief overview of these various phases of drug development is given in Chapter 1.

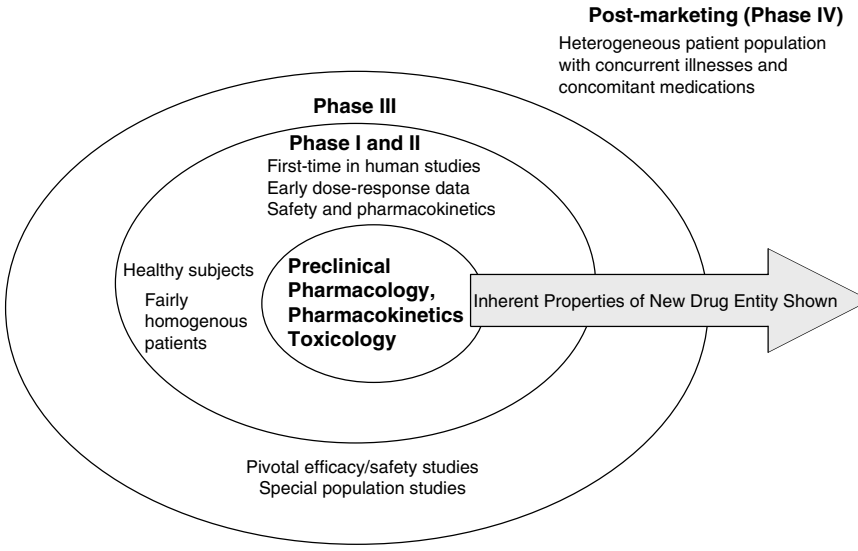

Figure 3.1. Information accumulates with each new phase of drug development and the drug's inherent properties become evident.

The information collected during drug development assists in determining the benefit-to-risk assessment for the heterogeneous population of patients that will be treated after the drug is approved for marketing. Adjustments are made to the

proposed dosing regimen throughout the drug development process upon review of information related to the drug's safety profile, efficacy, and PK.

This chapter focuses on Phase I studies that are designed to provide preliminary but essential information on safety, tolerability, PK and if possible the pharmacological actions of a compound. The term Phase I has two connotations: one refers to the earliest, first-time-in-humans (FIH) studies, while the other encompasses studies of PK, metabolism, drug interactions, special populations, and other clinical pharmacology trials (ICH-E8, 1997). Dose selection is a critical activity for Phase I studies to ensure that the data collected in these clinical trials are at doses to support the recommended therapeutic dose. The purpose of dose-finding studies in Phase I is to evaluate: the compound's mechanism of action in humans, the compound's metabolic actions and PK, AEs associated with increasing doses of the compound and to gain early evidence of the compound's effectiveness (Code of Federal Regulations, 2004). A well-designed and executed Phase I program permits the design of well-controlled, scientifically valid, Phase II studies.

Traditionally, Phase I studies have been conducted in 20 to 80 young, healthy, male subjects; however, this is not a regulatory requirement (ICH-E8, 1997). Women of non-childbearing potential and older healthy subjects are now being included in early studies especially if the drug is intended for these populations. Initial evaluations in patients may be preferable for drugs with a low safety margin and in certain life-threatening disease states (see Chapters 4 and 5). Given that healthy subjects derive no benefit from receiving a new chemical entity, risk minimization is a critical ethical concern for Phase I studies (FDA, 1997; Tishler and Bartholomae, 2002).

3.3 General Considerations for FIH Studies

Ascending dose studies are usually the first clinical trials in the drug development process. The upper limit of a compound's therapeutic window is partially characterized in Phase I as these ascending dose studies usually determine the dose-limiting AEs that prevent the titration to higher doses. The primary objectives of these ascending dose studies are to estimate a maximally tolerated dose (MTD), to characterize the most frequently occurring AEs, and to gain a general understanding of the drug's PK and PD profile. The MTD is defined as the dose level below that producing unacceptable but reversible toxicity and is considered the upper limit of patient tolerance. This chapter focuses on general design concerns of Phase I clinical trials. The reader is referred to Chapters 4 and 5 for discussions of issues related to the design of dose-finding trials in life-threatening diseases.

The same pharmacological mechanisms that account for a drug's efficacy can account for many of its toxic effects, as most drug-induced (or treatment-emergent) AEs are expected extensions of a drug's known pharmacological properties (Rawlins and Thompson, 1991). These AEs are usually dose-dependent and can

be predicted from animal studies. Thus, detailed knowledge of a drug's pharmacological actions assists in assessing for possible treatment-emergent AEs in the clinic. For example, both the AE of bradycardia (undesirable action) and the therapeutically desired reduction in blood pressure associated with the cardioselective beta-blocker, atenolol, are mediated through the drug's effect on beta-1 adrenergic receptors.

Treatment-emergent AEs may be unrelated to the drug's pharmacological action but may occur at higher doses or systemic exposures or upon chronic exposure to the drug. These types of AEs include withdrawal reactions, delayed reactions, failure of therapy and pharmacogenetic reactions (Edwards and Aronson, 2000). Unlike most treatment-emergent AEs, allergic drug reactions are unpredictable (Gruchalla, 2003). Some drugs (antimicrobial drugs, anticonvulsants, chemotherapeutic agents, heparin, insulin, protamine, and biologic response modifiers) are more likely to elicit clinically relevant immune responses.

Generally, ascending dose studies enroll too few subjects to observe treatment-emergent AEs that occur at a low to modest frequency. One way to visualize that only the most frequently occurring or common AEs are likely to be detected in FIH studies is to apply Hanley's Rule of Three (Hanley and Lippman-Hand, 1983). In order to ensure that one captures at least one occurrence of an AE happening at a frequency of 1:10 or greater at a 95% confidence level, the appropriate size of the safety database would be at least 30 subjects. Thus, given the small sample size of each dose group in the FIH study, it is common for these ascending dose studies to overestimate the MTD as the less frequently occurring treatment-emergent AEs and dose-limiting toxicities may not be detected (Buoen et al., 2003; Natarajan and O'Quigley, 2003).

3.3.1 Study Designs

A single dose is usually tested first, followed by multiple ascending dose studies; however, the study design is influenced by the type of compound. Study designs may be open-label, baseline-controlled or may use randomization and blinding. The most common study design used for these early studies is the parallel group, placebo-controlled, randomized, double-blind ascending dose study (Figure 3.2). Each group is typically made up of three to six subjects who receive single or multiple doses of the compound and one to two subjects who receive placebo. Safety and tolerability at the very least (in some cases PK and PD endpoints also) are evaluated before the next ascending dose group receives treatment.

Tolerability is an aspect of safety. It is a term used to indicate how well a patient is able to endure treatment such that AEs do not result in the discontinuation of treatment. A comparator drug, a marketed drug in the same class, can be included in the FIH study to evaluate the differences in tolerability between the two compounds if the comparator drug has a significant frequency of well-characterized AEs. The new chemical entity may possess a better tolerability profile than the comparator drug leading to a greater proportion of treated patients that successfully receive the full course of treatment.

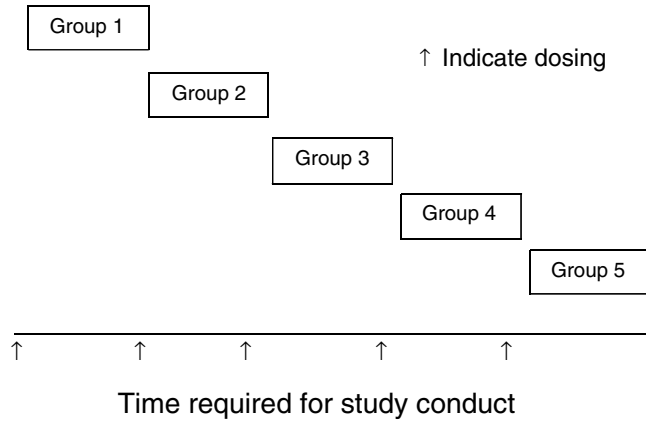

Figure 3.2. Parallel-group, placebo-controlled, randomized, double-blind ascending dose Phase I study design.

Given this early stage of drug development, not all subjects from the same dose group or cohort are dosed on the same day. This practice of spacing the dosing of subjects in a given dose group minimizes the number of subjects who are exposed to a given escalating dose of the drug and who are potentially at risk for a dose-limiting or irreversible toxicity of the drug. Should dose-limiting toxicities occur in the first few subjects of a dose group, dose escalation can be stopped without exposing all the subjects in a dose group.

Stopping rules for dose escalation need to be clearly described in the protocol. These may include reaching dose-limiting toxicities that define the MTD or seeing more frequent AEs than anticipated, that may influence the compliance of chronic administration (e.g., diarrhea or nausea). Stopping rules can also include clauses for evidence of unexpected or unique PK properties of the compound (e.g., dose- or time-dependent changes in clearance or volume of distribution, saturable absorption, presence of multiple active metabolites).

The stopping rules may be tailored for locally acting drugs or compounds with minimal toxicity. For these types of compounds, dose escalation may stop when the maximal feasible dose is reached. The maximal feasible dose is lower than the MTD, which cannot be estimated because it is not possible to administer high enough doses to reach the MTD. For some drugs where a good understanding of the pharmacological action of the drug exists in relation to the pathophysiology of the disease and efficacy of the drug, dose escalation may continue until the maximal pharmacological effect is reached in the absence of toxicity.

In general, only an average response for each dose group with respect to characterizing desirable or undesirable PD effects is obtained in the parallel-group, placebo-controlled, randomized, double-blind ascending dose study design. Although not easily appreciated, individual dose–response relationships may differ significantly from the population average relationship (see Chapter 1, Figure 1.1).

Table 3.1. Crossover, placebo-controlled, randomized, blinded study design

Group[a]	Treatment period 1	Treatment period 2	Treatment period 3
1	Placebo	Medium dose	High dose
2	Low dose	Placebo	High dose
3	Low dose	Medium dose	Placebo

[a] $N > 5$ per group

Another basic Phase I design is a crossover, placebo-controlled, randomized, blinded study (Table 3.1). In this design, a subject receives two dose levels and placebo in a randomized fashion. Like the parallel group design above, safety, tolerability, PK and PD data are evaluated before proceeding to the next treatment period. Stopping rules are clear and the study may be stopped or the doses modified based on information from the preceding treatment period. An individual subject's response is assessed at more than one dose level and before or after placebo treatment. There is a better understanding of an individual subject's contribution to the average dose response (FDA, 2003a).

The washout period between treatment periods in a crossover design is critical to ensure that there are no carryover effects from one period to another. This study design is inappropriate for drugs with long half-lives, for drugs with late toxicity, and if sensitization or tolerance develops. This study design is generally not used for FIH studies due to the general lack of information needed to rule out late toxicity, sensitization, tolerance, or to select an appropriate washout period. Sensitization is a phenomenon whereby the effects of a drug are augmented. Although it might sound counterintuitive, the same drug can evoke both tolerance and sensitization. Behavioral sensitization is a well-documented effect of repeated exposure to drugs such as amphetamine and cocaine (Pierce and Kalivas, 1997). Unlike transient drug effects, such as tolerance and withdrawal, behavioral sensitization can last as long as a year after the last drug administration in rats. The persistence of these effects implicates mechanisms distinct from those responsible for more transient drug effects.

Thus, for drugs with reversible desirable or undesirable actions, the crossover study design may provide a better understanding of the dose-concentration–response relationship than the parallel-group design as individuals receive two dose levels. In cases where it is unclear if the crossover design will be appropriate for a new chemical entity, a follow-up study to the traditional parallel-group FIH study may employ this design to better characterize individual dose- or exposure-response relationships.

3.3.2 Population

The description of the study population should identify important inclusion and exclusion criteria, demographic characteristics, baseline values of any clinically relevant variables that would be needed to understand the treatment effect related

to safety, tolerability or PD. Other characteristics of the population that have implications to the extent that results can be generalized need to be clearly described (Friedman et al., 1998). Inclusion and exclusion criteria are defined according to population studied (i.e., healthy subjects or patients).

Phase I studies often include healthy subjects between 18 and 65 years old, and groups are balanced for sex and racial distribution. General exclusionary criteria are written to prevent the enrollment of subjects that are not in good health (e.g., those with evidence of underlying diseases, abnormalities, or organ impairment). Subjects are excluded if they have participated in a study with another investigational agent in the recent past or have known allergies to any of the components in the formulation of the new chemical entity or to any of the related class of compounds. Specific exclusionary criteria that are related to safety concerns may vary with the compound being studied. These specific exclusionary criteria are likely to arise due to the compound's mechanism of action (e.g., subjects with flu-like symptoms for an interferon-like drug are excluded as endogenous levels of interferon are elevated during the flu). Exclusionary criteria may also be related to preclinical toxicology findings.

There are times, however, when initial studies are best performed in patients. Often patients present with a different tolerability profile than healthy subjects (e.g., antipsychotic drugs are tolerated at significantly higher doses in patients). In some cases, the AE profile can only be studied in patients. Typically, this occurs when a drug is suspected or known to be unavoidably toxic such as those used in oncology or other life-threatening diseases. The target patient population should be considered for FIH studies when there is evidence from toxicology studies of irreversible, severe effects (e.g., cytotoxicity) or damage to an organ system, a steep toxicity dose–response curve, or the effects are not easily monitored.

Drugs for the treatment of diseases that affect the elderly are tested early in elderly subjects. Similarly, drugs intended for the treatment of diseases that typically affect women need to be tested in female subjects. In addition, the pharmacodynamic effects of the drug may be measurable only in patients (e.g., anti-hypertensive medications such as nifedipine have little or no effect on blood pressure in normotensive subjects or the glucose-lowering effect of a drug is best assessed in a diabetic patient).

The most salient issue with the administration of protein drugs is that they may induce antibody formation. Antibodies could cross react with the naturally occurring protein, conceivably neutralizing desired physiological effects in healthy subjects. This is another factor to consider when including healthy subjects versus patients.

In general, if patients are required in Phase I studies for drugs to be used in non-life-threatening diseases, patients with comorbid conditions who are receiving concomitant therapies other than for the disease under study are excluded. Phase I studies for drugs to be used in life-threatening diseases, on the other hand, may include patients who have not responded to previously administered marketed or investigational treatments. These patients are ill, may have other underlying conditions or diseases and a shortened life expectancy.

A large pool of healthy subjects willing to participate allows the rapid enrollment and completion of studies. Healthy subjects are in a normal, relatively low-risk state of health. Studies in healthy subjects offer important advantages in that they generally have a greater physiological reserve than patients do. If an AE should occur, a healthy subject is more likely to recover without suffering long-term negative consequences. Also, healthy subjects are better able to provide more frequent measures of PD endpoints and give a greater number of blood samples for PK. The drawback to enrolling healthy subjects is that pathophysiological mechanisms of the targeted disease state cannot be observed and can only rarely be accurately simulated.

3.4 Dose Selection

The most important variable in FIH studies is dose. The choice of the starting dose, dose increment for subsequent doses, and the maximal dose to be investigated are common issues that need to be addressed in the study design. Selecting a starting dose and choosing the next dose levels are challenging. An overly conservative approach may lead to an endless study, whereas a too rapid escalation can lead to unacceptable toxicity. Although not always obvious, the maximal dose considered for testing should be stated in the protocol and the rationale for the upper range of doses selected should be clearly described. It is understandable that this maximal dose may never be reached.

3.4.1 Estimating the Starting Dose in Phase I

A strategy has been proposed to determine the highest recommended starting dose of new therapeutics in adult healthy volunteers (FDA, 2002). The draft guidance presents a fairly simple method of estimating the starting dose. The maximum recommended starting dose (MRSD) in adult healthy subjects is to be derived from the no-observed adverse effect levels (NOAELs) in toxicology studies of the most appropriate species, the NOAELs converted to human equivalent doses (HED), and a safety factor is then applied. The method assumes that NOAELs and MTDs scale reasonably well across species and that the conversion to HED is reasonably accurate after normalizing dose by a body surface area (BSA) conversion factor. Another major assumption is that the determination of a NOAEL is unambiguous.

The draft guidance method for estimating a starting dose in adult healthy subjects relies solely on dose and does not employ systemic exposure data directly (Figure 3.3). While more quantitative modeling approaches are presented in other guidelines (FDA, 2003a), the draft guidance on estimating the starting dose does not recommend these approaches. However, all of the relevant preclinical data, including information on the pharmacologically active dose, the compound's full toxicology profile, and the compound's PK (absorption, distribution, metabolism, and excretion) is likely to be considered when determining the MRSD.

Toxicology studies generate basically three types of findings that can be used to determine the NOAEL: (1) overt toxicity (clinical signs, macro and microscopic

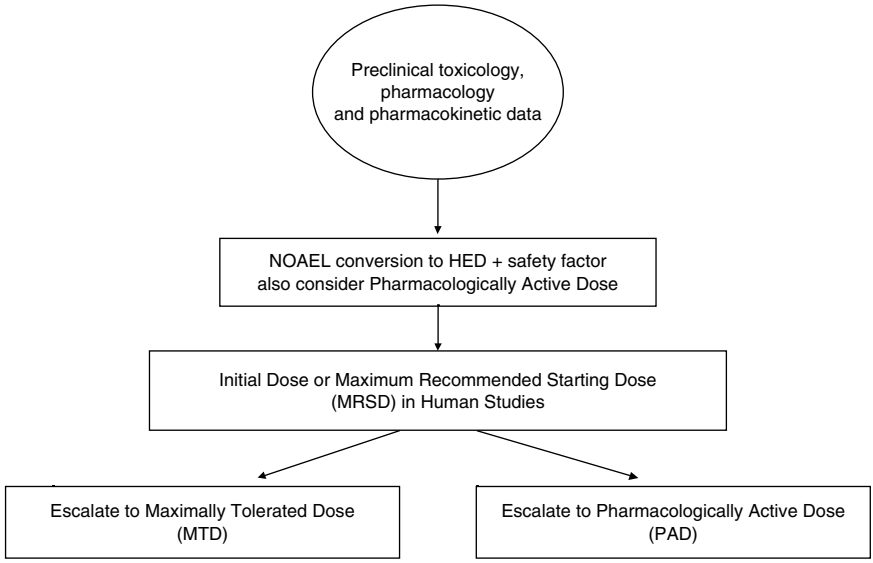

Figure 3.3. Overview of dose selection for FIH studies.

lesions), (2) surrogate markers of toxicity (serum liver enzyme levels), and (3) exaggerated pharmacodynamic effects (FDA, 2002; Sellers and du Souich, 2004). A recent review of current practices has revealed a lack of consistency in definition and application of frequently used terms such as no observed effect level (NOEL), NOAEL, adverse effect, biologically significant effect, or toxicologically significant effect (Lewis et al., 2002). Moreover, in review of current practices, no coherent criteria were found that were used to guide consistent interpretation of toxicity studies, including the recognition and differentiation between adverse effects and effects that are not considered adverse. As the interpretation of a compound's toxicology findings is the foundation of hazard and risk assessment, there is a need for consistent interpretation of toxicity (Lewis et al., 2002).

Toxicity should be avoided at the initial dose for the FIH study, but that does not necessarily mean that the starting dose will not possess any pharmacological activity. The pharmacologically active dose (PAD) should also be considered in that for a compound with limited toxicity, the PAD may be used to lower the estimate of the MRSD. However, in general, the HED is estimated from toxicology data in the most relevant species or alternatively, from the most sensitive species if the most relevant species is not known (FDA, 2002; Sellers and du Souich, 2004). Several factors could influence the choice of the most appropriate species including: (1) species differences in the compound's PK, (2) evidence indicating that a given species is predictive of human toxicity, and (3) limited cross-species pharmacological reactivity of the compound. This later point is especially important for biologic therapeutics in that many human proteins only bind to human or

nonhuman primate targets, and thus species other than nonhuman primates would not be appropriate for estimation of the HED (ICH-S6, 1997).

The draft guidance advocates that the NOAEL for systemically administered compounds can be accurately extrapolated to other species and humans when doses are normalized to BSA (mg/m^2) (FDA, 2002). The work of Freireich et al. (1966) and Schein et al. (1970) using approximately 33 anti-neoplastic drugs provide the basis for this assumption. For these limited number of anti-neoplastic drugs, doses lethal to ten percent of rodents (LD_{10}) and MTDs in nonrodents both correlated with the human MTD when the doses were normalized to the same administration schedule and expressed in terms of BSA (i.e., mg/m^2).

Body surface area was introduced into medical oncology practice in order to derive at a safe starting dose for Phase I studies of anticancer drugs from preclinical animal toxicology data (Sawyer and Ratain, 2001). While cardiac output does correlate with BSA, the relationship between BSA and other physiologic measures relevant for drug metabolism and disposition and thus systemic exposure, such as renal and hepatic function, is weak or nonexistent (Sawyer and Ratain, 2001, Boxenbaum and Dilea, 1995, Mahmood and Balian, 1999). An analysis of the impact of allometric exponent (0.67 vs. 0.75) on the conversion of an animal dose to the HED using Eq. (3.1) is presented in Appendix A of the draft guidance (FDA, 2002).

The approach recommended in the draft guidance to convert an animal NOAEL to an HED is by using the following equation:

$$\text{HED} = \text{animal NOAEL} \times (W_{\text{animal}}/W_{\text{human}})(1-b) \tag{3.1}$$

where W is the weight in kg, b (equal to 0.67) is a correction factor used to convert mg/kg to mg/m^2 and the interspecies scaling factor is $(W_{\text{animal}}/W_{\text{human}})(1-b)$.

The derivation of the interspecies scaling factor in Eq. (3.1) is presented in Appendix C of the same draft guidance. Inherent in the BSA normalization is the use of the factor, $W^{0.67}$. Other limited data besides that of Freireich et al. (1966) and Schein et al. (1970) suggest that the most accurate allometric exponent for normalizing MTDs of antineoplastic agents for interspecies extrapolation is $b = 0.75$ (FDA, 2002). Based on the analysis presented in Appendix A of the draft guidance and the premise that correcting for BSA increases clinical trial safety by resulting in a more conservative starting dose estimate, the guidance recommends that the approach of converting NOAEL doses to an HED based on BSA correction factors (i.e., $W^{0.67}$) be used for selecting starting doses of initial studies in adult healthy volunteers. Deviations from the surface area approach should be justified, and it is wise to calculate the initial dose to be used in adult healthy volunteer studies by multiple approaches (Reigner and Blesch, 2002).

Once the HED has been determined, a safety factor is applied to provide a margin of safety that allows for variability in extrapolating from animal toxicity studies to humans (FDA, 2002; Sellers and du Souich, 2004). This variability can result from: (1) uncertainties due to enhanced sensitivity to pharmacological activity in humans versus animals, (2) difficulties in detecting certain toxicities in

animals, (3) differences in receptor densities or affinities, (4) unexpected toxicities and (5) interspecies differences in PK. In practice, the MRSD for the clinical trial is determined by dividing the HED by a default safety factor of 10.

In certain situations, the use of a safety factor greater than 10 is required. Criteria for using a safety factor greater than 10 include those related to toxicity such as: (1) a steep dose–response curve for important toxicities in the most relevant species or in multiple species, (2) severe toxicity or damage to an organ system in animals, (3) irreversible toxicity in animals, (4) nonmonitorable toxicity, (5) presence of significant toxicities without prodromal indicators and (6) nonpredictable and unexplained mortality. Other factors to consider include: (1) variable bioavailability between species, with poor bioavailability in the test species used to derive the HED, (2) large variability in doses or AUC levels eliciting a toxic effect, (3) questionable toxicology study design or conduct, such as few dose levels, wide dosing intervals, or large differences in responses between animals within dosing groups and (4) novel therapeutic targets. The safety factor should be increased when animal models with limited ability to evaluate the compound's toxicity are used. This may result because of very limited interspecies cross-reactivity or pronounced immunogenicity (e.g., protein drugs likely to be pharmacologically active only in nonhuman primates), or because the compound's effect is elicited by mechanisms that are not known to be conserved between animals and humans (FDA, 2002; Sellers and du Souich, 2004).

Safety factors of less than 10 may be appropriate under some conditions (FDA, 2002; Sellers and du Souich, 2004): (1) the compound belongs to a well-characterized class, has a similar metabolic profile and bioavailability, presents similar toxicity across all the species tested including humans, and it is administered by the same route, schedule, and duration of administration, (2) the toxicity elicited is easily monitored, reversible and predictable, and a moderate to shallow dose–response relationship with toxicities are consistent across the tested species, and (3) the NOAEL is estimated from toxicity studies of longer duration than required for the proposed clinical schedule in healthy subjects. The toxicology testing in these cases should be of the highest caliber in both conduct and design.

3.4.2 Dose Escalation

It is not always necessary to escalate to doses as high as the MTD in the FIH studies. The highest single dose tested can also be defined as the pharmacologically active dose (PAD) giving the maximal effect in the absence of toxicity (Figure 3.3). However, the estimation of the PAD from preclinical pharmacology studies may not be possible if animal models of the disease are not available or the understanding of the fundamental biochemical or physiological aspects of the mechanism of action of the drug is lacking. Target site and receptors may be absent or modified in animal models precluding the estimation of the PAD in animals. Treatment in animals does not always lead to sufficiently sustained drug concentrations at the site of action in order to extrapolate the PAD to humans. PK may differ between species. Also, it is common to perform studies in animal models of disease using

the intravenous or intraperitoneal route of administration which are unlikely to be the intended route of administration for patients. However, an estimation of pharmacologically active doses or targeted plasma concentrations is often helpful in guiding the dose escalation (Reigner and Blesch, 2002).

The choice of the dose escalation scheme is usually based on the type of toxicity and the steepness of the dose–response curve seen in toxicology and pharmacology experiments. Several classical methods for dose escalation have been described (Spilker, 1991): (1) starting dose (x) increased by an equal amount (x, 2x, 3x, etc.), (2) dose increased by equal percentage (e.g., by 100%), (3) modified Fibonacci (x, 2x, 3x, 5x, 7x, 9x, 12x, and 16x), and (4) a variant of the modified Fibonacci scheme where doses are increased by 100% until the first hint of toxicity followed by the modified Fibonacci scheme. Many of these methods have been traditionally used in Phase I studies in patients with cancer. A number of new study design proposals for anticancer agents address ethical concerns about treating excessive numbers of patients at subtherapeutic doses. These new study designs aim to increase the overall efficiency of the process while enhancing the precision of the recommended Phase II dose (see Chapters 4 and 5; Zhou, 2004).

Methods based on concentrations or PK guided dose escalation utilize PK parameters such as AUC or C_{max} from the preceding dose group to rationalize the dose increments for escalation (Vaidya and Vaidya, 1981; Graham and Workman, 1992; Reigner and Blesch, 2002). Doses are escalated to the MTD if appropriate, and AUC or a given PK parameter is monitored. In general, doses are escalated by doubling the dose until 40% of the AUC at the mouse LD_{10} is reached, and then conventional dose escalation begins. The underlying theme of this approach is that the AUC at the mouse LD_{10} is close to the MTD in humans although a different dose may be needed to achieve that AUC value in humans.

The PK–PD guided dose escalation can utilize target plasma concentrations established in animal models of disease and may provide a more rapid and safe completion of the FIH study as well as decrease the number of patients receiving a subtherapeutic dose. At each dose level, the PK and PD data are incorporated into an interactively updated PD model. Difficulties arise when the compound's PK differs substantially among species, dose-dependent or time-dependent changes in PK occur, or there is considerable inter- and intra-individual variability in PK or PD. In addition, it is unknown if maintaining these target plasma concentrations will ultimately lead to efficacy in the patient population. When using PK to escalate the dose, a maximally tolerated systemic exposure instead of MTD may be determined. This type of strategy can be seen as an application of the "concentration controlled clinical trial" design (Kraiczi et al., 2003).

Biomarkers can be defined as "physical signs or laboratory measurements that may be detected in association with a pathologic process and that may have putative diagnostic or prognostic utility". These can be measured objectively as indicators of biological or pathological processes or of the response to a therapeutic intervention (Rolan et al., 2003). Biomarkers can help guide dose escalation and may assist in understanding the dose–response relationship for the primary efficacy endpoint in Phases II or III (e.g., blood pressure and cholesterol reduction have

been linked to heart attack or stroke-related mortality and have attained the status of surrogate endpoints; Temple, 1999). However, the shape of the dose–response relationship generated with biomarkers may differ from that of the primary efficacy endpoint, as long-term effects may not readily translate from the acute effects on the biomarker. Biomarkers may also be used to characterize the relationship between dose and undesirable effects (e.g., incidence and severity of neutropenia seen with interferon-like drugs) to facilitate the estimation of MTD.

3.5 Assessments

3.5.1 Safety and Tolerability

The use of randomization, blinding and a concurrent placebo-controlled group reduces the bias in safety assessments during the FIH study. Prespecified safety definitions (e.g., definition of dose-limiting toxicities and MTD) and stopping rules for dose escalation also ensure that safety and tolerability data are collected in an objective manner. Many have proposed that FIH studies be open-label and without concurrent placebo controls. For objective measures that are less susceptible to bias by the subject or investigator (e.g., AUC values), this could be a consideration. Unfortunately, AE reporting is often subjective.

The underlying objectives of safety and tolerability assessments in single dose FIH studies are to monitor for early signals of toxicity and to characterize the common treatment-emergent AEs. Consideration should be given to AEs that are likely after chronic use of the drug to reduce compliance in the intended patient population. Safety issues may result from the extension of the drug's pharmacological effects or be unrelated to the drug's pharmacological actions in that the toxicity is unexpected and was not seen in preclinical studies.

For multiple ascending dose studies, subchronic treatment-emergent AEs are characterized. The effect of multiple dosing on accumulation of a drug's systemic exposure is evaluated. For both single and multiple ascending dose studies, appropriate follow-up is needed to detect late toxicity (e.g., hepatotoxicity with fialuridine and antiretroviral agents; Styrt and Freiman, 1995; Kontorinis and Dieterich, 2003). Compounds that affect hematology parameters (e.g., red blood cells) may produce late toxicity like anemia, which may not appear until there has been enough time for the red blood cell population to turn over. In general, the follow-up period should not be less than four to five times the terminal half-life of the drug (provided this covers a significant portion of the AUC) or 4 weeks.

Early studies usually carefully monitor organ functions after single or multiple ascending doses (e.g., cardiovascular and pulmonary vital signs and electrocardiograms, hepatic, renal, and hematological laboratory parameters, and clinical signs and symptoms of target organ toxicity that have been identified in preclinical toxicology or pharmacology studies). One of the objectives of FIH studies is to monitor for early signals of severe toxicity, and humans are considered to be possibly more sensitive to the toxicity of the compound than the species used in toxicology studies. The critical organ functions to monitor are those identified

in toxicology studies as being affected by the compound. However, a number of compounds exhibit safety concerns that were not initially detected in toxicology studies (e.g., hepatotoxicity). It is prudent to ensure that adequate safety assessments are included in the protocol to characterize the expected AEs and to identify early signals of severe or unexpected toxicity.

3.5.2 Pharmacokinetics

Most drugs have inter- and intra-subject variability in PK parameters of at least 20% to as much as several fold. Overlap in systemic exposure across various dose levels occurs when the variability in PK parameters is large (e.g., > fourfold in clearance) or if the increment in each dose escalation is low. If significant treatment–emergent events occur during a given dose escalation, it may be reasonable to repeat the same dose in the next group or proceed with a minimal dose increase.

A major objective underlying PK assessments is to detect an unexpected or unusual PK profile that could lead to severe toxicity. While important to detect, dose-dependent and time-dependent changes PK may be masked by the small sample sizes and considerable inter-subject variability in PK parameters. However, the FIH study is often the best study to show that a compound exhibits dose-independent and time-independent PK (i.e., clearance and volume of distribution is constant across doses and over time), as there are generally several dose levels tested and the PK sampling is more extensive in early studies. Further study may be required to characterize the mechanism of a compound's dose-dependent or time-dependent PK and to identify its source. PK data should be obtained rapidly from all dose groups in the single and multiple ascending dose studies if dose-dependent or time-dependent PK is suspected. If a drug exhibits dose-dependent PK such that small changes in dose have a significant effect on AUC, the drug's pharmacological effect may be increased disproportionately as well as its duration of action with increasing doses.

In multiple ascending dose studies, subjects are usually treated for several days beyond that needed to achieve steady state. PK data from the single dose FIH study is used to estimate the dosing frequency for the multiple dose study. These data are used to predict accumulation and the time required to reach steady-state plasma concentrations. In a broad qualitative sense, the appearance of metabolites are characterized in humans and compared with animal data. As drug development progresses, the PK profile of a compound is continuously refined such that predictions can be made about routes of elimination and potential drug interactions, and special populations can be identified.

3.5.3 Pharmacodynamics

It is important to determine if the drug's desired pharmacological effects occur at dose levels that humans can tolerate. Without this information, the estimated MTD cannot be put into context of a therapeutic window. For drugs with reversible pharmacological action that is readily quantifiable, PD becomes an important

assessment in FIH studies. Desirable or undesirable pharmacodynamic effects may only be measurable in patients (e.g., anti-hypertensive agents). Often with antagonists, pharmacological activity can only be demonstrated with a provocative challenge. For example in exercise-induced asthma, a patient undergoes an exercise challenge to assess the pharmacological activity of a leukotriene antagonist as the targeted leukotriene pathway responsible for bronchoconstriction is operative only in the disease state (Adelroth et al., 1997).

Pharmacological effects, if these are related to exposure and are predictable from animal data, should be monitored by carefully observing subjects. If no exaggerated pharmacological effects are seen in healthy subjects and patients in early Phases I and II studies, then these exaggerated effects are unlikely to be seen in Phase III. However, it is possible that a drug could have an effect that might become apparent in patients, but was not seen in healthy subjects. The healthy subject's counter-regulatory system may be able to compensate whereas that of the patient may not. For example, counter-regulatory mechanisms induced by hypothermia include shivering, which can induce a fourfold increase in heat production, but at the expense of a 40 to 100% increase in oxygen consumption. Patients with coronary artery disease often have worse outcomes in hypothermia. However, for certain treatment–emergent events counter-regulatory mechanisms may be ineffective even in the healthy subject.

Major sources of variability in a patient's response to a given treatment are derived from PK, PD or the disease state itself provided that the patient is compliant. The drug may have a variable effect on the disease over time. For drugs having greater variability in PK than PD parameters, plasma concentration data may be better able than dose to predict the magnitude and duration of PD effects (FDA, 2003a). On the other hand, if PD variability is greater than PK variability, plasma concentration data may not predict the PD effect well. Sources of PK variability could include demographic factors (age, gender, and race), other diseases (renal or hepatic), diet, concomitant medications, and disease characteristics. Thus, assessing variability and identifying the sources of variability allows for a better understanding of the individual dose–response relationship for PD or efficacy endpoints.

Understanding a drug's pharmacological response is challenging due to the multifaceted nature of this endeavor. As a practical matter, it is easier to demonstrate a dose–response relationship for a PD effect that can be measured as a continuous or categorical variable, if the effect is obtained relatively rapidly after dosing and dissipates rapidly after therapy is stopped (e.g., blood pressure, analgesia, or bronchodilation) (FDA, 2003a). For drugs acting on the central nervous system, measuring the intensity of the pharmacological response is not always possible and several of the frequently used psychomotor performance tests suffer from limitations related to learning and practice effects (Di Bari et al., 2002). For this reason, it may not be possible to apply these tests repeatedly within the same subject.

For drugs used in the treatment of depression, anxiety and pain, rating scales are often used. The responses to rating scales may be subjective and variables such

as motivation or fatigue can influence the results (Demyttenaere et al., 2005). The assessment of visual acuity in age-related macular degeneration requires the use of sham or placebo-control to minimize bias as the patient may try harder to see and lean forward during visual acuity assessments if he believes he is benefiting from treatment (Gragoudas et al., 2004). Knowledge of the disease state in relation to the selection of PD endpoints and examples of successful efforts with other drugs for the same indication or having the same mechanism of action provides a greater certainty that these data will be collected and analyzed appropriately and be ultimately usable.

PD endpoints which can be readily measured and exhibit the ideal characteristics (continuity, repeatability or the ability to obtain multiple measurements over time, reproducibility, sensitivity, and objectivity) often have an unclear relationship to the primary efficacy endpoint (Lesko et al., 2000). Sometimes the efficacy endpoint is delayed, persistent, or irreversible (e.g., stroke prevention, arthritis treatments with late onset response, survival in cancer, treatment of depression). Thus, it is not inconceivable that the shape of the dose or exposure or concentration–response relationship for the PD endpoint differs from that of the efficacy dose or concentration–response relationship (Figure 3.4).

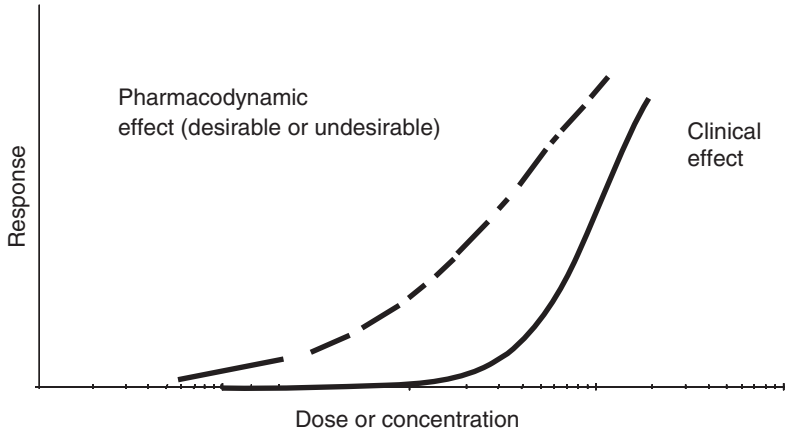

Figure 3.4. PD effect vs. clinical efficacy dose– or exposure–response relationships.

Clinical PK/PD data arise from complex and dynamic systems. Data from early studies are limited to single and short-term multiple dosing from a small number of individuals, and these data are unlikely to represent the full breadth of the intended patient population. Nonetheless, these data are invaluable in establishing exposure–response relationships that are further characterized in Phases II and III to provide a basis for dosage adjustment in subpopulations of interest and a rationale for the intended clinical dose (see Chapter 6). Various approaches have been used to model PK–PD or PD versus dose data (e.g., effect compartment, lag-time, PK–PD link, physiological feedback, indirect response models). These models in their most general form can be seen as relating PD effects to dose or

exposure (see Chapter 14 for E_{max} model) to more extensive modeling efforts with successive links from dose to exposure to PD or efficacy endpoints (see Chapter 6).

3.6 Dose Selection for Phase II

In addition to examining dose or concentration response information from studies specifically designed to provide it, the entire database should be examined for possible desirable or undesirable PD effects that could be related to dose or concentration. If possible have an estimate from Phase I studies of the smallest dose that could provide any benefit. If quantifiable, select reasonable PD parameters to measure in Phase II in order to gain further information on the variability in PD and an early understanding of the influence of disease state on PD effects in Phase II. In addition, information about the relationship between PD and the proposed efficacy endpoint can be gathered in Phase II if not already known. The careful selection of PD endpoints or biomarkers are invaluable in understanding the dose or exposure response data as the development progresses from Phase I to II and reduces the likelihood of a failed Phase III study or a Phase III study where all doses rest on the plateau of the efficacy dose–response curve. Information on the duration of a PD effect along with PK data obtained in Phase I studies provides a basis for dosage interval or frequency. Identification of the common AEs and those associated with dose is extremely helpful in planning Phase II studies. Setting the upper limit of the dose range that will be explored by estimating the MTD in Phase I guides the selection of doses. While Phase I studies are generally small in size and have many limitations with respect to the breadth of information that can be gathered, a well designed Phase I program is essential for formulating hypotheses on how the drug works and forms the basis for the design of scientifically valid Phase II dose-ranging studies.

References

Adelroth, E., Inman, M.D., Summers, E., Pace, D., Modi, M., and O'Byrne, P.M. 1997. Prolonged protection against exercise-induced bronchoconstriction by the leukotriene D4-receptor antagonist cinalukast. *Journal of Allergy and Clinical Immunology* 99: 210–215.

Andrews, E., and Dombeck, M. 2004. The role of scientific evidence of risks and benefits in determining risk management policies for medications. *Pharmacoepidemiology and Drug Safety* 13:599–608.

Boxenbaum, H., and DiLea, C. 1995. First-time-in-human dose selection: Allometric thoughts and perspectives. *Journal of Clinical Pharmacology* 35:957–966.

Buoen, C., Holm, S., and Thomsen, M.S. 2003. Evaluation of the cohort size in Phase I dose escalation trials based on laboratory data. *Journal of Clinical Pharmacology* 43: 470–476.

Bush, J.K., Dai, W.S., Dieck, G.S., Hostelley, L.S., and Hassall, T. 2005. The art and science of risk management: A US research-based industry perspective. *Drug Safety* 28:1–18.

Code of Federal Regulations. 2004. *21 CFR312.21—Investigational new drug application—Phases of an investigation*.

Cross, J., Lee, H., Westelinck, A. Nelson, J., Grudzinskas, C., and Peck, C. 2002. Postmarketing drug dosage changes of 499 FDA-approved new molecular entities, 1980–1999. *Pharmacoepidemiology and Drug Safety* 11:439–446.

Demyttenaere, K., De Fruyt, J., and Stahl, S.M. (2005) The many faces of fatigue in major depressive disorder. *The International Journal of Neuropsychopharmacology* 8, 93–105.

Di Bari, M., Pahor, M., Barnard, M., Gades, N., Graney, M., Franse, L.V., Penninx, B.W., Marchionni, N., and Applegate, W.B., 2002. Evaluation and correction for a 'training effect' in the cognitive assessment of older adults. *Neuroepidemiology* 21:87–92.

Edwards, I.R., and Aronson, J.K. 2000. Adverse drug reactions: Definitions, diagnosis, and management. *Lancet* 356:1255–1259.

Food and Drug Administration. 1997. *Guidance for Industry: General Considerations for the Clinical Evaluation of Drugs*.

Food and Drug Administration. 2002. *Guidance for Industry: Estimating the Safe Starting Dose in Clinical Trials for Therapeutics in Adult Healthy Volunteers*.

Food and Drug Administration. 2003a. *Guidance for Industry: Exposure-Response Relationships – Study Designs, Data Analysis, and Regulatory Applications*.

Food and Drug Administration. 2003b. *Concept Paper: End-of-phase-2A Meetings with Sponsors Regarding Exposure-Response of IND and NDA Products*.

Freireich, E.J., Gehan, E.A., Rall, D.P., Schmidt, L.H., and Skipper, H.E. 1966. Quantitative comparison of toxicity of anticancer agents in mouse, rat, hamster, dog, monkey and man. *Cancer Chemotherapy Reports* 50:219–244.

Friedman, L.M., Furberg, C.D., and DeMets, D.L. 1998. "Study population," in *Fundametals of Clinical Trials.* 3rd ed., New York: Springer-Verlag, pp. 30–40.

Gragoudas, E.S., Adamis, A.P., Cunningham, E.T.Jr., Feinsod, M., and Guyer, D.R. 2004. Pegaptanib for neovascular age-related macular degeneration. *The New England Journal of Medicine* 351:2805–2816.

Graham, M.A., and Workman, P. 1992. The impact of pharmacokinetically guided dose escalation strategies in Phase I clinical trials: Critical evaluation and recommendations for future studies. *Annals of Oncology* 3:339–347.

Gruchalla, R.S. 2003. Drug allergy. *Journal of Allergy and Clinical Immunology* 111(2 Suppl):S548–S559.

Hanley, J.A., and Lippman-Hand, A. 1983. If nothing goes wrong, is everything all right? Interpreting zero numerators. *JAMA* 249:1743–1745.

Herxheimer, A. 1991. How much drug in the tablet? *Lancet* 337:346–348.

ICH-E4 Harmonized Tripartite Guideline. 1994. *Dose-Response Information to Support Drug Registration*.

ICH-E8 Harmonized Tripartite Guideline. 1997. *General Considerations for Clinical Trials*.

ICH-S6 Harmonized Tripartite Guideline. 1997. *Preclinical Safety Evaluation of Biotechnology-Derived Pharmaceuticals*.

Kontorinis, N., and Dieterich, D. 2003. Hepatotoxicity of antiretroviral therapy. *AIDS Reviews* 5, 36–43.

Kraiczi, H., Jang, T, Ludden, T., and Peck, C.C. 2003. Randomized concentration-controlled trials: Motivations, use and limitations. *Clinical Pharmacology and Therapeutics* 74:203–214.

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. *Journal of Clinical Pharmacology* 40:803–814.

Lewis, R.W., Billington, R., Debryune, E., Gamer, A., Lang, B., and Carpanini, F. 2002. Recognition of adverse and nonadverse effects in toxicity studies. *Toxicologic Pathology* 30:66–74.

Mahmood, I., and Balian, J.D. 1999. The pharmacokinetic principles behind scaling from preclinical results to Phase I protocols. *Clinical Pharmacokinetics* 36, 1–11.

Natarajan, L., and O'Quigley, J. 2003. Interval estimates of the probability of toxicity at the maximum tolerated dose for small samples. *Statistics in Medicine* 22:1829–1836.

Peck, C.C., Barr, W.H., Benet, L.Z., Collins, J., Desjardins, R.E., Furst, D.E., Harter, J.G., Levy, G., Ludden, T., Rodman, J.H., et al. 1994. Opportunities for integration of pharmacokinetics, pharmacodynamics, and toxicokinetics in rational drug development. *Journal of Clinical Pharmacology* 34:111–119.

Pierce, R.C., and Kalivas, P.W. 1997. A circuitry model of the expression of behavioural sensitization to amphetamine-like psychostimulants. *Brain Research. Brain Research Reviews* 25:192–216.

Rawlins, M.D., and Thompson, J.W. 1991. "Mechanisms of adverse drug reactions," in *Textbook of adverse drug reactions* (D.M. Davies, editor), Oxford: Oxford University Press, pp. 18–45.

Reigner, B.G., and Blesch, K.S. 2002. Estimating the starting dose for entry into humans: Principles and practice. *European Journal of Clinical Pharmacology* 57:835–845.

Rolan, P., Atkinson, A.J.Jr., and Lesko, L.J. 2003. Use of biomarkers from drug discovery through clinical practice: Report of the Ninth European Federation of Pharmaceutical Sciences Conference on Optimizing Drug Development. *Clinical Pharmacology and Therapeutics* 73:284–291.

Sawyer, M., and Ratain, M.J. 2001. Body surface area as a determinant of pharmacokinetics and drug dosing. *Investigational New Drugs* 19:171–177.

Schein, P.S., Davis, R.D., Carter, S., Newman, J., Schein, D.R., and Rall D.P. 1970. The evaluation of anticancer drugs in dogs and monkeys for the prediction of qualitative toxicities in man. *Clinical Pharmacology and Therapeutics* 11:3–40.

Sellers, E.M., and du Souich, P. 2004. "Phase I Studies (Human Pharmacology)," in *The IUPHAR compendium of basic principles for pharmacological research in humans* (P. du Souich, M. Orme and S. Erill, editors), Irvine, CA: International Union of Basic and Clinical Pharmacology, pp. 40–49.

Spilker, B. 1991. *Guide to Clinical Trials*. New York: Raven Press.

Styrt, B., and Freiman, J.P. 1995. Hepatotoxicity of antiviral agents. *Gastroenterology Clinics of North America* 24:839–852.

Temple, R. 1999. Are surrogate markers adequate to assess cardiovascular disease drugs? *JAMA* 282:790–795.

Tishler, C.L., and Bartholomae, S. 2002. The recruitment of normal healthy volunteers: A review of the literature and the use of financial incentives. *Journal of Clinical Pharmacology* 42:365–375.

Vaidya, A.B., and Vaidya, R.A. 1981. Initial human trials with an investigational new drug (Phase I and II) : Planning an management. *Journal of Postgraduate Medicine* 27: 197–213.

Zhou, Y. 2004. Choice of designs and doses for early phase trials. *Fundamental and Clinical Pharmacology* 18:373–378.