# Stochastic Approximation and Modern Model-based Designs for Dose-Finding Clinical Trials

**Ying Kuen Cheung[Associate Professor]**
Department of Biostatistics, Columbia University, New York, New York, USA 10032

Ying Kuen Cheung: yc632@columbia.edu

## Abstract

In 1951 Robbins and Monro published the seminal paper on stochastic approximation and made a specific reference to its application to the "estimation of a quantal using response, non-response data". Since the 1990s, statistical methodology for dose-finding studies has grown into an active area of research. The dose-finding problem is at its core a percentile estimation problem and is in line with what the Robbins-Monro method sets out to solve. In this light, it is quite surprising that the dose-finding literature has developed rather independently of the older stochastic approximation literature. The fact that stochastic approximation has seldom been used in actual clinical studies stands in stark contrast with its constant application in engineering and finance. In this article, I explore similarities and differences between the dose-finding and the stochastic approximation literatures. This review also sheds light on the present and future relevance of stochastic approximation to dose-finding clinical trials. Such connections will in turn steer dose-finding methodology on a rigorous course and extend its ability to handle increasingly complex clinical situations.

### Key words and phrases

Coherence; Dichotomized data; Discrete barrier; Ethics; Indifference interval; Maximum likelihood recursion; Unbiasedness; Virtual observations

## 1. Introduction

Dose-finding in phase I clinical trials is typically formulated as estimating a pre-specified percentile of a dose-toxicity curve. That is, the objective is to identify a dose $\theta$ such that $\pi(\theta) = p$, or equivalently,

$$\theta = \pi^{-1}(p)$$

(1)

where $\pi(x)$ is the probability of toxicity at dose $x$ and is assumed continuous and increasing in $x$. Percentile estimation, often seen in bioassay, is a well-studied problem for which statisticians have an extensive set of tools; see Finney (1978) and Morgan (1992). There are, however, two practical aspects of clinical studies that distinguish phase I dose-finding from the classical bioassay problem. First, the experimental units are humans. An implication is that the subjects should be treated sequentially with respect to some ethical constraints (e.g. Section 2.3.1). As such, dose-finding is as much a *design* problem as an analysis problem. Second, the actual doses administered to the subjects are confined to a discrete panel of levels, denoted by $\{d_1, \ldots, d_K\}$, with $\pi(d_1) < \cdots < \pi(d_K)$. Therefore, it is possible that $\pi(d_k) \neq p$ for all $k$, and the working objective then is to identify the dose

$$v = \operatorname*{argmin}_{d_k} |\pi(d_k) - p|. \tag{2}$$

Apparently, the continuous dose-finding objective $\theta$ and the discrete objective $v$ are close to each other. However, in this article, we will see that discretized versions of methods developed for $\theta$ are not necessarily good solutions for $v$. This special section of *Statistical Science* also consists of four other articles that review some benchmarks in the recent development of the so-called model-based methods for dose-finding studies. In a nutshell, a model-based method makes dose decisions based on the explicit use of a dose-toxicity model. That is, the toxicity probability at dose $x$, $\pi(x)$, is postulated to be $F(x, \phi_0)$ for some true parameter value $\phi_0$. This is in contrast to the class of algorithm-based designs whereby a set of dose-escalation rules are pre-specified for any given dose without regard to the observations at the other doses. Section 2 of this article will present a brief history of the development of the modern dose-finding methods and define the scope of this special issue.

In addition, this article complements the other articles in two ways. First, it consolidates the key theoretical dose-finding criteria that are otherwise scattered in the literature (Section 2.3). Second, it compares and contrasts the dose-finding literature with the large literature on stochastic approximation (Section 3); the former primarily addresses the discrete objective $v$, whereas the latter deals with $\theta$. While this literature synthesis is of intellectual interest, it also sheds light on how we may tailor the well-studied stochastic approximation method to meet the practical needs in dose-finding studies (Section 4). Section 5 will end this article with some future directions in dose-finding methodology.

## 2. Modern Dose-Finding Methods

### 2.1 A brief history

This article uses Storer and DeMets (1987) as a historical line to define the modern statistical literature of dose-finding. Little discussion and formal formulation of the dose-finding problem existed in the pre-1987 statistical literature; an exception was Anbar (1984). While dose-finding in cancer trials were discussed as early as in the 1960s in the biomedical communities, a well-defined quantitative objective such as (2) was absent in the communications; see Schneiderman (1965) and Geller (1984) for example. Storer and DeMets (1987) is the earliest reference, to the best of my knowledge, that engages the clinical readership with the idea of percentile estimation. The authors point out the arbitrary estimation properties associated with the traditional 3+3 algorithm used in actual dose-finding studies in cancer patients. The 3+3 algorithm identifies the so-called maximum tolerated dose (MTD) using the following dose escalation rules after enrolling every group of three subjects: let $x_j$ denote the dose given to the $j$th group of subjects and suppose $x_j = d_k$, then

$$x_{j+1} = \begin{cases} d_{k+1} & \text{if } z_k/n_k < 0.33, \\ d_k & \text{if } z_k = 1 \text{ and } n_k = 3, \\ d_{k-1} & \text{if } z_k \geq 2, \end{cases} \tag{3}$$

where $n_k$ and $z_k$ respectively denote the *cumulative* sample size and number of toxicities at dose $d_k$. The trial will be terminated once a de-escalation occurs, and the next lower dose will be called the MTD. In the sequel, Storer (1989) deduces from the 3+3 algorithm (3) that a cancer dose-finding study aims to estimate the 33rd percentile (i.e., $p = .33$). While it has now emerged that the target is likely lower than the 33rd percentile with $p$ being between .16

and .25, their work has shaped the subsequent development of dose-finding methods in both the statistical and biomedical literatures, and the MTD has since been defined invariably as a dose associated with a pre-specified toxicity probability $p$.

O'Quigley and colleagues proposed the continual reassessment method (CRM) in 1990. The CRM is the first model-based method in the modern dose-finding literature. The main idea of the method is to treat the next subject or group of subjects at the dose with toxicity probability estimated to be closest to the target $p$. Precisely, suppose we have observations from the first $j$ groups of subjects and compute the posterior mean $\hat{\phi}_j$ of $\phi$ given these observations. Then the next group of subjects will be treated at

$$x_{j+1} = \arg \min_{d_k} |F(d_k, \widehat{\varphi}_j) - p|.$$

(4)

A similar idea is adopted in most model-based designs proposed since 1990. One example is the escalation with overdose control (EWOC) by Babb et al. (1998), who apply the continual reassessment notion but estimate the MTD with respect to an asymmetric loss function which places heavier penalties on over-dosing than under-dosing. O'Quigley and Conaway (2010) and Tighiouart and Rogatko (2010) in this special issue review the CRM and the EWOC and their respective extensions. Another CRM-like design is the curve-free method by Gasparini and Eisele (2000) who estimate the dose-toxicity curve using a Bayesian nonparametric method in an attempt to avoid bias due to model misspecification. Leung and Wang (2001) propose an analogous frequentist version that uses isotonic regression for estimation. Other model-based designs include the Bayesian decision-theoretic design (Whitehead and Brunier, 1995), the logistic dose-ranging strategy (Murphy and Hall, 1997), and Bayesian $c$-optimal design (Haines et al., 2003).

The late 1990s saw an increasing interest in algorithm-based designs. Durham et al. (1997) propose a biased coin design by which the dose for the next subject is reduced if the current subject has a toxic outcome, and the dose is escalated with a probability $p/(1 - p)$ otherwise. The biased coin design is a randomized version of the Dixon and Mood (1948) up-and-down design. Motivated by its similarity to the traditional design, Cheung (2007) studies a class of stepwise procedures that includes (3) as a special case. Yet another algorithm-based method is proposed by Ji et al. (2007) who make interim decisions based on the posterior toxicity probability interval associated with each dose. The impetus for these algorithm-based designs is simplicity: the decision rules can be charted prior to the trial, so that the clinical investigators know exactly how doses will be assigned based on the observed outcomes.

In order to make dose-finding techniques relevant to clinical practice, statisticians have responded to the realistically complicated clinical situations such as time-to-toxicity endpoints (Cheung and Chappell, 2000) and combination treatments (Thall et al., 2003). While the core dose-finding objective remains a percentile estimation problem, the complexity of dose-finding methods has grown rapidly in the literature, with most innovations taking the model-based approach. Thall (2010) in this special issue will review the major development of these complex designs.

Most (model-based) designs in the literature up to this point take the myopic approach by which the dose assignment is optimized with respect to the next immediate subject without regard to the future subjects. Bartroff and Lai (2010) in this issue break away from this direction and propose a model-based method from an adaptive control perspective. While this work attempts to solve a specific Bayesian optimization problem, it also sets a new direction in the modern dose-finding techniques; see Section 3 of this article.

## 2.2 Why model-based now

A model-based design allows borrowing strength from information across doses. This characteristic appeals to statisticians and clinicians alike, especially because of the typically small-to-moderate sample sizes seen in early-phase clinical studies. As clinicians begin to appreciate the crucial role of dose-finding in the entire drug development program and the value of statistical inputs to reconcile the ethical and research aspects in early phase trials, their discussions have revolved around model-based innovations such as the CRM (Ratain et al., 1993) and the EWOC (Eisenhauer et al., 2000). The increasing number of applications in actual trials (Muller et al., 2004) indicates the clinical awareness and readiness for these model-based methods.

When compared to the simplicity of algorithm-based methods, the model-based approaches are computationally complex and require special programming before and during the implementation of a trial. Thanks to the advances of computing algorithms (e.g., Markov Chain Monte Carlo) and computer technology however, trial planning with extensive simulation has become feasible. This being the case, a full-scale dynamic programming can still stretch the computing resource; see Bartroff and Lai (2010) for some comparison of computational times. In addition, statistician-friendly software has become increasingly available for the planning and execution of these model-based designs, e.g. the dfcrm package in R (Cheung, 2008). These indicators of computational maturity transform the model-based designs into practical tools for dose-finding trials.

Finally, the development of dose-finding theory and dose-response models in the past two decades lends scientific rigor to the complexity of the model-based methods. Indeed, the goal of this special issue is to review the theoretical and modeling progress made in the modern dose-finding literature, and thereby demonstrate the full promise, and perhaps challenges, of the model-based methods.

## 2.3 Some theoretical criteria

In a typical dose-finding trial, subjects are enrolled in small groups of size $m \geq 1$. The enrollment plan is said to be fully sequential when $m = 1$. Let $x_i$ denote the dose given to the $i$th group of subject(s). Thus, the sequence $\{x_i\}$ forms the design of a dose-finding study. As most dose-finding methods are outcome-adaptive, each design point $x_i$ is random and depends on the previous observation history. Evaluation of a dose-finding method therefore involves the study of its design space with respect to some ethical and estimation criteria. This section will review some key dose-finding criteria including coherence, rigidity, indifference intervals, and unbiasedness.

**2.3.1 Coherence—**First, consider fully sequential trials with $m = 1$, so that each human subject is an experimental unit. An ethical principle, coined coherence by Cheung (2005), dictates that no escalation should take place for the next enrolled subject if the current subject experiences some toxicity, and that dose reduction for the next subject is not appropriate if the current subject has no sign of toxicity. Precisely, let $Y_i$ denote the toxicity outcome of the $i$th subject. An escalation for the subject is said to be coherent only when $Y_{i-1} = 0$; likewise, a de-escalation is coherent only when $Y_{i-1} = 1$. Extending the notion of coherence for each move, one can naturally define coherence as a property of a dose-finding method:

PROPERTY 1 (Coherence). A dose-finding design $\mathcal{D}$ is said to be coherent in escalation if with probability one

$$P_{\mathcal{D}}(U_i>0|Y_{i-1}=1)=0 \tag{5}$$

for all $i$, where $U_i = x_i - x_{i-1}$ is the dose increment from subject $i-1$ to $i$, and $P_{\mathcal{D}}(\cdot)$ denotes probability computed under the design $\mathcal{D}$. Analogously, the design is said to be coherent in de-escalation if with probability one

$$P_{\mathcal{D}}(U_i<0|Y_{i-1}=0)=0 \tag{6}$$

for all $i$.

It is important to note that coherence is motivated by ethical concerns, and hence may not correspond to efficient estimation of the dose-toxicity curve. For example, in bioassay, an efficient design obtained by sequentially maximizing some function of the information may induce incoherent moves, and thus is not appropriate for human trials; see McLeish and Tosh (1990) for example.

An algorithm-based design can explicitly incorporate dose decision rules that respect the coherence principles; cf. the biased coin design. For a model-based design, on the other hand, it is not immediately clear that whether coherence necessarily holds. There are three general ways to ensure coherence in practice. First, one could adopt model-based methods that have been proven coherent analytically. This includes the one-stage Bayesian CRM. Second, one could take a numerical approach. Let $N$ denote the sample size of a trial. Then the design space is completely generated by the first $N-1$ binary toxicity observations, and thus consists of $2^{N-1}$ possible design outcomes. Therefore, one could establish coherence (for a given $N$) by enumerating all possible outcomes and verifying that there is no incoherent move. In some cases, the number of outcomes can be immensely reduced to the order of $N$; see Theorem 1 in Cheung (2005). Third, one could enforce coherence by restriction when the model-based dose assignment is incoherent. Applying coherence restrictions is common in practice (Faries, 1994) and is the most straightforward approach for complex designs. On the other hand, the restricted moves need to be examined carefully lest they should cause an incompatibility problem as defined in Cheung (2005).

In practice, the enrollment plan is often small-group sequential, i.e., $m > 1$, in order to reduce the number of interim decisions and hence trial duration. In this case, each group of subjects may be viewed as an experimental unit. A generalized version of Property 1 can be stated as:

PROPERTY 1′ (Group coherence). A dose-finding design $\mathcal{D}$ is said to be group coherent in escalation if with probability one

$$P_{\mathcal{D}}(U_i>0|\overline{Y}_{i-1} \geq p)=0 \tag{7}$$

for all $i$, where $U_i = x_i - x_{i-1}$ now denotes the dose increment from group $i-1$ to $i$ and $\overline{Y}_{i-1}$ is the observed proportion of toxicities in group $i-1$. Analogously, the design is said to be group coherent in de-escalation if with probability one

$$P_{\mathcal{D}}(U_i < 0 | \overline{Y}_{i-1} \leq p) = 0 \tag{8}$$

for all $i$.

It is easy to see that (7) and (8) reduce to (5) and (6) respectively when $m = 1$ for $p \in (0, 1)$.

**2.3.2 Rigidity and sensitivity**—A design sequence $\{x_n\}$ is strongly consistent for $\theta$ if $x_n \to \theta$ with probability one. For trials allowing only a discrete number of test doses as in (2), consistency means $x_n = v$ eventually with probability one. Consistency, a desirable statistical property in general, has an ethical connotation in dose-finding studies because it implies all subjects enrolled after a certain time point will be treated at $v$, which is the desired dose.

PROPERTY 2 (Rigidity). A dose-finding design $\mathcal{D}$ is said to be rigid if for every $0 < p_L < \pi(v) < p_U < 1$ and all $n \geq 1$,

$$P_{\mathcal{D}}\{x_n \in \mathcal{I}_\pi(p_L, p_U)\} < 1,$$

where $\mathcal{I}_\pi(p_L, p_U) = \{x : p_L \leq \pi(x) \leq p_U\}$.

It is easy to see that consistency excludes the rigidity problem. In other words, Property 2 implies that a design is inconsistent. In particular, rigidity occurs when a CRM-like procedure is applied in conjunction with nonparametric estimation. Hence, such a nonparametric design is inconsistent. This is quite interesting and somewhat counter-intuitive, because nonparametric estimation is introduced with an intention to remove bias and to enhance the prospect of consistency.

To illustrate, consider a design that starts at dose level 1, enrolls subjects in groups of size $m = 2$, and assigns the next group at $\arg\min_k |\tilde{p}_k - p|$ where $\tilde{p}_k$ is an estimate of $\pi(d_k)$ based on isotonic regression, and the target is $p = .20$. Now suppose that none of the subjects in the first group has a toxic outcome. Then suppose the second group enters the trial at dose level 2, with one of the two experiencing toxicity. Based on these observations, the isotonic estimates are $\tilde{p}_1 = .00$ and $\tilde{p}_2 = .50$, which bring the trial back to dose level 1. From this point on, because there is no parametric extrapolation to affect the estimation of $\pi(d_2)$ by the data collected at $d_1$, the isotonic estimate $\tilde{p}_2$ will be no smaller than .50 regardless of what happens at $d_1$, i.e., $|\tilde{p}_2 - .20| \geq .30$. As a result, $|\tilde{p}_1 - .20| < .30 \leq |\tilde{p}_2 - .20|$ if $\tilde{p}_1 \leq .20$. That is, the trial will stay at dose level 1 even if there is a long string of non-toxic outcomes there!

This example demonstrates that nonparametric estimation and the sequential sampling plan together cause rigidity through an "extreme" over-estimate of $\pi(d_2)$ based on small sample size. The probability of this extreme over-estimation is non-negligible indeed: if dose level 2 is the true MTD with $\pi(d_2) = .20$, then the probability that the trial is confined to the suboptimal dose 1 is at least .36 by a simple binomial calculation. Cheung (2002) constructs a similar numerical example for the Bayesian nonparametric curve-free method, and suggests that the rigidity probability can be reduced by using an informative prior to add smoothness to the estimation.

Due to ethical constraints such as coherence and the discrete design space, it may be challenging to achieve consistency without strong model assumptions. For example, the CRM has been shown to be consistent under certain model misspecifications, but is not generally so (Shen and O'Quigley, 1996). In this context, Cheung and Chappell (2002)

introduced the indifference interval as a sensitivity measure of how close a design may approach $v$ on the probability scale:

PROPERTY 3 (Indifference interval). The indifference interval of a dose-finding design $\mathcal{D}$ exists and is equal to $p \pm \delta$ if there exist $N > 0$ and $\delta \in (0, p)$ such that

$$P_{\mathcal{D}} \{x_n \in \mathscr{I}_\pi(p - \delta, p+\delta) \, for \, all \, n \geq N\} = 1.$$

Apparently, the smaller the half-width $\delta$ of a design's indifference interval is, the closer the design converges to the MTD; whereas a large $\delta$ indicates the design is sensitive to the underlying $\pi$. The sensitivity of the design $\mathcal{D}$ can thus be measured by $\delta$. Specifically, a design with half-width $\delta$ (for some $\delta < p$) will be called a $\delta$-sensitive design.

It is clear that if a design $\mathcal{D}$ is consistent for $v$, then it is $\delta$-sensitive; i.e., one may choose $\delta$ so that $\pi(v) \in p \pm \delta$. Also, if $\mathcal{D}$ is $\delta$-sensitive, then it is non-rigid. Thus, while consistency appears to be too difficult and non-rigidity too non-discriminatory for a dose-finding design, $\delta$-sensitivity seems to be a reasonable design property. Cheung and Chappell (2002) prescribed a way to calculate the indifference interval of the CRM, i.e., the CRM is $\delta$-sensitive. Moreover, Lee and Cheung (2009) showed that the CRM can be calibrated to achieve any $\delta$ level of sensitivity. However, it should be noted that indifference interval is an asymptotic criteria. As such, a small $\delta$ does not necessarily yield good finite sample properties.

**2.3.3 Unbiasedness—**The performance of a reasonable dose-finding design is expected to improve as the underlying dose-toxicity curve $\pi$ becomes steep. This property, called unbiasedness by Cheung (2007), is formulated as follows:

PROPERTY 4 (Unbiasedness). Let $p_i = \pi(d_i)$ denote the true toxicity probability at dose $d_i$. A design $\mathcal{D}$ is said to be unbiased if

   **a.**  $P_{\mathcal{D}} (x_n = d_k)$ is nonincreasing in $p_{i'}$ for $i' \leq k$, and

   **b.**  $P_{\mathcal{D}} (x_n = d_k)$ is nondecreasing in $p_i$ for $i > k$.

For the special case with $d_k = v$ and $\pi(v) = p$, unbiasedness implies that the probability of correctly selecting $v$ increases as the doses above the MTD become more toxic (i.e., $p_i \gg p$), or the doses below less toxic (i.e., $p_{i'} \ll p$). In other words, the design will select the true MTD more often as it becomes more separated from its neighboring doses in terms of toxicity probability. A design that satisfies this special case is called *weakly unbiased*.

One may argue that $\delta$-sensitive designs (e.g. the CRM) are *asymptotically weakly unbiased*, in that they will be consistent if the underlying dose-toxicity curve $\pi$ becomes sufficiently steep around the MTD; see Figure 1 for an illustration. Unbiasedness has been established only for few designs in the dose-finding literature; an example is the class of stepwise procedures (Cheung, 2007). In practice, extensive simulations are usually required, and are often adequate, to confirm that a design is (weakly) unbiased.

## 3. Stochastic Approximation

### 3.1 The Robbins-Monro (1951) procedure

Robbins and Monro (1951) introduce the first formal stochastic approximation procedure for the problem of finding the root of a regression function. Precisely, let $M(x)$ be the mean of

an outcome variable $Y = Y(x)$ at level $x$, and suppose $M(x) = \alpha$ has a unique root $\theta$ and $\sup_x E\{Y^2(x)\} < \infty$. Then the stochastic approximation recursion approaches $\theta$ sequentially:

$$x_{i+1} = x_i - \frac{1}{ib}(Y_i - \alpha) \tag{9}$$

for some constant $b > 0$. It is well established that $x_n \to \theta$ with probability one. If in addition, the constant $b$ is chosen properly (namely $b < 2M'(\theta) \equiv 2\beta$), then $n^{1/2}(x_n - \theta)$ will converge in distribution to a normal variable with mean 0 and variance $\sigma^2\{b(2\beta - b)\}^{-1}$ where $\sigma^2 = \lim_{x \to \theta} \text{var}\{Y(x)\}$; see Sacks (1958) and Wasan (1969).

It is immediately clear that (9) is applicable to address objective (1) in a clinical trial setting with $M = \pi$ and $\alpha = p$. For one thing, the recursion output is coherent (Property 1) thus passing the first ethical litmus test. It is also easy to see that a small-group sequential version of (9), i.e., replace $Y_i$ with $\overline{Y}_i$, is group coherent (Property 1′). There are, however, several practical considerations.

The choice of $b$ is crucial. In view of efficiency, the asymptotic variance is minimized when we set $b = \beta$, which is typically unknown in most applications. This leads to the idea of adaptive stochastic approximation where $b$ is replaced by a sequence $b_i$ that is strongly consistent for $\beta$ (Lai and Robbins, 1979). However, when the sample size is small-to-moderate, the numerical instability induced by the adaptive choice $b_i$ may offset its asymptotic advantage. In this article, for a reason described in Section 4.2, we assume that a good choice of $b$ is available.

The next practical issue is that (9) entails the availability of a continuum of doses. This is seldom feasible in practice. In drug trials, dose availability is often limited by the dosage of a tablet. For treatments involving combination of drugs administered multiple times over a fixed period, each subsequent dose may involve increasing doses and/or frequency of different drugs. For example, Table 1 describes the dose schedules of bortezomib used in a dose-finding trial in patients with lymphoma (Leonard et al., 2005). The first three levels prescribe bortezomib at a fixed dose 0.7 mg/m$^2$ with increasing frequency, whereas the next two increments apply the same frequency with increasing bortezomib doses. While we are certain that the risk for toxicity increases over each level, there is no natural scale of dosage (e.g. mg/m$^2$). Thus, assuming that the toxicity probability $\pi(x)$ is well-defined on a continuous range of $x$ is artificial.

To tailor the stochastic approximation for the discrete objective $v$, an obvious approach is to round the output of (9) to its closest dose at each iteration. For example, suppose that the dose labels are $\{1,\ldots,K\}$, i.e., $d_k = k$. Then a discretized stochastic approximation may be expressed as

$$x_{i+1} = C\left\{x_i - \frac{1}{ib}(\overline{Y}_i - p)\right\} \tag{10}$$

where $C(x)$ is the rounded value of $x$ if $.5 \leq x < K + .5$, and is set equal to 1 and $K$ respectively if $x < .5$ or $> K + .5$. Unfortunately, the discretized stochastic approximation is rigid (Property 2). To illustrate, consider applying (10) with $b = .2$ and a target $p = .20$ in a trial with $x_1 = 1$ and $m = 2$. Then no toxicity event in the first group, i.e., $\overline{Y}_1 = 0$, gives $x_2 = 2$. Further suppose that the second group has a 50% toxicity rate ($\overline{Y}_2 = 0.5$). This will bring

the trial back to $x_3 = C(2 - 0.75) = 1$; it is easy to see that the remaining subjects will receive dose 1. To see how rigidity occurs for a general variable type, we observe that since $x_i$ is an integer, the update $x_{i+1}$ according to (10) will stay the same as $x_i$ if $|(ib)^{-1}(\overline{Y}_i - p)| <$ 0.5, whose probability approaches 1 at a rate of $O(i^{-2})$ according to Chebyshev's inequality if $Y_i$ has a finite variance. If $Y_i$ is bounded (e.g., binary), the term $C\{(ib)^{-1}(\overline{Y}_i - p)\}$ will always be zero as $i$ becomes sufficiently large, and will not contribute to future updates. This problem, called *discrete barrier*, is thus built by rounding and the fact that the design points take on a discrete set of levels. In the context of the CRM, Shen and O'Quigley (1996) point out similar difficulties in establishing the theoretical properties of dose-finding methods due to the discrete barrier. This is where the modern dose-finding literature departs from the elegant stochastic approximation approach.

### 3.2 Stochastic approximation and model-based methods

The Robbins-Monro stochastic approximation is a nonparametric procedure in that the convergence results depend only very weakly on the true underlying $M(x)$. For the case of normal $Y$, interestingly Lai and Robbins (1979) show that the recursion output in (9) is identical to the solution $\tilde{x}_{i+1}$ of

$$\sum_{j=1}^{i} Y_j - \{\alpha + b(\tilde{x}_j - \tilde{x}_{i+1})\} = 0$$

(11)

which amounts to maximum likelihood estimation of $\theta$ under a simple linear regression model. This connection between the stochastic approximation and a model-based approach motivates the study of the maximum likelihood recursion in Wu (1985), Wu (1986), and Ying and Wu (1997) for data arising from the exponential family. In particular, for binary $Y$, Wu (1985) proposes the logit-MLE that uses the logistic working model

$$F(x, \theta) = \frac{p \exp\{\widetilde{b}(x - \theta)\}}{1 - p + p \exp\{\widetilde{b}(x - \theta)\}} \text{ for some } \widetilde{b} > 0$$

(12)

and replaces the estimating equation (11) with $\sum_{j=1}^{i}\{Y_j - F(\tilde{x}_j, \tilde{x}_{i+1})\} = 0$. Here, we focus on the non-adaptive version, i.e., where $\tilde{b}$ is a fixed constant. A maximum likelihood version of the CRM (4) would clearly yield the same design point as $\tilde{x}_{i+1}$ if the design space was continuous. In this regard, the likelihood CRM is an analogue of the logit-MLE for the discrete objective $v$.

In order to establish the asymptotic distribution of the logit-MLE (and the maximum likelihood recursion in general), Ying and Wu (1997) show that the sequence $\tilde{x}_{i+1}$ is asymptotically equivalent to an adaptive Robbins-Monro recursion; see the proof of Theorem 3 in Ying and Wu (1997). While the justification of the model-based logit-MLE relies on its asymptotic equivalence to the non-parametric Robbins-Monro procedure, Wu (1985) showed by simulation that the former is superior to the latter in finite-sample settings with binary data. Similarly, O'Quigley and Chevret (1991) demonstrated that the CRM performs better than the discretized stochastic approximation (10) for the objective $v$.

These observations regarding the stochastic approximation, the logit-MLE, and the CRM bear two practical suggestions. First, in typical dose-finding trial settings with binary data

and small sample sizes, a model-based approach seems to retain some information that is otherwise lost when using nonparametric procedures. This speculation is made without assuming much confidence about the working model. Second, one may study the theoretical (asymptotic) properties of the modern model-based method (e.g., CRM) by tapping the rich stochastic approximation literature, thus giving guidance on the choice of design parameters such as $\tilde{b}$ in (12). This can be achieved, of course, only if we can resolve the discrete barrier —we will return to this in Section 4.2.

### 3.3 Stochastic approximation and adaptive control

Maximum likelihood recursion attempts to optimize the prospect for the next subject by setting the next design point at the current estimate of $\theta$, and is myopic in that it does not consider the dose assignments of future subjects. The Robbins-Monro procedure is therefore myopic by (asymptotic) equivalence. Lai and Robbins (1979) study the adaptive cost control aspect of the stochastic approximation for normal $Y$ where $C_n = \sum_{i=1}^{n}(x_i - \theta)^2$ is defined as the cost of a design sequence $\{x_i\}$ at stage $n$. Specifically they show that the cost of (9) is of the order $\sigma^2 \log n$ if $b < 2\beta$. Under some simple linear regression models, Han et al. (2006) show that the myopic Bayesian rule is optimal when the slope parameter is known. This suggests that the myopic Robbins-Monro method may also have good adaptive control properties.

The control aspect of the stochastic approximation is less clear for binary data. Bartroff and Lai (2010) address the control problem by using techniques in approximate dynamic programming to minimize some well-defined global risk, such as the expectation of the design cost $C_n$. The authors demonstrate reduction of the global risk by non-myopic approaches when compared to the myopic ones including the stochastic approximation and the logit-MLE. The scope of the simulations, however, is confined to situations where the logistic model correctly specifies $\pi$. In addition, their approach is intended for the continuous objective $\theta$, instead of $\nu$.

Further research on the use of non-myopic approaches in dose-finding is warranted, especially for practical situations with a discrete set of test doses. The design cost at stage $n$ for the discrete objective $\nu$ can be analogously defined as $C_n' = \sum_{i=1}^{n}(x_i - \nu)^2$. Then a dose-finding design $\mathcal{D}$ is consistent if and only if $\lim_{n \to \infty} C_n'$ is finite almost everywhere. As mentioned earlier, the myopic CRM is not necessarily consistent (as it tries to treat each subject at the current "best" dose). By contrast, designs that spread out the design points (e.g., the biased coin design) allow consistent estimation of $\nu$ at the expense of the enrolled subjects. Neither guarantees a finite $\lim_n C_n'$. An optimal $\mathcal{D}$ for the infinite-horizon control of $C_n'$ thus seems to resolve the inherent tension between the welfare of enrolled subjects (i.e., the cost is kept low) and the estimation of $\nu$ (i.e., $x_n$ is consistent).

## 4. Ongoing Relevance

### 4.1 Binary versus dichotomized data

As mentioned above, with a binary outcome and small samples, the Robbins-Monro procedure is generally less efficient than model-based methods, and hence may not be suitable for clinical dose-finding where the study endpoint is classified as toxic and non-toxic. In many situations, however, the binary toxic outcome $T$ is defined by dichotomizing an observable biomarker expression $Y$, namely, $T = \mathbf{1}(Y > t_0)$ for some fixed safety threshold $t_0$, where $\mathbf{1}(E)$ denotes the indicator of the event $E$. The biomarker $Y$ apparently contains

more information than the dichotomized $T$, and may be used to achieve the dose-finding objective (1) with greater efficiency.

To illustrate, consider the regression model

$$Y = M(x) + \sigma(x)\varepsilon \tag{13}$$

where $\varepsilon$ has a known distribution $G$ with mean 0 and variance 1. Under (13), the toxicity probability can be expressed as $\pi(x) = 1 - G\left[\{t_0 - M(x)\}/\sigma(x)\right]$ and the continuous dose-finding objective (1) can be shown to be equivalent to the solution to

$$f(x) \equiv M(x) + z_p\sigma(x) = t_0 \tag{14}$$

where $z_p$ is the upper $p$th percentile of $G$. To focus on the comparison between the use of $Y$ and $T$, suppose for the moment that a continuum of dose $x$ is available. Further suppose that a trial enrolls patients in small groups of size $m$. Let $x_i$ denote the dose given to the $i$th group, and $Y_{ij}$ the biomarker expression of the $j$th subject in the group. With this experimental setup, we note that

$$O_i = \overline{Y}_i + \left[E\{S_i/\sigma(x_i)\}\right]^{-1} z_p S_i \tag{15}$$

is an unbiased realization of $f(x_i)$, where $S_i$ is the sample standard deviation of the observations in group $i$. The expectation in (15) can be computed for any given $G$, because $S_i/\sigma(x_i)$ depends on the error variable $\varepsilon$ but not $M$ and $\sigma$ under model (13). In other words, $O_i$ is observable and is a continuous variable that can be used to generate a stochastic approximation recursion

$$x_{i+1} = x_i - (ib)^{-1}(O_i - t_0). \tag{16}$$

The design $\{x_n\}$ generated by (16) is consistent for $\theta$ under the condition that $\theta$ is the unique solution to (14). This condition holds, for example, when $M$ is strictly increasing and $\sigma$ is nondecreasing in $x$. This is a reasonable assumption for many biological measurements, for which the variability typically increases with the mean. Furthermore, if $b < 2\beta$, where $\beta = f'(\theta)$ here, then the asymptotic variance of $x_n$ is $v_O = \lim_{x \to \theta} \mathrm{var}(O_i)\{b(2\beta - b)\}^{-1}$. In particular, when $\varepsilon$ is standard normal,

$$v_o = \frac{\sigma^2(\theta)\{1 + mz_p^2(\lambda_m - 1)\}}{mb(2\beta - b)},$$

where

$$\lambda_m = \frac{(m-1)\Gamma^2\{(m-1)/2\}}{2\Gamma^2(m/2)}.$$

Now, instead of using the recursion (16), suppose that we apply the logit-MLE based on the dichotomized outcomes by solving $\sum_{j=1}^{i}\{\overline{T}_j - F(\tilde{x}_j, \tilde{x}_{i+1})\}=0$ where $F$ is defined in (12). Then using the results in Ying and Wu (1997), we can show that $\sqrt{n}(\tilde{x}_n - \theta)$ converges in distribution to a mean zero normal with variance $\upsilon_T = p(1-p)\{mb(2\tilde{\beta} - b)\}^{-1}$ where $\tilde{\beta} \equiv \pi'(\theta) = \beta G'(z_p)/\sigma(\theta)$.

The asymptotic variances of $\upsilon_O$ and $\upsilon_T$ are minimized when $b = \beta$ and $\tilde{b} = \tilde{\beta}$ respectively. Thus, the optimal choice depends on unknown parameters. For the purpose of comparing efficiencies, suppose we could set $b$ and $\tilde{b}$ to their respective optimal values. Then the variance ratio is equal to

$$\frac{\upsilon_T}{\upsilon_O} = \frac{p(1-p)}{\{G'(z_p)\}^2\{1 + mz_p^2(\lambda_m - 1)\}} \tag{17}$$

for normal noise, and also represents the asymptotic efficiency of $x_n$ relative to $\tilde{x}_n$. For $m = 3$, the ratio (17) attains a minimum of 1.238 when $p = 0.12$ or $0.88$. As shown in Figure 2, the efficiency gain can be substantial for any group sizes larger than 2, especially when the target $p$ is extreme.

## 4.2 Virtual observations

A particular obstacle to the use of stochastic approximation is the discrete design space used in clinical studies, which creates the discrete barrier (Section 3.1). To overcome the discrete barrier, Cheung and Elkind (2010) introduce the notion of virtual observations. Precisely, the virtual observation of the $i$th group of subjects is defined as

$$V_i = O_i + b(x_i^* - x_i) \tag{18}$$

where $x_i^*$ denotes the *assigned* dose of the group which can take values on a continuous conceptual scale that represents an ordering of doses. In the situations where the *actual given* dose $x_i$ can take on any real value, we have $x_i^* \equiv x_i$ and $V_i \equiv O_i$, and thus, the recursion (16) may be used to approach the target dose $\theta$. When $x_i$ is confined to $\{1, \ldots, K\}$, Cheung and Elkind (2010) propose generating a stochastic approximation recursion based on the virtual observations:

$$x_{i+1}^* = x_i^* - \frac{1}{ib}(V_i - t_0), \tag{19}$$

and treating the next group of subjects at $x_{i+1} = C(x_{i+1}^*)$. To initiate the virtual observation recursion, one may set $x_1^* = x_1 \in \{1, \ldots, K\}$.

Cheung and Elkind (2010) prove, under mild conditions, that $x_{i+1}^*$ generated by (19) is consistent (hence non-rigid) for $\theta_b$ for some $\theta_b = v \pm 0.5$, and hence $x_{i+1}$ for $v$. Briefly, for any given $b$, consistency will occur if the neighboring doses of the MTD are sufficiently apart from the MTD in terms of toxicity probability. This is in essence asymptotically weakly unbiased as defined in Section 2.3.3, and can be easily derived from Propositions 2 and 3 of Cheung and Elkind (2010).

With the use of continuous $V$'s, the notion of coherence needs to be reexamined. In particular, the virtual observation recursion (19) will de-escalate if the biomarker expression of the current subjects has a high average ($\overline{Y}_i$) or a large variability ($S_i$). This is a sensible dose-escalation principle for situations where the variability increases proportionally to the mean.

The idea of virtual observation is to create an objective function

$$h(x) = E(V_i | x_i^* = x) = f\{C(x)\} + b\{x - C(x)\}$$

that is defined on the real line, and has a local slope at $\{1, \ldots, K\}$, such that the solution $\theta$ of (14) can be approximated by the solution $\theta_b$ of $h(x) = t_0$. Quite importantly, since now the objective function $h$ has a known slope $b$ around $\theta_b$ (under some Lipschitz-type regularity conditions), we can use the same $b$ in the recursion (19) as in the definition of virtual observations (18). This design feature enables us to achieve optimal asymptotic variance without resorting to adaptive estimation of the slope of the objective function. It is particularly relevant to early phase dose-finding studies where adaptive stochastic approximation can be unstable due to small sample sizes.

## 5. Looking to the Future

Statistical methodology for dose-finding trials is by its nature an application-oriented discipline. Consequently, much of the emphasis in the dose-finding literature has been on empirical properties via simulation. However, as the (model-based) methods become increasingly complicated, it is imperative to check their properties against some theoretical criteria so as to avoid pathological behaviors that may not be detected in aggregate via simulations; rather, pathologies such as incoherence and rigidity are *pointwise* properties that can be found by careful analytical study. As a case in point, the virtual observation recursion (19) is presented in light of the properties described in Section 2.3. Granted, as the data content becomes richer, these theoretical criteria have to be re-examined. Cheung (2010), in another instance, extends the notion of coherence for bivariate dose-finding in the context of phase I/II trials—see Thall (2010) for a review of the bivariate dose-finding objective—and shows how coherence can be used to simplify dose decisions in the complex "black-box" approach of the bivariate model-based methods, and to provide clinically sensible rules.

The idea of virtual observation bridges the stochastic approximation and the modern (model-based) dose-finding literatures. As the Robbins-Monro method has motivated a large number of extensions and refinements for a wide variety of root-finding objectives, there exists a reservoir of ideas from which we can borrow and apply to dose-finding methods for specialized clinical situations. To name a few, consult Kiefer and Wolfowitz (1952) for finding the maximum of a regression function, and Blum (1954) for multivariate contour-finding. While studying the analytical properties of model-based designs in these specialized situations can be difficult, connection to the theory-rich stochastic approximation procedures allows us to do so with relative ease and elegance, as is the case for the virtual observation

recursion (19). In this light, extending the idea of virtual observations for data types other than continuous and multivariate data appears to be a promising "crosswalk" that warrants further research.

## Acknowledgments

## References

Anbar D. Stochastic approximation methods and their use in bioassay and phase I clinical trials. Communications in Statistics 1984;13(19):2451–2467.

Babb J, Rogatko A, Zacks S. Cancer phase I clinical trials: efficient dose escalation with overdose control. Statistics in Medicine 1998;17(10):1103–1120. [PubMed: 9618772]

Bartroff J, Lai TL. Approximate dynamic programming and its applications to the design of phase I cancer trials. Statistical Science. 2010

Blum JR. Multidimensional stochastic approximation methods. Annals of Mathematical Statistics 1954;25:737–744.

Cheung YK. On the use of nonparametric curves in phase I trials with low toxicity tolerance. Biometrics 2002;58(1):237–240. [PubMed: 11890321]

Cheung YK. Coherence principles in dose-finding studies. Biometrika 2005;92(4):863–873.

Cheung YK. Sequential implementation of stepwise procedures for identifying the maximum tolerated dose. Journal of the American Statistical Association 2007;102:1448–1461.

Cheung, YK. Dose-finding by the continual reassessment method (dfcrm). R package version 0.1-2. 2008. URL http://www.r-project.org

Cheung, YK. Dose Finding by the Continual Reassessment Method. Chapman & Hall; New York: 2010. In preparation

Cheung YK, Chappell R. Sequential designs for phase I clinical trials with late-onset toxicities. Biometrics 2000;56(4):1177–1182. [PubMed: 11129476]

Cheung YK, Chappell R. A simple technique to evaluate model sensitivity in the continual reassessment method. Biometrics 2002;58(3):671–674. [PubMed: 12230003]

Cheung YK, Elkind MSV. Stochastic approximation with virtual observations for dose-finding on discrete levels. Biometrika 2010;97:109–121.10.1093/biomet/asp065

Dixon, WJ.; Mood, AM. A method for obtaining and analyzing sensitivity data. Vol. 43. Journal of the American Statistical Association; 1948. p. 109-126.

Durham SD, Flournoy N, Rosenberger WF. A random walk rule for phase I clinical trials. Biometrics 1997;53(2):745–760. [PubMed: 9192462]

Eisenhauer EA, O'Dwyer PJ, Christian M, Humphrey JS. Phase I clinical trial design in cancer drug development. Journal of Clinical Oncology 2000;18(3):684–692. [PubMed: 10653884]

Faries D. Practical modifications of the continual reassessment method for phase I cancer trials. Journal of Biopharmaceutical Statistics 1994;4:147–164. [PubMed: 7951271]

Finney, DJ. Statistical Method in Biological Assay. Griffin; London: 1978.

Gasparini M, Eisele J. A curve-free method for phase I clinical trials. Biometrics 2000;56(2):609–615. [PubMed: 10877324]

Geller NL. Design of phase I and II clinical trials in cancer: A statistician's view. Cancer Investigation 1984;2:483–491. [PubMed: 6509359]

Haines LM, Perevozskaya I, Rosenberger WF. Bayesian optimal design for phase I clinical trials. Biometrics 2003;59:591–600. [PubMed: 14601760]

Han J, Lai TL, Spivakovsky V. Approximate policy optimization and adaptive control in regression models. Computational Economics 2006;27(4):433–452.

Ji Y, Li Y, Bekele BN. Dose-finding in phase I clinical trials based on toxicity probability intervals. Clinical Trials 2007;4(3):235–244. [PubMed: 17715248]

Kiefer J, Wolfowitz J. Stochastic estimation of the maximum of a regression function. Annals of Mathematical Statistics 1952;23:462–466.

Lai TL, Robbins H. Adaptive design and stochastic approximation. Annals of Statistics 1979;7:1196–1221.

Lee SM, Cheung YK. Model calibration in the continual reassessment method. Clinical Trials 2001;6:227–238. [PubMed: 19528132]

Leonard JP, Furman RR, Cheung YK, et al. Phase I/II trial of bortezomib plus CHOP-Rituximab in diffuse large B cell (DLBCL) and mantle cell lymphoma (MCL): phase I results. Blood 2005;106:147A–147A.

Leung DHY, Wang YG. Isotonic designs for phase I trials. Controlled Clinical Trials 2001;22(2):126–138. [PubMed: 11306151]

McLeish DL, Tosh D. Sequential designs in bioassay. Biometrics 1990;46:103–116.

Morgan, BJT. Analysis of Quantal Response Data. Chapman & Hall; New York: 1992.

Muller JH, McGinn CJ, Normolle D, Lawrence T, Brown D, Hejna G, Zalupski MM. Phase I trial using a time-to-event continual reassessment strategy for dose escalation of cisplatin combined with gemcitabine and radiation therapy in pancreatic cancer. Journal of Clinical Oncology 2004;22:238–243. [PubMed: 14665608]

Murphy JR, Hall DL. A logistic dose-ranging method for phase I clinical investigations trials. Journal of Biopharmaceutical Statistics 1997;7:636–647.

O'Quigley J, Chevret S. Methods for dose finding studies in cancer clinical trials: a review and results of a Monte Carlo study. Statistics in Medicine 1991;10(11):1647–1664. [PubMed: 1792460]

O'Quigley J, Conaway M. Continual reassessment and related dose finding designs. Statistical Science. 2010 accepted.

O'Quigley J, Pepe M, Fisher L. Continual reassessment method: A practical design for phase I clinical trials in cancer. Biometrics 1990;46(1):33–48. [PubMed: 2350571]

Ratain MJ, Mick R, Schilsky RL, Siegler M. Statistical and ethical issues in the design and conduct of phase I and phase II clinical trials of new anticancer agents. Journal of the National Cancer Institute 1993;85:1637–1643. [PubMed: 8411243]

Robbins H, Monro S. A stochastic approximation method. Annals of Mathematical Statistics 1951;22(3):400–407.

Sacks J. Asymptotic distribution of stochastic approximation procedures. Annals of Mathematical Statistics 1958;29(2):373–405.

Schneiderman MA. How can we find an optimal dose. Toxicology and Applied Pharmacology 1965;7:44–53.

Shen LZ, O'Quigley J. Consistency of continual reassessment method under model misspecification. Biometrika 1996;83:395–405.

Storer B. Design and analysis of phase I clinical trials. Biometrics 1989;45:925–937. [PubMed: 2790129]

Storer B, DeMets D. Current phase I/II designs: are they adequate? Journal of Clinical Research Drug Development 1987;1:121–130.

Thall PF. Bayesian models and decision algorithms for complex early phase clinical trials. Statistical Science. 2010 in press.

Thall PF, Millikan RE, Müller P, Lee SJ. Dose-finding with two agents in phase I oncology trials. Biometrics 2003;59(3):487–496. [PubMed: 14601749]

Tighiouart M, Rogatko A. Dose finding with escalation with overdose control (EWOC) in cancer clinical trials. Statistical Science. 2010 under review.

Wasan, MT. Stochastic approximation. Cambridge University Press; 1969.

Whitehead J, Brunier H. Bayesian decision procedures for dose determining experiments. Statistics in Medicine 1995;14:885–893. [PubMed: 7569508]

Wu CFJ. Efficient sequential designs with binary data. Journal of the American Statistical Association 1985;80:974–984.

Wu, CFJ. Maximum likelihood recursion and stochastic approximation in sequential designs. In: Van Ryzin, J., editor. Adaptive Statistical Procedures and Related Topics. Vol. 8. 1986. p. 298-314.IMS Monograph

Ying Z, Wu CFJ. An asymptotic theory of sequential designs based on maximum likelihood recursion. Statistica Sinica 1997;7(1):75–91.
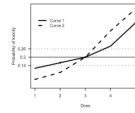
**Fig 1.**
Two dose-toxicity curves under which dose 3 is the MTD with $p = .20$. A $\delta$-sensitive design with $\delta = .06$ will eventually select doses 2 or 3 under the shallow curve (Curve 1), but will be consistent for dose 3 under the steep curve (Curve 2). The horizontal dotted lines indicate the indifference interval.
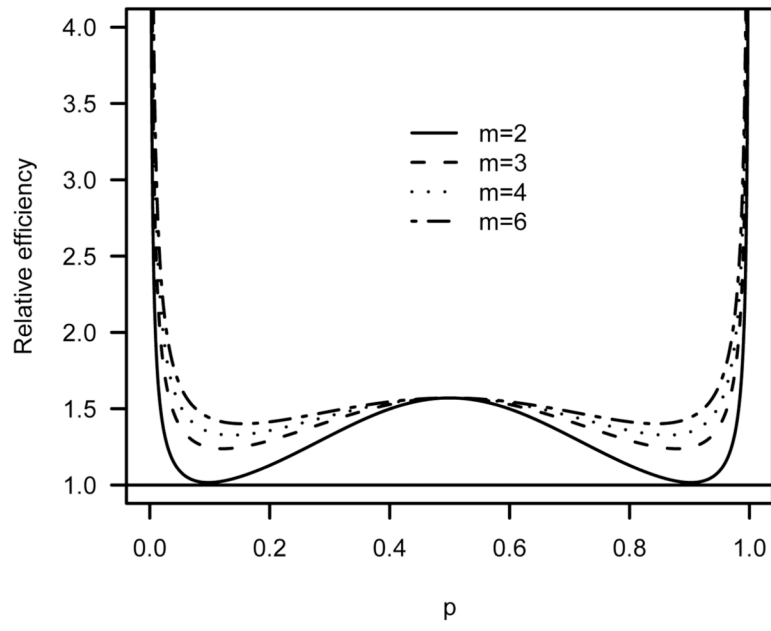
**Fig 2.**
Asymptotic efficiency of $x_n$ based on recursion (16) relative to the logit-MLE $\tilde{x}_n$.

**Table 1**

Dose schedules of bortezomib used in Leonard et al. (2005).

| Level | Dose and schedule within cycle |
|-------|-------------------------------|
| 1 | 0.7 mg/m$^2$ on day 1 of each cycle |
| 2 | 0.7 mg/m$^2$ on days 1 and 8 of each cycle |
| 3 | 0.7 mg/m$^2$ on days 1 and 4 of each cycle |
| 4 | 1.0 mg/m$^2$ on days 1 and 4 of each cycle |
| 5 | 1.3 mg/m$^2$ on days 1 and 4 of each cycle |