

Part I

General Principles and Controversial Issues in Dose-Findings

1

Basic concepts in dose-finding

Sylvie Chevret

*Département de Biostatistique et Informatique Médicale, U717 Inserm,
Hôpital Saint-Louis, Paris, France*

1.1 Main concepts

Dose-finding trials aim at coming up with a safe and efficient drug administration in humans [1]. They are defined as early phase clinical experiments in which different doses of a new drug are evaluated to determine the optimal dose that elicits a certain response to be recommended for the treatment of patients with a given medical condition. They are to be distinguished from preclinical quantal bioassays, though statistical analysis of both experiments may be closely related [2–5], as detailed below. They should also be distinguished from dose-ranging experiments, though there is a frequent confusion in literature between dose-ranging and dose-finding [6]. Actually, we will consider dose-ranging as the (often random) comparison of two or more doses of a drug in terms of response, which is not within the scope of this book.

Besides this controversy, the definition of dose-finding studies always uses several distinct concepts of *dose*, *response* and *optimal dose*, which should be clearly defined. The *dose* is the amount of active substance given in a single administration or repeated over a given period, according to a certain administration schedule of equal or unequal single doses at equal or unequal intervals. Of note, it is sometimes important to include the patient's body weight (notably in children) in the expression of doses. The *response* is the patient's outcome of interest, defined either in terms of toxicity/tolerability or therapeutic points of view. The tolerability is commonly measured

by the appearance of unwanted signs and symptoms and changes in some clinical and laboratory findings, denoting some adverse reactions following drug administration. By contrast, the therapeutic response is measured by the change observed in one or more variables taken as indications of the intensity of the patient's condition, i.e. depending on the disease under study. According to the measured outcome, either toxicity or therapeutic response, two *optimal doses* are defined in dose-finding, namely the maximal tolerated dose (MTD) and the minimal effective dose (MED). The maximal tolerated dose is often defined as the dose that produces an 'acceptable' level of toxicity [7] or which, if exceeded, would put patients at 'unacceptable' risk for toxicity, or the dose producing a certain frequency of (medically unacceptable) reversible, dose-limiting toxicity (DLT) within the treated patient population. DLT includes host effects up to the point that is acceptable to the patient, based on several severity grading scales of adverse events, such as the common toxicity criteria developed by the National Cancer Institute of the United States. For instance, in oncology, DLT is usually defined as any nonhaematological grade III or grade IV toxicity (except alopecia, nausea, vomiting or fever, which can be rapidly controlled with appropriate measures), or an absolute neutrophil count $<500/\text{ml}$ for at least 7 days, or febrile neutropenia (absolute neutrophil count $<500/\text{ml}$ for at least 3 days and fever above 38.5°C for 24 hours), or thrombocytopenia grade IV. The minimal effective dose (MED) is the dose that elicits a prescribed lowest therapeutic response. Such definitions are particularly important by conditioning the definition of the endpoint in dose-finding experiments; they will be discussed in Section 1.3.3.2. The difference between these two doses defines the therapeutic area (window) of interest.

Dose-finding experiments are common for evaluating tolerability, while dose-finding studies of efficacy are unusual. Dose-finding experiments that focus on the evaluation of MTD are referred to as phase I clinical trials, whereas those focusing on the evaluation of MED are referred to as (early) phase II clinical trials. Most of the time, dose-finding experiments are restricted to phase I trials.

Traditionally, phase I trials are considered 'first in human' studies, following extensive preclinical testing. However, it is important to recognize that phase I studies are not limited to 'first in human' studies. Subsequent phase I trials often evaluate new schedules or combinations with established drugs. In addition, these secondary phase I studies may focus on a particular population that was excluded in prior studies, such as children.

The primary goal of phase I trials is to determine the appropriate dose for phase II evaluation, i.e. attempting to define a standardized treatment schedule to be safely applied to humans and worth being further investigated for efficiency. Phase I clinical trials are performed in many medical areas, but are particularly important in cancer, where they are the essential gateways to the development of new therapies. Otherwise, for most drugs, as long as the expected toxicity is mild and can be controlled without harm, dose-finding phase I trials involve the administration of usually subtherapeutic doses of a new agent to healthy adult volunteers, in specially dedicated clinical pharmacology units, to investigate the initial safety and tolerability, and also the pharmacokinetic drug profile and the pharmacodynamics [8]. By contrast, in

life-threatening diseases, such as cancer and AIDS, because of the toxicity that generally is observed in preclinical studies, phase I trials of new agents cannot be conducted in healthy volunteers. In these settings, participants in phase I trials are almost always patients with refractory disease or for whom there is no standard therapy, often at a very high risk of death, who consent to participate in the trial only as a last resort in seeking a cure.

Actually, most attention has been devoted to the design, conduct and analysis of phase I dose-finding experiments in cancer patients. Besides the main difference described above in enrolled subjects, there are many other differences in the design, the endpoints and the analysis of dose-finding phase I experiments conducted in healthy volunteers and those conducted in cancer patients. Therefore, dose-finding methods for phase I healthy volunteer studies will be treated separately (in Chapter 9).

By contrast, there are fewer differences in statistical methods between dose-finding experiments for cancer phase I studies and those for phase II studies than between dose-finding phase I studies in cancer and in healthy volunteers. Most of the methods presented below focus on dose-finding in phase I cancer studies, but would be extended easily to dose-finding phase II studies in noncancer populations.

1.2 Main issues from a pharmaceutical point of view

Drug development is a continuous process through which the knowledge of efficacy and safety of an experimental drug is gradually accumulated. For designing confirmatory (phase III) trials, it is essential to have sufficient knowledge of one or two doses of the drug that can be considered as optimal for safety and efficacy. A well-designed dose-finding study should provide valuable information for this purpose. Actually, it constitutes one of the most important steps in the drug's development [9].

Notably, it should be kept in mind that failure to identify an appropriate dose can cause delays throughout the drug development process, increase cost and, at last not at least, compromise the commercial success of the marketed drug, and even cause widespread harm to patients. This underlines the need for understanding the scientific and regulatory aspects of dose selection, as well as effective statistical methods for identifying dosing regimens with maximized likelihood of efficacy and minimized risk of toxicity. This will be detailed in Chapter 2.

1.3 Statistical issues of dose-finding phase I trials

As stated above, the dose-finding phase I trials usually represent the first application of a proposed drug to humans. Despite the centrality of these early-phase trials to the process of drug evaluation, they are not well understood, are subject to many popular misconceptions [10] and are still unfamiliar to most physicians. Moreover, besides some guidelines established for phase I trial execution during the 1970s, statistical

considerations have been mostly absent from the design, analysis and reporting of these studies until the 1990s. Nevertheless, phase I trials pose challenging problems for the ethical conduct, for efficient design and for inference (with increased confidence that our estimate is accurate) of these studies, all of which have important statistical content [11].

1.3.1 Ethical concerns

Dose-finding studies involve humans, either healthy volunteers or patients. Therefore, investigators conducting dose-finding trials must adhere to the ethical norms of clinical research. Moreover, as first-in-man studies, the safety of the participating subjects is of primary concern. Phase I trials are studies of, for and (increasingly) with patients. However, the research goals may differ from the patients' goals, and there is no consensus on how to achieve the researcher's goals in the most efficient and ethically appropriate way. Actually, ethical issues have markedly influenced the sample size and the design of these studies.

Firstly, typical sample sizes of dose-finding studies are small, commonly about 20 subjects. As mentioned by Gatsonis and Greenhouse [12], perhaps the limited sample size and the ethical concerns, closely related, are the two most serious problems currently facing phase I study designs.

Secondly, some rather special designs are used in dose-finding. Indeed, the design has long been governed by the ethical constraint of minimizing subjects treated at toxic doses. Therefore, randomly assigning patients between several dose levels appeared unacceptable. Doing so, some subjects will be assigned to low doses that are known to be suboptimal in terms of efficacy while other subjects will be exposed to high doses that are very toxic. Initial standard phase I designs were rule-based (or algorithm-based) dose-escalation schemes in which the dose is gradually increased throughout the experiment from the lowest dose level until some desired response or unwarranted toxicity is observed (see Chapter 3). Inpatient dose-escalation is frequently not permitted because of concerns about cumulative toxicity obscuring effects at a subsequent dose level. However, accelerated titration designs have been developed using such an within-patient dose escalation [13], notably in healthy volunteers. Such designs will be developed in Chapter 4.

In traditional between-patient group escalation designs, because of fears for safety, an escalation of doses by group takes place. Thus, the first dose will be examined in one group of patients before proceeding to study the next dose and so forth. The starting dose and the escalation scheme, which both determine the distribution of patients at potentially toxic dose levels, deal with the issue of safety. Issues raised in determining the starting dose and escalation scheme will be discussed below (Section 1.3.2).

Nevertheless, when conducted in AIDS or cancer patients, phase I trials also have a therapeutic aim. Thus, another ethical obligation, more recently reported, is to maximize the chance that the dose which an individual receives has the potential for therapeutic value or, in other words, that the number of patients treated at ineffective doses should also be minimized [14]. Therefore, the process of dose escalation is

governed by a fundamental conflict. On the one hand, there is a need to go slowly in order to avoid a sudden jump from no observable toxicity to a lethal dose; on the other hand, there is a need to go rapidly, so that large numbers of patients are not treated at ineffective doses. In this setting, there is a need to balance the concern for patient safety when being treated with an unknown agent, as reflected in careful dose escalation and the desire to treat at doses that will be close to the recommended phase II dose, thus increasing the likelihood of benefit. Thus, one should design cancer phase I trials to minimize both the number of patients treated at low, nontherapeutic doses, as well as the number given severely toxic overdoses. Attention has been directed towards the objective of treating as many phase I patients as possible close to the MTD (see Chapter 6). Alternate designs were developed to treat as few patients as possible at a biologically inactive dose level. This will be discussed in Chapter 8.

Finally, two fundamental ethical challenges are often raised about phase I cancer trials, namely the risk–benefit ratio and informed consent, but available data do not support these objections [15].

1.3.2 Design

During the past decade, the importance of properly designed early (phase I) trials has led to dramatic changes in their traditional 3 + 3 design (see Chapter 3), including selection of starting dose and rapidity of dose-escalation.

1.3.2.1 Starting dose

To begin human testing, a safe starting dose is needed. This is an important step in dose-findings for anticancer drugs [16]. Guidelines for determining phase I starting doses for chemotherapeutic agents were developed following retrospective reviews that compared specific toxic dose levels in animals and in humans [17]. Actually, the current preclinical toxicology does provide the basis for a safe starting dose, tailored to potency in rodents. As safety was the primary concern, the starting dose has been long calculated as the highest fraction of a specific toxic dose level, in a particular animal species. Early in the history of the development process for phase I trials, preclinical toxicology studies were used to define the toxic dose low (TDL), or first toxic dose, in dogs and monkeys for safety considerations in clinical trials; one-third of the TDL in the more sensitive species was chosen as a safe starting dose for phase I trials of antitumor agents [18]. However, later studies suggested that the LD10 in mice, i.e. the dose that was lethal to 10 % of animals, would be a safe starting dose and would decrease the number of dose-escalation steps required to reach the MTD as compared with one-third of the TDL in dogs and monkeys [19]. One-tenth of the mouse LD10, expressed in milligrams per meters squared, has historically been found to be a safe starting dose in humans, as long as that dose is not lethal or life-threatening to a second species (e.g. rat, dog). This provides the commonly used basis for initial doses in phase I studies [20]. The question of whether higher starting doses can be

safely used was recently pointed out [21], notably through the use of interspecies scaling [16]. This could help to save time and avoid unnecessary steps in attaining MTD in dose-finding experiments.

1.3.2.2 Dose-escalation

Standard phase I design is a rule-based dose-escalation scheme, in which the dose is gradually increased throughout the experiment until some desired response or unwarranted toxicity is observed. Until recently, it was expected that dose-escalation for all anticancer drugs would continue until limited by toxicity.

The traditional phase I design is algorithm-based. For escalating doses above the starting dose, the modified Fibonacci search scheme [22] with escalations in decreasing increments was initially applied to phase I studies on nitrosourea [23]. Once the starting dose has been evaluated, the rate of escalation is empirically defined by a modified Fibonacci series. The initial dose escalation is thus rapid (100 %, or doubling of the dose) and narrows down with successive increases until the 30–35 % range is reached (actually, dose increases of 100 %, 65 %, 50 %, 40 % and 30–35 % thereafter). In theory, this approach would decrease the risk of overshooting the MTD as it is approached, but scientific justification for it is lacking.

The widely used standard phase I design (so called 3 + 3 design) is very simple. It uses three patients at each dose level until one of three patients have DLT; then three more patients are added. If none of the three additional patients have DLT, escalation continues; if one or more of the additional patients have DLT and only three patients were evaluated at the next lower dose, additional patients are added to the lower dose level. Otherwise, the trial is stopped.

This standard approach treats the MTD as being identifiable from the data [24], and thus is a statistic rather than a parameter. Actually, no estimation is involved. Of note, the derivation of exact formulae for statistical quantities of interest has recently been proposed [25, 26].

By design, the MTD estimate relies heavily on the actual cohort of patients treated and the order in which they enter the study. A review of phase I studies undergone during the period 1977–89 has documented that the usual methods for choosing starting doses and dose-escalation schemes for phase I studies are overly conservative and delay opportunities for therapeutic benefit in phase I and subsequent phase II trials [27]. Because of the small number of patients who actually receive the recommended phase II dose, there is a large uncertainty in the toxicity rate associated with that dose. For instance, if all three patients of a hypothetical cohort experience dose-limiting toxicity, most investigators would agree that the recommended phase II dose has been exceeded. However, even in that setting, one must acknowledge that the 95 % exact confidence interval for the incidence of dose-limiting toxicity is 37–100 %. Many investigators would not accept a 33% rate of dose-limiting toxicity (two out of six patients) at the recommended phase II dose; however, the 95 % exact confidence interval for this dose-limiting toxicity rate is quite broad (6–73 %). Otherwise, statisticians feel that conventional methods in widespread use for dose-finding in phase I clinical trials are unreliable, pointing out the need for alternative methods [28]. As

an illustration, all participants in a workshop on phase I trial designs in cancer drug development stated that they no longer would routinely use standard dose escalation when designing a phase I trial [21].

Therefore, other rules for escalating and de-escalating dose levels have been proposed. For more rapid and efficient dose-escalation schemes, pharmacologically based designs have been proposed in recent years. Like the traditional modified Fibonacci-based design, these methods use toxicity, and specifically DLT, as the endpoint of the trial, using toxicologic projections that are based on pharmacological information from preclinical models. Such pharmacokinetically guided determination of the MTD has been suggested in cancer [11, 14, 29], using preclinical and pharmacokinetic information in the choice of starting dose, rate of escalation and design. They will be discussed in Chapter 4.

Finally, statistical-based designs, which treat the MTD as a population parameter, are driven by accumulating patient observations that refine a model predictive of the optimal dose. They will be developed below. Whatever the inference, iterative fitting of the dose–effect model, as data accumulate, allows model-guided dosing. A new data-driven dose-escalation method, adaptive dose-finding (ADF), has been proposed, which combines the standard rule-based method with a model-guided method [30].

Most of these new proposals will be discussed in Parts III, IV and V.

1.3.3 Inference

The primary purpose of a phase I clinical trial is efficiently and accurately to determine the dose of a new drug or therapeutic agent being studied for future applications (in phase II and then in phase III trials). Thus, phase I trials address an estimation problem rather than the testing of a hypothesis. The fact that phase I clinical trials are not hypothesis-driven has been considered as potentially the reason why statistical considerations have largely been ignored in these trials [24]. Moreover, as exposed above, for a long time phase I studies were often done in small numbers of patients, with the optimal dose usually administered to only a minority of the patients treated in the trial. As a result, neither efficacy nor safety data were considered to be reliable.

In the traditional algorithm-based designs, it is common to define the MTD as the previous lower dose level reached in the trial. Nonetheless, designs for phase I clinical trials should be concerned with efficiency of estimation of the MTD. Actually, it has been argued that efficiency is a more relevant consideration than convergence in such very small trials [31]. Notably, any statistical discussion of issues in dose-finding trials should begin with a proposal for a formal inferential framework for analysis. Within this framework, one would specify a form for the dose–response relationship, define the quantities of interest and proceed to quantify the degree of uncertainty about these quantities.

1.3.3.1 Specify a form for the dose–response relationship

The basic and central assumption in cancer dose-finding is that the therapeutic and toxic effects of a treatment are related to the concentration of the treatment in blood,

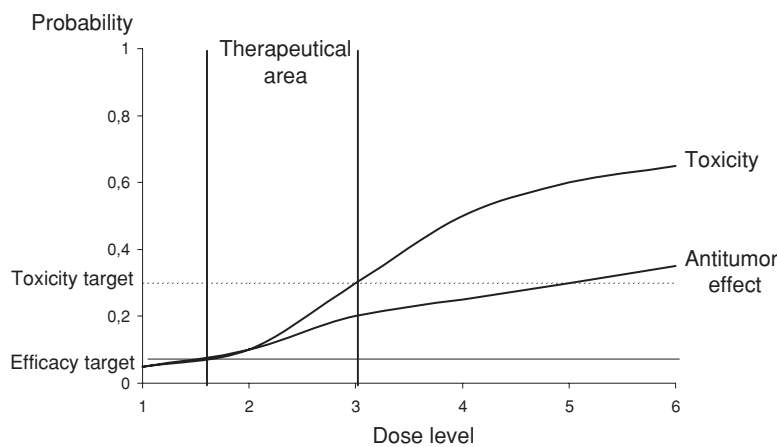


Figure 1.1 Dose–response relationships.

which are dose-dependent (Figure 1.1). In the case of cytotoxic agents, this concept of dose response has greatly influenced the thought process of oncologists, assuming that the higher the dose, the greater the likelihood of efficacy (‘more seems better’ [32]). A monotonic relationship is commonly stated, although some authors have developed methodology for the design and analysis of nonmonotonic relationships.

The mathematical function describes the hypothesized relationship between the incidence of DLT and dose in the target population. This parametric quantal response model is reasonably predicted to assume a sigmoid shape that can be generically described by a probit or logit function, and for which the MTD must be estimated first [6, 12, 33, 34]. This is more in line with the statistical development in bioassay with a binary response, which dates back from more than five decades [35]. Usually, maximum likelihood or least-squares techniques are used to estimate the model parameters [6, 13, 26], though Bayesian approaches have also been proposed since 1972 [5], such as the original continual reassessment method (CRM) [34] (see Chapter 6), the decision–theoretic approach of Whitehead and Brunier [36] (see Chapter 7), escalation with overdose control (EWOC) [33] (see Chapter 8) and Bayesian optimal designs for phase I trials of Haines, Perevozskaya and Rosenberger [31].

Of note, completely nonparametric approaches have been proposed more recently [33, 37–41].

1.3.3.2 Define the quantities of interest

Because most of these agents exhibit a monotonic dose-toxicity relationship, dose–related toxicity is regarded, in general, as a surrogate for efficacy. In this setting, the achievement of significant, but reversible, toxicity becomes desirable, so that most statistical methods for dose-finding in cancer phase I clinical trials determine a

maximum tolerable dose (MTD) based on toxicity alone, while ignoring therapeutic response [34, 42–46]. Virtually, all designs for dose-finding in phase I characterize patient response by a binary [45, 47] or possibly ordinal [48] toxicity variable.

In the traditional algorithm-based designs (either the 3 + 3 design, see Chapter 3, or the up-and-down designs, see Chapter 5), the MTD was first defined as the dose level at which at least one-third (conventionally in no more than two of six patients) of the patients experience DLT. Many use the term MTD for the next lower dose, which is frequently the recommended phase II dose. It is important to note, however, that the MTD will depend upon the criteria set for DLT. These sometimes vary because investigators may be more conservative or more aggressive in their definition of acceptable and unacceptable toxicity. In all cases, the MTD is treated as being observable from the data. Such a definition of MTD appeared vague from a statistician's point of view. As stated by Storer [45], a strict quantitative definition of the MTD is seldom acknowledged, although it is often taken as a specific percentile of the tolerance distribution of the treatment, i.e. the dose that would be expected to produce a toxic response in a specified proportion of patients. The percentile of interest could vary according to the nature and seriousness of the toxicities produced by the drug under investigation. The 33rd percentile has been proposed for cancer chemotherapy drugs [45], but other percentiles may be reasonable choices. In this setting, the MTD is an unknown population parameter that can be estimated from the data. Such an interpretation of the MTD is more in line with historical statistical developments in bioassay with quantal response curves for a binary response.

Phase I trials are designed to determine the recommended phase II dose for a population, not an individual patient. Since there is usually substantial interpatient variability in toxic effects, the recommended phase II dose will always be an imprecise estimate of the optimal dose for the individual patient. An alternative target dose, the patient-specific MTD, has been proposed explicitly to take into account a patient's history via an augmented dose–response model [49]. Recently, other researchers have defined the MTD as the highest dose with toxicity risk not exceeding the tolerable toxicity [33, 50]. Otherwise, some authors have developed methods to account for differences among toxicities, using ordinal-valued toxicities according to severity (or grade). This will be developed in Chapter 12. Strategies for dose-finding where both toxicity and efficacy are considered have been developed. They will be discussed in Chapter 14.

Finally, toxicity may no longer be an appropriate endpoint in some phase I studies [51], so it is no longer expected that determination of the MTD will be the universal endpoint of a phase I investigation. These include studies of drugs characterized by a dose–response curve that reaches a plateau at nontoxic doses, or that have the desired effect at doses without significant toxicity. Probable examples include angiogenesis inhibitors and colony stimulating factors. Toxicity would also not be an appropriate endpoint for phase I studies of agents with a bell-shaped dose–response curve and a dose–toxicity curve that rises after the maximal therapeutic dose. There are already several classes of molecules undergoing clinical evaluation for which the MTD was not determined and/or relevant to the drug's use. Probable examples include

interferons, some interleukins and negative regulators of hematopoiesis. Thus, with the development of novel biologic agents such as these, the use of toxicity as an endpoint would result in unnecessary toxicity and wasted resources. The need for new endpoints is also apparent for new modalities including monoclonal antibodies and gene therapy, which may produce nonspecific and sporadic toxicities that are not clearly dose-related. In many cases, plasma concentration has served as the alternative endpoint. Pharmacokinetics assessment and correlation of drug levels with target effects, in addition to pharmacodynamic measures, will be necessary. At last, it is critical to validate that the proposed target endpoints are correlated with activity [52].

Of note, a recent survey of completed phase I trials for targeted, noncytotoxic agents showed that nontraditional endpoints (such as measures of molecular drug effects in tumor or surrogate tissue) were not routinely incorporated in the study design and rarely formed the basis for dose selection, which is still based on toxicity traditional endpoints [53].

1.3.3.3 Proceed to quantify the degree of uncertainty about these quantities

Finally, it is important to include measures of variability in such studies. They are usually nonreported from traditional algorithm-based designs, in contrast to statistical-based designs.

The main approach is to quote confidence limits. Confidence limits, due to Neyman, are the limits of usually the 95 % confidence interval, which represents all the parameter values compatible with the data accumulated in the trial. In a Bayesian framework, credible intervals for the model parameter can be computed. This gives an interval such that (say) the model parameter lay between these limits with 95 % probability. Of note, the numerical integration procedures are often ignored in reporting these Bayesian studies, though of prime interest in evaluating its reliability [24].

1.4 Conclusion

Dose-finding experiments refer to a broad class of early development trial designs whose purpose is to find a dose of treatment that is optimal with respect to simple criteria, namely toxicity, efficacy and a low risk of side effects. The tendency has been to focus on a dose–safety association. Therefore, dose-finding studies are typically phase I clinical trials, based on very small, uncontrolled, sequential studies of subjects to determine the maximum tolerated dose (MTD) of an experimental drug that will be used in further phase II or III trials. While performed in many areas of medicine, they are mostly devoted to cancer studies due to the severe toxicity of cytotoxic drugs, assumed to be related to the dose and, expectedly, to the efficacy of the drug. Such phase I trials are mostly concerned with protecting patients from being assigned to highly toxic doses and with efficiency of estimation.

A number of philosophical and statistical issues emerge from the design and analysis of dose-finding studies, with two main divergent schools. While algorithm-based designs define the MTD as a statistic computed from data, statistical (or model)-based designs consider the MTD as a parameter of a monotonic dose-response curve to be estimated from data. However, whatever the statistical grounds of the proposed techniques, there is still an obvious gap between the statistical approaches developed in this setting and the common use of standard methods that have been shown to be inefficient for several decades. Indeed, despite proposed new methodologies for phase I trials, very few are being used in practice and many of the methods currently used in phase I trial design date back to the 1970s [52]. This was notably observed through a review of 46 phase I trials of single cytotoxic agents in adult solid tumors published between 1993 and 1995 [54].

The purpose of this book is to review the main innovative statistical approaches for dose-finding in phase I/II clinical trials, from both algorithm- and model-based designs, with emphasis on general concepts and practical considerations.

References

1. J. Kotz, Z. Knapik, W. Lubczynska-Kowalska, M. Houszka, J. Lapinska and M. Rybak (1988) Effect of a single dose of 96 % ethanol on rat gastric mucosa. Part I. Morphological studies. *Mater. Med. Policy*, **20**(2), 67–9.
2. P.R. Freeman (1970) Optimal Bayesian sequential estimation of the median effective dose. *Biometrika*, **57**, 79–89.
3. M. Mendoza (1990) A Bayesian analysis of the slope ratio bioassay. *Biometrics*, **46** (4), 1059–69.
4. F.L. Ramsey (1972) A Bayesian approach to bioassay. *Biometrics*, **28** (3), 841–58.
5. R. K. Tsutakawa (1972) Design of experiment for bioassay. *J. Am. Statist. Assoc.* **67**, 584–90.
6. J.R. Murphy and D.L. Hall (1997) A logistic dose-ranging method for phase I clinical investigations trials. *J. Biopharm. Statistics*, **7**(4), 635–47.
7. B.E. Storer (2001) Phase I trials, in *Biostatistics in Clinical Trials* (eds C. Redmond and T. Colton), John Wiley & Sons, Ltd, Chichester, pp. 337–42.
8. G. Emilien, W. van Meurs and J.M. Maloteaux (2000) The dose-response relationship in phase I clinical trial design and beyond: use, meaning and assessment. *Pharmacology and Therapeutics*, **88**, 33–58.
9. D.D. Von Hoff, J. Kuhn and G.M. Clark (1984) Design and conduct of phase I trials, in *Cancer Clinical Trials: Methods and Practice* (eds M.E. Buyse, M.J. Staquet and R.J. Sylvester), Oxford University Press, Oxford, pp. 210–20.
10. ASCO (1997) Clinical practice guidelines for the treatment of unresectable non-small-cell lung cancer. Adopted on May 16, 1997 by the American Society of Clinical Oncology, *J. Clin. Oncology*, **15** (8), 2996–3018.
11. M.J. Ratain, R. Mick, R.L. Schilsky and M. Siegler (1993) Statistical and ethical issues in the design and conduct of phase I and II clinical trials of new anticancer agents. *J. Natl Cancer Inst.*, **85**(20), 1637–43.

12. C. Gatsonis and J.B. Greenhouse (1992) Bayesian methods for phase I clinical trials. *Statistics in Medicine*, **11**(10), 1377–89.
13. R. Simon, B. Freidlin, L. Rubinstein, S.G. Arbuck, J. Collins and M.C. Christian (1997) Accelerated titration designs for phase I clinical trials in oncology. *J. Natl Cancer Inst.*, **89**(15), 1138–47.
14. J.M. Collins, D.S. Zaharko, R.L. Dedrick and B.A. Chabner (1986) Potential roles for preclinical pharmacology in phase I clinical trials. *Cancer Treatment Rep.*, **70**(1), 73–80.
15. M. Agrawal and E.J. Emanuel (2003) Ethics of phase 1 oncology studies. *JAMA*, **290**(8), 1075–82.
16. I. Mahmood (2001) Interspecies scaling of maximum tolerated dose of anticancer drugs: relevance to starting dose for phase I trials. *Am. J. Therapeutics*, **8**(2), 109–16.
17. E.J. Freireich, E.A. Gehan, D.P. Rall, L.H. Schmidt and H.E. Skipper (1966) Quantitative comparison of toxicity of anticancer agents in mouse, rat, hamster, dog, monkey, and man. *Cancer Chemotherapy Rep.*, **50**(4), 219–44.
18. D.J. Prieur, D.M. Young, R.D. Davis, D.A. Cooney, E.R. Homan, R.L. Dixon and A.M. Guarino (1973) Procedures for preclinical toxicologic evaluation of cancer chemotherapeutic agents: protocols of the laboratory of toxicology. *Cancer Chemotherapy Rep.* **3**, **4**(1), 1–39.
19. J.S. Penta, M. Rozenzweig, A.M. Guarino and F.M. Muggia (1979) Mouse and large-animal toxicology studies of twelve antitumor agents: relevance to starting dose for phase I clinical trials. *Cancer Chemotherapy and Pharmacology*, **3**(2), 97–101.
20. M. Rozenzweig, D.D. Von Hoff, M.J. Staquet, P.S. Schein, J.S. Penta, A. Goldin, F.M. Muggia, E.J. Freireich and V.T. DeVita Jr (1981) Animal toxicology for early clinical trials with anticancer agents. *Cancer Clin. Trials*, **4**(1), 21–8.
21. E.A. Eisenhauer, P.J. O'Dwyer, M. Christian and J.S. Humphrey (2000) Phase I clinical trial design in cancer drug development. *J. Clin. Oncology*, **18**(3), 684–92.
22. M.A. Goldsmith, M. Slavik and S.K. Carter (1975) Quantitative prediction of drug toxicity in humans from toxicology in small and large animals. *Cancer Res.*, **35**(5), 1354–64.
23. H.H. Hansen, O.S. Selawry, F.M. Muggia and M.D. Walker (1970) Clinical studies with 1-(2-chloroethyl)-3-cyclohexyl-1-nitrosourea (nsc 79037). *Cancer Res.*, **31**(3), 223–7.
24. W.F. Rosenberger and L.M. Haines (2002) Competing designs for phase I clinical trials: a review. *Statistics in Medicine*, **21**(18), 2757–70.
25. Y. Lin and W.J. Shih (2001) Statistical properties of the traditional algorithm-based designs for phase I cancer clinical trials. *Biostatistics*, **2**(2), 203–15.
26. M. Stylianou and N. Flournoy (2002) Dose finding using the biased coin up-and-down design and isotonic regression. *Biometrics*, **58**(1), 171–7.
27. J.S. Penta, G.L. Rosner and D.L. Trump (1992) Choice of starting dose and escalation for phase I studies of antitumor agents. *Cancer Chemotherapy and Pharmacology*, **31**(3), 247–50.
28. T.L. Smith, J.J. Lee, H.M. Kantarjian, S.S. Legha and M.N. Raber (1996) Design and results of phase I cancer clinical trials: three-year experience at M.D. Anderson Cancer Center. *J. Clin. Oncology*, **14**(1), 287–95.
29. J.M. Collins, C.K. Grieshaber and B.A. Chabner (1990) Pharmacologically guided phase I clinical trials based upon preclinical drug development. *J. Natl Cancer Inst.*, **82**(16), 1321–26.

30. D.M. Potter (2002) Adaptive dose finding for phase I clinical trials of drugs used for chemotherapy of cancer. *Statistics in Medicine*, **21** (13), 1805–23.
31. L.M. Haines, I. Perevozskaya and W.F. Rosenberger (2003) Bayesian optimal designs for phase I clinical trials. *Biometrics*, **59**(3), 591–600.
32. W.M. Hryniuk (1988) More is better. *J. Clin. Oncology*, **6**(9), 1365–7.
33. J. Babb, A. Rogatko and S. Zacks (1998) Cancer phase I clinical trials: efficient dose escalation with overdose control. *Statistics in Medicine*, **17**(10), 1103–20.
34. J. O’Quigley, M. Pepe and L. Fisher (1990) Continual reassessment method: a practical design for phase I clinical trials in cancer. *Biometrics*, **46**(1), 33–48.
35. J. Berkson (1944) Applications of the logistic function to bio-assay. *J. Am. Statist. Assoc.*, **39**, 357–65.
36. J. Whitehead and H. Brunier (1995) Bayesian decision procedures for dose determining experiments. *Statistics in Medicine*, **14**(9), 885–93.
37. Y.K. Cheung and R. Chappell (2002) A simple technique to evaluate model sensitivity in the continual reassessment method. *Biometrics*, **58**(3), 671–4.
38. S.D. Durham, and N. Flournoy (1994) Random walks for quantile estimation, in *Statistical Decision Theory and Related Topics V* (eds S.S. Gupta and J.O. Berger), Springer, New York, pp. 467–76.
39. M. Gasparini and J. Eisele (2000) A curve-free method for phase I clinical trials. *Biometrics*, **56**(2), 609–15.
40. J. O’Quigley (2002) Curve-free and model-based continual reassessment method designs. *Biometrics*, **58**(1), 245–9.
41. J. O’Quigley, X. Paoletti and J. Maccario (2002) Non-parametric optimal design in dose finding studies. *Biostatistics*, **3**(1), 51–6.
42. S.N. Goodman, M.L. Zahurak and S. Piantadosi (1995) Some practical improvements in the continual reassessment method for phase I studies. *Statistics in Medicine*, **14**(11), 1149–61.
43. E.L. Korn, D. Midthune, T.T. Chen, L.V. Rubinstein, M.C. Christian and R.M. Simon (1994) A comparison of two phase I trial designs. *Statistics in Medicine*, **13**(18), 1799–806.
44. S. Moller (1995) An extension of the continual reassessment methods using a preliminary up-and-down design in a dose finding study in cancer patients, in order to investigate a greater range of doses. *Statistics in Medicine*, **14**(9–10), 911–22.
45. B.E. Storer (1989) Design and analysis of phase I clinical trial. *Biometrics*, **45**(3), 925–37.
46. P.F. Thall, J.J. Lee, C.H. Tseng and E.H. Estey (1999) Accrual strategies for phase I trials with delayed patient outcome. *Statistics in Medicine*, **18**(10), 1155–69.
47. N.L. Geller (1984) Design of phase I and II clinical trials in cancer: a statistician’s view. *Cancer Invest.*, **2**(6), 483–91.
48. B.N. Bekele and Y. Shen (2004) A Bayesian approach to jointly modeling toxicity and biomarker expression in a phase I/II dose-finding trial. Technical report, M. D. Anderson Cancer Center, The University of Texas.
49. A.T.R. Legedza and J.G. Ibrahim (2001) Heterogeneity in phase I clinical trials: prior elicitation and computation using the continual reassessment method. *Statistics in Medicine*, **20**(6), 867–82.
50. D.H. Leung and Y.G. Wang (2002) An extension of the continual reassessment method using decision theory. *Statistics in Medicine*, **21**(1), 51–63.

51. E.L. Korn, S.G. Arbuck, J.M. Pluda, R. Simon, R.S. Kaplan and M.C. Christian (2001) Clinical trial designs for cytostatic agents: are new approaches needed? *J. Clin. Oncology*, **19**(1), 265–72.
52. S.G. Arbuck (1996) Workshop on phase I study design. Ninth NCI/EORTC New Drug Development Symposium, Amsterdam, March 12, 1996. *Ann. Oncology*, **7**(6), 567–73.
53. W.R. Parulekar and E.A. Eisenhauer (2004) Phase I trial design for solid tumor studies of targeted, non-cytotoxic agents: theory and practice. *J. Natl Cancer Inst.*, **96**(13), 990–7.
54. S.F. Dent and E.A. Eisenhauer (1996) Phase I trial design: are new methodologies being put into practice? *Ann. Oncology*, **7**(6), 561–6.