PART I

# Clinical Trials

CHAPTER 1

# Phase I Clinical Trials

***Anastasia Ivanova***
*Department of Biostatistics, University of North Carolina, Chapel Hill, North Carolina*

***Nancy Flournoy***
*Department of Statistics, University of Missouri, Columbia, Missouri*

## 1.1 INTRODUCTION

Phase I trials are conducted to find a dose to use in subsequent trials. They provide data on the rate of adverse events at different dose levels and provide data for studying the pharmacokinetics and pharmacology of the drug. Dose-finding studies that involve therapies with little or no toxicity often enroll healthy volunteers and usually have a control group. Trials in oncology and other life-threatening diseases such as HIV enroll patients because treatments are usually highly toxic and to enroll healthy volunteers would not be ethical. The primary outcome for phase I trials in oncology and HIV is typically dose-limiting toxicity. Such studies require different design strategies.

In Section 1.2, we review dose-finding procedures used in healthy volunteers. In Section 1.3 we describe dose-finding procedures for trials with toxic outcomes enrolling patients. In Section 1.4, we list some other design problems in dose finding.

## 1.2 PHASE I TRIALS IN HEALTHY VOLUNTEERS

Buoen et al. [7] reviewed designs that are used for studying first-time-in-human drugs by looking at 105 studies published in five major pharmacology journals since 1995. In this section we briefly summarize their findings. Bouen et al. found that first-time-in-human studies usually enroll healthy volunteers; most are placebo-controlled and more than half are double-blind. The placebo group is included to reduce observer bias and sometimes to

*Statistical Advances in the Biomedical Sciences*, edited by Atanu Biswas, Sujay Datta, Jason P. Fine, and Mark R. Segal

enable comparison of the active drug with placebo. Usually three to eight dose levels are investigated. Doses are selected using linear, logarithmic, Fibonacci, modified Fibonacci dose escalation patterns, or some combinations of these. The popular modified Fibonacci procedure escalates doses in relative increments of 100%, 65%, 50%, 40%, and 30% thereafter.

The simplest pattern of dose administration being used in first-time-in-human studies is the *parallel single-dose design* in which a single dose is administered once. Multiple administrations of the same dose are referred to as *parallel multiple-dose design*. Parallel dose administration was found to be the most frequently used procedure in first-time-in-human studies. In a typical trial with parallel dose administration, subjects are assigned in cohorts consisting of eight subjects, with six assigned to the active treatment and two assigned to a control. All treated subjects in a cohort receive the same dose. Doses are increased by one level for each subsequent cohort. The trial is stopped when an unacceptable number of adverse events is observed, the highest dose level is reached, or for other reasons. The "target dose," the dose recommended for future trials, is usually determined on the basis of the rates of adverse events at dose levels studied and/or on pharmacokinetic parameters.

More complex dose administration patterns were found to involve the administration of several different dose levels to each patient. In such trials, the healthy subjects are given some rest time between administrations to minimize the carryover effect. One such pattern is referred to as an *alternating crossover design*. An example of an alternating crossover design for a study with six doses is as follows:

| | | | |
|---|---|---|---|
| Cohort 1: | Dose 1 | REST | Dose 4 |
| Cohort 2: | Dose 2 | REST | Dose 5 |
| Cohort 3: | Dose 3 | REST | Dose 6 |

Another dose administration pattern is the grouped crossover escalation. An example of this pattern for a trial with four dose levels is as follows:

Cohort 1

| | | | |
|---|---|---|---|
| Subject 1 | Placebo | Dose 1 | Dose 2 |
| Subject 2 | Dose 1 | Placebo | Dose 2 |
| Subject 3 | Dose 1 | Dose 2 | Placebo |

Cohort 2

| | | | |
|---|---|---|---|
| Subject 1 | Placebo | Dose 3 | Dose 4 |
| Subject 2 | Dose 3 | Placebo | Dose 4 |
| Subject 3 | Dose 3 | Dose 4 | Placebo |

Sheiner et al. [41] reviewed parallel and crossover designs and methods for analyzing the data obtained in such studies. They point out ethical problems and a lack of representativeness in these designs. Sheiner et al. [41] advocated using a dose administration pattern that they call the *dose escalation design*:

> According to the dose-escalation design all subjects are given a placebo dose first. If after some predefined time period the response fails to satisfy a certain clinical endpoint and no unacceptable toxicity is seen, the dose is increased by one level. This process is repeated at

each dose level until either the clinical endpoint is reached or the highest dose is attained. If the response is adequate at any dose, the dose is maintained at that level for the duration of the study.

The main obstacle to using this design is the lack of formal statistical methods for data analysis.

Girard et al. [17] studied the effects of several confounding factors on trials that use parallel dose, crossover and dose escalation designs by simulations. They concluded that the presence of nonresponders biases the estimate of the dose producing 50% of the maximum effect, in all three designs. However, other confounders such as carryover effects only bias the results of trials in which the dose escalation design is used.

Buoen et al. [7] conclude that, although "the development of study designs and evaluation methods for cancer trials is extensive, ... formal statistically based methods ... are unusual in phase I dose-escalation trials in healthy volunteers." This lack and the recognition of need present both challenges and opportunities to the statistical research community.

## 1.3 PHASE I TRIALS WITH TOXIC OUTCOMES ENROLLING PATIENTS

In many phase I trials in which the subjects are patients, rather than healthy volunteers, the goal is to find the dose that has a prespecified toxicity rate. This is particularly true in oncology. In these trials, the primary outcome is typically binary: dose-limiting toxicity? Yes or no. For example, the dose-limiting toxicity (DLT) in radiotherapy and chemotherapy studies is usually defined as treatment-related nonhematological toxicity of grade 3 or higher or treatment-related hematological toxicity of grade 4 or higher. The *maximally tolerated dose* (MTD) is statistically defined as the dose at which the probability of DLT is equal to the some prespecified rate $\Gamma$. The typical underlying model assumption is that the probability of toxicity is a nondecreasing function of dose, even though decreasing toxicity rates at high doses have been observed [43].

Preclinical studies in animals often attempt to determine the dose with approximately 10% mortality (e.g., the murine $LD_{10}$). In first-in-human toxicity studies, one-tenth or two-tenths of the dose considered to be equivalent to the murine equivalent, expressed in milligrams per meter squared ($mg/m^2$), is generally used as a starting dose in escalation procedures. The starting dose is anticipated to be 5–10-fold below the dose that would demonstrate activity in humans. In trials with oral drugs, only certain doses can be used; therefore, the set of possible doses is fixed in advance. The set of possible doses is often chosen according to the modified Fibonacci sequence.

In dose-finding trials in oncology, patients may receive a single dose of a drug or multiple administrations of the same dose. To address ethical concerns similar to those of Sheiner et al. [41] and to shorten trial duration, Simon et al. [42] introduced acceleration titration designs. Such designs allow intrapatient dose escalation if no toxicity is observed in a patient at the current dose. A patient goes off study or the patient's dose is reduced if toxicity is observed. Although appealing from an ethical perspective, this approach is not widely used for the same reason as in the hesitation to use Sheiner's dose escalation design. In the rest of this chapter, we review methods with parallel dose administration.

One cannot begin to detail all designs that have been used with parallel administration for dose finding in patients with dose-limiting toxicity. Some popular procedures are ad hoc, as are

the designs used in healthy volunteers. Others were developed with various desirable characteristics. We discuss the most popular procedures, but our choice is admittedly biased by our own interests.

### 1.3.1 Parametric versus Nonparametric Designs

Designs for dose finding can be classified as parametric or nonparametric. Non-parametric designs are attractive because they are easy to understand and implement; the decision rules are intuitive and their implementation does not involve complicated calculations. By *nonparametric*, we mean that no parametric representation of the dose–response relationship is used in the design's treatment allocation rule. In this chapter, we discuss several Markovian and Markovian-motivated non-parametric up-and-down designs and the $A + B$ designs of Lin and Shih [31]. We also discuss non-parametric designs in which the treatment allocation rule is based on isotonic estimates of the dose–response function. These are called *isotonic designs*.

Then we describe some parametric designs that assume one- or two-parameter models for the dose–toxicity relationship. Popular parametric designs include the continual reassessment method [33] and escalation with overdose control [2].

With the Markovian and Markovian-motivated designs, treatment assignments typically cluster unimodally around a specific dose, and the key to their effectiveness is to select design parameters so as to center the treatment distribution judiciously [11]. For example, for toxicity studies with increasing dose–response functions, these designs can be constructed to cluster treatments around the unknown dose with prespecified "target" toxicity rate $\Gamma$.

In other designs that allow multiple escalations and deescalations of dose, treatment assignments first fluctuate around the MTD and then converge assignments to the MTD. Such designs include, for example, the continual reassessment method [33] and isotonic designs [29].

### 1.3.2 Markovian-Motivated Up-and-Down Designs

In up-and-down designs, the next dose assigned is never more than one level distant from the dose given to the current cohort of patients. Such designs are appealing in dose-limiting toxicity studies because of the potentially devastating consequences of abruptly making major changes in dosing. Many ad hoc up-and-down procedures exist, including the most widely cited design in oncology, that is, the $3 + 3$ design [44,28]. The $3 + 3$ design is a special case of the $A + B$ designs [31]. It is important in trials with patients who are critically ill not to assign too many patients to low, ineffective doses. The $A + B$ designs address this concern by assigning $A$ patients to the lower doses and assigning $A + B$ patients to doses closer to the target.

Before describing the $A + B$ designs, we review a fundamental theorem that is useful for characterizing the Markovian up-and-down design. Let $p_k$, $q_k$, and $r_k$ denote the probability of increasing, decreasing, and repeating dose $d_k$, respectively. Assume that these probabilities depend only on $d_k$, $k = 1, \ldots, K$. Furthermore, assume that $p_k$ decreases with dose, whereas $q_k$ increases with dose. Let $d_\kappa$ denote the largest dose such that $p_{\kappa-1} \geq q_\kappa$. The stationary distribution for Markov chain designs with transition probabilities $p_k$, $q_k$, $r_k$ exists uniquely if the Markov chain is recurrent, irreducible, and aperiodic. Under these conditions, Durham and Flournoy [11] proved that the stationary distribution of the dose assignments is unimodal and the mode occurs at $d_\kappa$. Additionally, if $p_{\kappa-1} = q_\kappa$, then the mode spans $d_{\kappa-1}$ as well as $d_\kappa$.

Convergence of the dose assignments to their stationary distribution is reached exponentially rapidly, so asymptotic results apply well with a relatively small number of treatment

assignments, regardless of the initial dose. Because of the discreteness of the dose space, as a practical approximation, we say that a Markovian up-and-down design "targets" $d_\kappa$ if $p_\kappa = q_\kappa$; treatments will cluster unimodally around this dose. Alternatively, we say that the design targets the toxicity rate $\Gamma$, for which $P\{\text{toxicity} \mid d_\kappa\} = \Gamma$. Markovian up-and-down designs can be characterized using this and other asymptotic and finite sample theory for Markov chains. Techniques are given in Durham and Flournoy [12], Durham et al. [13], Giovagnoli and Pintacuda [16], and Bortot and Giovagnoli [5].

A corollary of the Durham–Flournoy theorem is that treatments from the traditional up-and-down design of Dixon and Mood [10] are distributed unimodally around $d_\kappa = \text{LD}_{50}$, regardless of the underlying (increasing) dose–response model. In this procedure, the dose is decreased if a toxicity is observed and increased otherwise. So $p_k = P\{\text{toxicity} \mid d_k\}$ and $q_k = 1 - p_k = P\{\text{toxicity} \mid d_k\}$ (except at $k = 1$ or $K$). Solving $p_k = q_k$ yields $p_\kappa = 0.50$.

Durham and Flournoy [11,12] generalized the Dixon–Mood decision rule by using a biased coin, together with the Durham–Flournoy theorem, to provide a procedure that targets any given toxicity rate $\Gamma$. This procedure was not well received in oncology trials because clinicians were averse to using randomization in phase I treatment allocation rules.

Using cohorts at each dose, the Durham–Flournoy theorem was employed by Gezmu and Flournoy [15] to devise treatment allocation rules without randomization that still target a given toxicity rate $\Gamma$. However, the set of possible targets is limited by the group size. Some examples they give of $\Gamma$ that are possible with groups of size 2 are 0.29, 0.50, and 0.71; with groups of size 3, they are 0.21, 0.35, 0.50, 0.65, and 0.79; and with groups of size 4, they are 0.16, 0.27, 0.38, 0.39, 0.50, 0.61, 0.62, 0.73, and 0.84. Procedures for values of $\Gamma$ greater than 0.5 are useful for efficacy studies, but not toxicity studies. Gezmu and Flournoy [15] show that each of these target values can be found as a direct application of the Durham–Flournoy theorem; details justifying this application are given by Ivanova et al. [25]. Antognini et al. [1] generalize the Gezmu–Flournoy group procedure to target any $\Gamma \in (0,1)$ by introducing a randomization procedure. This is clever, but will probably not be any more attractive to oncologists than was the biased coin design of Durham and Flournoy [11].

Ivanova et al. [25] take a different approach to adapting the group up-and-down design so that it will target any given $\Gamma \in (0,1)$. They call their procedure the *cumulative cohort design*, which is as follows.

*Cumulative Cohort Design* Suppose that the most recent assignment was to dose $d_j$. Let $\hat{q}_j$ be the cumulative proportion of toxicities at $d_j$, and let $\Delta > 0$ denote a design parameter. Then

1. If $\hat{q}_j \leq \Gamma - \Delta$, the next group of subjects is assigned to dose $d_{j+1}$.
2. If $\hat{q}_j \geq \Gamma + \Delta$, the next group of subjects is assigned to dose $d_{j-1}$.
3. If $\Gamma - \Delta < \hat{q}_j < \Gamma + \Delta$, the next group of subjects is assigned to dose $d_j$.

Appropriate adjustments are made at the lowest and highest doses.

An intuitive choice of the parameter $\Delta > 0$ in the cumulative cohort design is close to 0. For example, with $\Delta = 0.01$ and moderate sample sizes, the dose will be repeated if the estimated toxicity rate is exactly equal to $\Gamma$, and changed otherwise. Ivanova et al. [25] suggested choosing $\Delta$ to maximize the total number of subjects assigned to the MTD over a set of dose–toxicity scenarios. For example, for moderate sample sizes they recommended using $\Delta = 0.09$ if $\Gamma = 0.10$, 0.15, 0.20, or 0.25; $\Delta = 0.10$ if $\Gamma = 0.30$ or 0.35; $\Delta = 0.12$ if $\Gamma = 0.40$; and $\Delta = 0.13$ if $\Gamma = 0.45$ or 0.50. Ivanova et al. [25] demonstrated via simulations that $\Delta = 0.01$ and choosing their recommended values of $\Delta$ yield similar frequency of correctly

selecting the MTD. However, the cumulative cohort design with their recommended D values assigns significantly more patients to the MTD.

The $A + B$ designs as given by Lin and Shih [31] begin like the first run of a Markovian group up-and-down design, but the design is switched when the dose would otherwise be repeated and stopped (for designs without deescalation) when the dose would otherwise be decreased.

*$A + B$ Design*: Let $A$ and $B$ be positive integers. Let $c_L$, $c_U$, and $C_U$ be integers such that $0 \le c_L < c_U \le A$, $c_U - c_L \ge 2$, and $c_L \le C_U < A + B$. Let $X_A(d_j)$ be the number of toxicities in a cohort of size $A$ assigned to dose $d_j$, and let $X_{A+B}(d_j)$ be the number of toxicities in a cohort of size $A + B$. Subjects are treated in cohorts of size $A$ starting with the lowest dose. Suppose that the most recent cohort was a cohort of $A$ subjects that has been treated at dose $d_j, j = 1, \ldots, K - 1$. Then

1. If $X_A(d_j) \le c_L$, the next cohort of $A$ subjects is assigned to dose $d_{j+1}$.
2. If $c_L < X_A(d_j) < c_U$, the cohort of $B$ subjects is assigned to dose $d_j$; then, if in the combined cohort assigned to $d_j$, $X_{A+B}(d_j) \le C_U$, the next cohort of size $A$ receives dose $d_{j+1}$; otherwise the trial is stopped.
3. If $X_A(d_j) \ge c_U$, the trial is stopped.

The dose that is one level below the dose where unacceptable numbers of toxicities are observed ($\ge c_U$ toxicities in a cohort of size $A$ or $> C_U$ toxicities in a cohort of size $A + B$) is the estimated MTD.

In an $A + B$ design, the frequency of stopping dose escalation at a certain level depends on toxicity rate at this dose as well as on toxicity rate at all lower dose levels. Ivanova [21] used the Durham–Flournoy theorem to derive recommendations for constructing escalation designs and explains how to compute the toxicity rate $\Gamma$ that will be targeted by any given $A + B$ design. The algorithm for selecting parameters $A$, $B$, $c_L$, $c_U$, and $C_U$ for a given target quantile $\Gamma$ is as follows (where Bin = binomial distribution):

1. Find $A$, $c_L$, and $c_U$, $0 \le c_L < c_U \le A$, $c_U - c_L \ge 2$, so that $\Gamma_A$, the solution to the equation $\Pr\{\text{Bin}(A, \Gamma_A) \le c_L\} = \Pr\{\text{Bin}(A, \Gamma_A) \ge c_U\}$, is equal to or slightly exceeds $\Gamma$.
2. Set $B$ (the choice $A \le B$ yields more efficient designs), and given that $\Gamma_{A+B}$ is the solution to the equation $\Pr\{\text{Bin}(A + B, \Gamma_{A+B}) \le C_U\} = 0.5$, find $C_U$ such that $C_U/(A + B) < \Gamma < \Gamma_{A+B}$.

The $3 + 3$ design is a special case of the $A + B$ design with $A = B = 3$, $c_L = 0$, $c_U = 2$, and $C_U = 1$ that target quantiles around $\Gamma = 0.2$. Applying the algorithm above, we obtain

1. $\Gamma_A = 0.35$ is the solution of the equation $\Pr\{\text{Bin}(3, \Gamma_A) \le 0\} = \Pr\{\text{Bin}(3,\Gamma_A) \ge 2\}$; $\Gamma_A = 0.35$ is slightly higher than $\Gamma = 0.2$.
2. $\Gamma_{A+B} = 0.26$ is the solution of the equation $\Pr\{\text{Bin}(3 + 3, \Gamma_{A+B}) \le 2\} = 0.5$, and $C_U(A + B) = 0.17$. Hence, approximate bounds for $\Gamma$ targeted by the $3 + 3$ design are $0.17 < \Gamma < 0.26$.

Exact probability calculations and simulation studies for several dose–response scenarios by Reiner et al. [39], Lin and Shih [31], Kang and Ahn [26,27] and He et al. [20] are consistent with the theoretical calculation above establishing that the $3 + 3$ design selects a dose with

toxicity rate near 0.2. He et al. [20] also showed that if dose levels are selected close to each other, the mean toxicity rate at the dose selected by the 3 + 3 design is slightly lower than the dose selected by trials with a sparser set of possible dose levels.

### 1.3.3 Isotonic Designs

Isotonic designs assume that the dose–toxicity relationship is isotonic and use isotonic estimates of the toxicity rates in the treatment allocation rule. We first review isotonic estimation of the toxicity rates, which are maximum-likelihood estimates for the isotonic model of the data. Let $N(d_j, n)$ be the number of patients assigned to dose $d_j$, and let $X(d_j, n)$ be the number of toxicities at $d_j$ after $n$ patients have been treated. Define $\hat{q}_j = X(d_j, n)/N_j(n)$ for all $j \in \{1, \ldots, K\}$ such that $N(d_j, n) > 0$, and let $(\hat{q}_1, \ldots, \hat{q}_k)$ be the vector of these proportions. The vector of isotonic estimates $(\tilde{q}_1, \ldots, \tilde{q}_K)$ can be obtained from $(\hat{q}_1, \ldots, \hat{q}_K)$ by using the pool adjacent violators algorithm (see, e.g., Ref. 3). At the end of the trial the dose with the value $\bar{q}_i$ closest to $\Gamma$ is the estimated MTD. If there are two or more such doses, the highest dose with the estimated value below $\Gamma$ is chosen. If all the estimated values at these doses are higher than $\Gamma$, the lowest of these doses is chosen. The cumulative cohort decision rule [25] described in Section 1.3.1 when used with isotonic estimates of toxicity rates is an isotonic design. A few other isotonic designs have been proposed, including the isotonic design of Leung and Wang [29]. Ivanova and Flournoy [24] compared several isotonic designs with the cumulative cohort design via simulations for a number of target quantiles and dose–toxicity models and concluded that the cumulative cohort design performs better than others.

### 1.3.4 Bayesian Designs

Parametric methods require assumptions about the model for the dose–toxicity relationship. In addition, Bayesian methods require priors on the model parameters. The continual reassessment method (CRM) is a Bayesian design proposed in 1990 [36]. The CRM starts with a working model for the dose–toxicity relationship. Let $y_i = 1$ if the $i$th patient experiences toxicity and let $y_i = 0$ otherwise, $i = 1, \ldots, n$. For example

$$F(d,\theta) := P\{y_i = 1 \mid d\} = [(\tanh d + 1)/2]^{\theta}. \tag{1.1}$$

The CRM uses Bayes' theorem to update a prior distribution $g(\Sigma)$ of $\Sigma$, for example, $g(\theta) = \exp(-\theta)$. After each patient's response is observed, the mean posterior density of the parameter is computed. Let $x_i \in D$ be the dose received by the $i$th patient. So after the $n$th patient's response, $\Omega_n = \{(x_1, y_1), \ldots, (x_n, y_n)\}$ are the accumulated data and

$$\hat{\theta}^{(n)} = E(\theta \mid \Omega_n) = \int_0^{\infty} \theta f(\theta \mid \Omega_n) d\theta \tag{1.2}$$

is the posterior mean of $\theta$. Here $f(\theta \mid \Omega_n) = L_{\Omega n}(\theta) g(\theta) / \int_0^{\infty} L_{\Omega n}(u) g(u) du$ and $L_{\Omega n}(\theta)$ is the likelihood function.

In the CRM, no prespecified set of doses is required and subjects are assigned one at a time. However, doses can be restricted to a prespecified ordered set $D = \{d_1, \ldots, d_K\}$ [34]. In this case, the model above can also be written as $F(d_i,\theta) = b_i^{\theta}$, where $(b_1, \ldots, b_k)$ is a set of constants, $b_i = (\tanh d_i + 1)/2$.

The first patient receives the first dose level, $x_1 = d_1$. Assume that $n$ patients have been assigned so far. The dose to be administered to the next patient is the dose $x_{n+1}$ such that the absolute difference between $\Pr\{y = 1 \mid x_{n+1}, \hat{\theta}^{(n)}\}$ and $\Gamma$ is minimized. If a prespecified set $D$ is chosen, this quantity is minimized over $D$. Dose $x_{n+1}$ can be used as an estimate of the MTD after $n$ patients have been assigned. Other estimators were explored by O'Quigley [35]. Necessary conditions for the CRM to converge to the target dose were given in Shen and O'Quigley [40], and more relaxed conditions were given by Cheung and Chappell [9]. Also, subjects can be assigned in groups [14,28,18] to shorten the total duration of the trial.

The CRM is a special case of a Bayesian decision procedure with the next dose $x_{n+1}$ selected to maximize the gain function [47]:

$$G(\hat{\theta}^{(n)}, d) = (F(d,\hat{\theta}^{(n)}) - \Gamma)^{-2}. \tag{1.3}$$

Another Bayesian design for dose-finding studies is the escalation with overdose control [2]. This design is from a class of Bayesian feasible designs. It uses a loss function to minimize the predicted amount by which any given patient is overdosed. Bayesian decision procedures for dose-finding studies were described in McLeish and Tosh [32], Whitehead and Brunier [47], and Whitehead and Williamson [48]. Leung and Wang [30] point out that the CRM is a myopic strategy and might not be globally optimal. A globally optimal strategy requires comparison of all possible sets of actions that could be taken, and this remains computationally formidable for designs having more than three dose levels [19].

### 1.3.5 Time-to-Event Design Modifications

If a follow-up time is required for each patient as, for example, in many radiation therapy trials, the dose-finding trial can be impractically long. Cheung and Chappell [8] suggested a modification of the CRM that allows treatment assignments to be staggered so as to shorten the trial duration.

In the original CRM [33], the calculation of the posterior mean of $\theta$ at the time when the $(n + 1)$th patient enters the trial is based on the likelihood

$$L_n(\theta) = \prod_{i=1}^{n} F(x_i, \theta)^{y_i} \{1 - F(x_i, \theta)\}^{1-y_i}, \tag{1.4}$$

where $F(x_i, \theta)$ is a working model as before. Cheung and Chappell [8] introduced the so-called TITE-CRM for trials with long follow-up. They redefined the toxicity rate at dose $d_i$ to be the probability of observing toxicity at $d_i$ during a time period of length $T$ after initiation of therapy. Data for the $i$th patient, $i = 1, \ldots n$, are $\{x_i, y_{i,n}, u_{i,n}\}$ when the $(n + 1)$st patient enters the trial, where $x_i$ is the dose, $y_{i,n}$ is the toxicity indicator, and $u_{i,n}$ is the time that has elapsed from the moment when the $i$th patient entered the trial to the time $(n + 1)$th patient enters the trial.

Cheung and Chappell [8] suggested using a weighted likelihood for TITE-CRM:

$$\tilde{L}_n(\theta) = \prod_{i=1}^{n} \{w_{i,n} F(x_i,\theta)\}^{y_i} \{1 - w_{i,n} F(x_i,\theta)\}^{1-y_i}, \tag{1.5}$$

where $w_{i,n}$ is the weight assigned to the $i$th observation prior to entry of the $(n+1)$th patient. For example, a weight of $w_{i,n} = \min(\mu_{i,n}/T, 1)$ reflects an assumption that the density of time to toxicity is flat in $(0, T)$. Other choices for weights can be considered [8].

Similar modifications can be applied to any treatment allocation rules that are based on the likelihood function. In particular, the isotonic designs can be extended using this idea for trials with long follow-up. Such extension of the cumulative cohort design is described in Ivanova et al. [25].

## 1.4 OTHER DESIGN PROBLEMS IN DOSE FINDING

Below we list various other design problems that arise in the dose-finding context. We have not included designs for bivariate outcomes, but note that dose-finding designs whose goals combine toxicity with efficacy form a growing area of research. Otherwise, we apologize in advance if we have overlooked one of your favorites.

*Ordered Groups* Sometimes patients are stratified into two subpopulations, for example, heavily pretreated and not, where the first subpopulation is more likely to experience toxicity. The goal is to find two MTDs, one for each subpopulation. One of the subpopulations is often very small, rendering the running of two separate trials, one for each subpopulation, unfeasible. O'Quigley and Paoletti [37] proposed a parametric design for this problem. Their method is an extension of the CRM. Ivanova and Wang [22] proposed an isotonic approach where bivariate isotonic regression is used to estimate toxicity rates in both populations simultaneously.

*Multitreatment Trials* Multi-treatment trials are very common. The goal is usually to find the maximum tolerated dose combination. Often only the dose of one agent is varied, with doses of all the other agents held fixed. Thall et al. [45] propose a Bayesian design for trials with two agents in which the doses of both agents are changed simultaneously.

Ivanova and Wang [22] and Wang and Ivanova [46] considered a two-agent trial where two doses of one of the agents, say, the second agent, have already been selected. The problem is to find two maximum tolerated doses of the first agent, one MTD for each dose of the second agent. Ivanova and Wang [22] described an isotonic design, and Wang and Ivanova [46] described a Bayesian design for the problem.

*Ordinal Outcomes* Toxicity in oncology, and many other settings, is measured as an ordinal variable. Bekele and Thall [4] gave an example of a dose-finding trial where different grades of toxicity are combined to obtain a toxicity score for each patient. The goal was to find the dose with a certain weighted sum of probabilities of toxicity grades corresponding to different toxicity types. They [4] suggested a Bayesian design for this problem. Ivanova [21] described a trial where three grades of toxicity (none, mild, and dose-limiting) are combined in a single score. A design in the spirit of the $A + B$ designs to target the dose with the score of 0.5 was used in that trial [21].

Paul et al. [38] considered a different problem in which, target toxicity rates are specified for each grade of toxicity. The goal is to find the vector of doses that have the prespecified rates of toxicity. A multistage random-walk rule with a multidimensional isotonic estimator is proposed.

*Finding a Maximum Tolerated Schedule* In chemotherapy trials treatment is usually administered in cycles. The goal is to find a maximum tolerated schedule for an agent used in chemotherapy administration. Braun, et al. [6] presented a parametric design for this problem.

## 1.5 CONCLUDING REMARKS

We have given an overview of dose-finding designs. There has been much progress in the area of dose-finding designs; new dose-finding problem are being formulated and new methods developed. Statistical methods for dose-finding designs are most advanced for trials in oncology and other life-threatening diseases. Ad hoc designs, such as the $3 + 3$ or $A + B$ designs, are often criticized for being inflexible with regard to their objectives. It is true that $A + B$ designs do not converge to a certain quantile because they invoke the stopping rule and use small sample size. Increasing the size of the second cohort or using an $A + B + C$ design will lead to better performance of these types of design. The major limitation of and $A + B$ design is that no modifications of the design exist to use the design in trials with delayed toxicity outcome. The CRM had been shown to converge to the MTD under certain conditions. It performs very well for small to moderate sample sizes. The CRM can be used for trials with delayed outcomes [8]. However, attempts to design a stopping rule (for use with the CRM) that performs very well have been unsuccessful. Therefore, one needs to specify the total sample size in advance when the CRM is used. A number of publications on isotonic designs or "model free" designs have appeared in the literature and are discussed by Ivanova and Flournoy [24]. These designs do not use any assumption other then toxicity monotonicity. As extension of nonparametric designs, isotonic designs allow using all the data available to obtain toxicity estimates. From the parametric design perspective, model-free designs bring flexibility to modeling when needed.

## REFERENCES

1. Antognini, A. B., Bortot, P., and Giovagnoli, A., Randomized group up and down experiments, *Ann. Ins. Statist. Math.* **58** (in press).
2. Babb, J., Rogatko, A., and Zacks, S., Cancer phase I clinical trials: Efficient dose escalation with overdose control, *Statist. Med.* **17**, 1103–1120 (1998).
3. Barlow, R. E., Bartholomew, D. J., Bremner, J. M., and Brunk, H. D., *Statistical Inference under Order Restrictions*, Wiley, London, New York, 1972.
4. Bekele B. N. and Thall F. T., Dose-finding based on multiple toxicities in a soft tissue sarcoma trial, *J. Am. Statist. Assoc.* **99**, 26–35 (2004).
5. Bortot, P. and Giovagnoli, A., Up-and-down designs of the first and second order, *J. Statist. Plan. Infer.* **134**, 236–253 (2005).
6. Braun, T. M., Thall, P. F., and Yuan, Z., Determining a maximum tolerated schedule of a cytotoxic agent, *Biometrics* **61**, 335–343 (2005).
7. Buoen, C., Bjerrum, O. J., and Thomsen, M. S., How first-time-in-human studies are being performed: A survey of phase I dose-escalation trials in healthy volunteers published between 1995 and 2004, *J. Clin. Pharmacol.* **45**,1123–1136 (2005).

8. Cheung, Y. K. and Chappell, R., Sequential designs for phase I clinical trials with late-onset toxicities, *Biometrics* **56**, 1177–1182 (2000).
9. Cheung, Y. K. and Chappell, R., A simple technique to evaluate model sensitivity in the continual reassessment method, *Biometrics* **58**, 671–674 (2002).
10. Dixon, W. J. and Mood, A. M., A method for obtaining and analyzing sensitivity data, *J. Am. Statist. Assoc.* **43**, 109–126 (1954).
11. Durham, S. D. and Flournoy, N., Random walks for quantile estimation, in *Statistical Decision Theory and Related Topics V*, Berger, J. and Gupta, S., eds., Springer-Verlag, New York, 1994, pp. 467–476.
12. Durham, S. D. and Flournoy, N., Up-and-down designs I: Stationary treatment distributions, in *Adaptive Designs. Lecture-Notes Monograph Series*, Flournoy, N. and Rosenberger, W. F. eds., Institute of Mathematical Statistics, Hayward, CA, Vol. 25, 1995, pp. 139–157.
13. Durham, S. D., Flournoy, N., and Haghighi, A. A., Up-and-down designs II: Exact treatment moments, in *Adaptive Designs. Lecture-Notes Monograph Series*, Flournoy, N. and Rosenberger, W. F., eds., Institute of Mathematical Statistics, Hayward, CA, Vol. 25, 1995, pp. 158–178.
14. Faries, D., Practical modifications of the continual reassessment method for phase I cancer clinical trials, *J. Biopharm. Statist.* **4**, 147–164 (1994).
15. Gezmu, M. and Flournoy, N., Group up-and-down designs for dose finding, *J. Statist. Plan. Infer.* **136**, 1749–1764 (2006).
16. Giovagnoli, A. and Pintacuda, N., Properties of frequency distributions induced by general up-and-down methods for estimating quantiles, *J. Statist. Plan. Infer.* **74**, 51–63 (1998).
17. Girard, P., Laporte-Simitsidis, S., Mismetti, P., Decousus, H., and Boissel, J. P., Influence of confounding factors on designs for dose-effect relationship estimates, *Statist. Med.* **14**, 987–1005 (1995).
18. Goodman, S. N., Zahurak, M. L., and Piantadosi, S., Some practical improvements in the continual reassessment method for phase I studies, *Statist. Med.* **14**, 1149–1161 (1995).
19. Hardwick, J. P., Oehmke, R., and Stout, Q. F., A program for sequential allocation of three Bernoulli populations, *Comput. Statist. Data Anal.* **31**, 397–416 (1999).
20. He, W., Liu, J., Binkowitz, B., and Quan, H., A model-based approach in the estimation of the maximum tolerated dose in phase I cancer clinical trials, *Statist. Med.* **25**, 2027–2042 (2006).
21. Ivanova, A., Escalation, up-and-down and A + B designs for dose-finding trials, *Statist. Med.* **25**, 3668–3678 (2006).
22. Ivanova, A. and Wang, K., A nonparametric approach to the design and analysis of two-dimensional dose-finding trials, *Statist. Med.* **23**, 1861–1870 (2004).
23. Ivanova, A. and Wang, K., Bivariate isotonic design for dose-finding with ordered groups, *Statist. Med.* **25**, 2018–2026 (2006).
24. Ivanova, A. and Flournoy, N., Comparison of isotonic designs for dose-finding, Submitted (in press).
25. Ivanova, A., Flournoy, N., and Chung, Y., Cumulative cohort design for dose-finding, *J. Statist. Plan. Infer.* (in press).
26. Kang, S. H. and Ahn, C., The expected toxicity rate at the maximum tolerated dose in the standard phase I cancer clinical trial design, *Drug Inform. J.* **35**(4), 1189–1200 (2001).

27. Kang, S. H. and Ahn, C., An investigation of the traditional algorithm-based designs for phase I cancer clinical trials, *Drug Inform. J.* **36**, 865–873 (2002).
28. Korn, E. L., Midthune, D., Chen, T. T., Rubinstein, L. V., Christian, M. C., and Simon, R. M., A comparison of two phase I trial designs, *Statist. Med.* **13**, 1799–1806 (1994).
29. Leung, D. H. and Wang, Y. G. Isotonic designs for phase I trials, *Control. Clin. Trials* **22**, 126–138 (2001).
30. Leung, D. H. and Wang, Y. G., An extension of the continual reassessment method using decision theory, *Statist. Med.* **21**, 51–63 (2002).
31. Lin, Y. and Shih, W. J., Statistical properties of the traditional algorithm-based designs for phase I cancer clinical trials, *Biostatistics* **2**, 203–215 (2001).
32. McLeish, D. L. and Tosh, D. Sequential designs in bioassay, *Biometrics* **46**, 103–116 (1990).
33. O'Quigley, J., Pepe, M., and Fisher, L., Continual reassessment method: A practical design for phase I clinical trials in cancer, *Biometrics* **46**, 33–48 (1990).
34. O'Quigley, J. and Chevret, S., Methods for dose-finding studies in cancer clinical trials: A review and results of a Monte Carlo study, *Statist. Med.* **10**, 1647–1664 (1991).
35. O'Quigley, J., Estimating the probability of a toxicity at recommended dose following a phase I clinical trial in cancer, *Biometrics* **48**, 853–862 (1992).
36. O'Quigley, J., Pepe, M., and Fisher, L. Continual reassessment method: A practical design for phase I clinical trials in cancer, *Biometrics* **46**, 33–48 (1990).
37. O'Quigley, J. and Paoletti, X., Continual reassessment method for ordered groups, *Biometrics* **59**, 430–440 (2003).
38. Paul, R. K., Rosenberger, W. F., and Flournoy, N., Quantile estimation following non-parametric phase I clinical trials with ordinal response, *Statist. Med.* **23**, 2483–2495 (2004).
39. Reiner, E., Paoletti, X., and O'Quigley, J., Operating characteristics of the standard phase I clinical trial design, *Comput. Statist. Data Anal.* **30**, 303–315 (1999).
40. Shen, L. Z. and O'Quigley, J., Consistency of continual reassessment method under model misspecification, *Biometrika* **83**, 395–405 (1996).
41. Sheiner, L. B., Beal, S. L., and Sambol, N. C., Study designs for dose-ranging, *Clin. Pharmacol. Ther.* **46**, 63–77 (1989).
42. Simon, R., Friedlin, B., Rubinstein, L., Arbuck, S. G., Collins, J., and Christian, M. C., Accelerated titration designs for phase I clinical trials in oncology, *J. Nat. Cancer Inst.* **89**(15), 1138–1147 (1997).
43. Simpson, D. G. and Margolin, B. H., Recursive nonparametric testing for dose-response relationships subject to downturns at high doses, *Biometrika* **73**, 589–596 (1986).
44. Storer, B. E., Design and analysis of phase I clinical trials, *Biometrics* **45**, 925–937 (1989).
45. Thall, P., Millikan, R., Mueller, P., and Lee, S., Dose finding with two agents in phase I oncology trials, *Biometrics* **59**, 487–496 (2003).
46. Wang, K. and Ivanova, A., Two-dimensional dose finding in discrete dose space, *Biometrics* **61**, 217–222 (2005).
47. Whitehead, J. and Brunier, H., Bayesian decision procedures for dose determining experiments, *Statist. Med.* **14**, 885–893 (disc: p897–899) (1995).
48. Whitehead, J. and Williamson, D. Bayesian decision procedures based on logistic regression models for dose-finding studies, *J. Biopharm. Statist.* **8**, 445–467 (1998).