Best Intention Designs in Dose–finding Studies

Valerii V. Fedorov
Quintiles, Inc.
Durham, NC, USA
V.V.Fedorov2011gmail.com

Nancy Flournoy
Department of Statistics
University of Missouri
Columbia, MO, USA
flournoyn@missouri.edu

Yuehui Wu
GlaxoSmithKline
Collegeville, PA, USA
Yuehui.2.Wu@gsk.com

and

Rongmei Zhang
Department of Biostatistics and Epidemiology
University of Pennsylvania
Philadelphia, PA, USA
rongmei@alumni.upenn.edu

December 19, 2011

Abstract

In the dose–finding setting, we discuss and compare adaptive designs, including those which allocate each new patient (or a cohort of patients) to the dose currently viewed as the best one, for example, the seemingly most efficacious. We call this type of allocation "best intention" designs and found that, similar to other areas of life, the best intention may pave the road to rather disastrous situations. Theoretical results originating in research on optimal control strategies suggest that the best intention approach needs modifications to insure convergence to the "best" dose. Monte Carlo simulations provide corroborating evidence that caution is needed. Indeed, even for very simple models, best intention designs may converge to a wrong dose with non–zero probability and in some cases almost for sure.

keywords: response-adaptive designs, sequential treatment allocation, clinical trials, optimal design

1 Introduction

In early dose–finding clinical trials, balancing between gathering information and treating the patients enrolled in these studies is an ever lasting conflict, and it generates a massive literature which gravitate to two still contradictory approaches. One approach targets the most effective gathering of information and is essentially based on the ideas of optimal design theory. The second approach follows a natural intention to allocate patients to doses that are the best according to the current knowledge with a hope that some valuable information can be still collected. The genesis of most proposed designs of this type can be found in optimal control theory. Unfortunately, some warnings well known in optimal design and in optimal control theory are not taken seriously in drug development practice. We call the corresponding designs "optimal designs" and "best intention (BI)" designs, respectively. In what follows, we discuss potential pitfalls for both approaches and propose some remedies based on the idea of quantifying potential harm. One pioneering paper in which both approaches were considered together was published by Wetherill [1]. He was probably the first to build locally D-optimal and c-optimal designs for the binary logistic model.
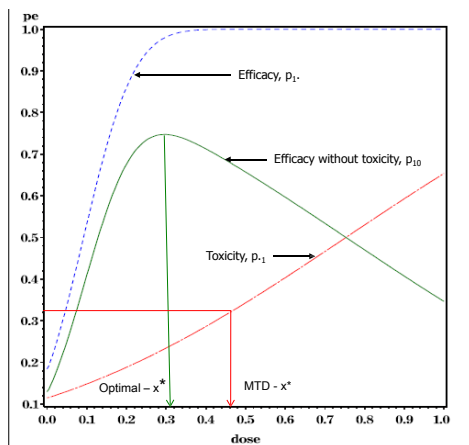
Figure 1: Typical Dose–Response Curves

The allocation of subjects to the dose currently believed to be best, or to doses close to the best one, have become very popular in clinical dose–finding studies, for example, when the intention is to identify the maximum tolerated dose (MTD), the minimum efficacious dose, or the most efficacious dose. Examples are found, for instance, in the already mentioned paper by Wetherill [1], Lai and Robbins [2], Q'Quigley, Pepe and Fisher [3], Li, Durham and Flournoy [4], and Thall and Cook [5]. Best intention designs are promoted as being ethnically attractive and caring about the sampled subjects, but doubts about convergence and informativeness of best intention designs were raised long ago, and cases were found in which such designs led to allocations converging to the wrong point; cf. Lai and Robbins [6], Bozin and Zarr [7], Pronzato [8], Chang and Ying [9], Oron, Azriel and Hoff [10], and Azriel [11].

In Section 2, we define two types of "best" doses in terms of utility functions. Common examples of such best doses are depicted in Figure 1. We describe Best Intention (BI) and Penalized Adaptive $\mathscr{D}$–Optimal (PAD) designs in Section 3. Although there are a number of alternative approaches to dose–finding in the literature (see [12] and references within for a recent review), for simplicity, we restrict comparisons of BI designs to PAD, which are much more general in their goals.

In Sections 3.1 and 3.2, we focus on two types of dose–finding problems,

3

staying in the framework of two very simple models, namely, the one dimensional continuous linear and quadratic response functions, respectively. For these models, one can easily understand the drawbacks of naive BI designs and see remedies. To show that the issues we raise translate directly to nonlinear models, we create a probit model in Section 3.3 by dichotomizing continuous responses in those linear models. In Section 3, we assume that doses are selected from a continuous sample space.

In Section 4, we discuss the effects of restricting the search for a best dose to a finite number of doses; and we give concluding remarks in Section 5.

2 Utility and Penalty Functions

Consider responses Y that are continuous, dichotomous or ordinal, assuming that their expectation,

$$\mathrm{E}[\mathbf{Y}|x] = \boldsymbol{\eta}(x, \boldsymbol{\theta}),$$

are known functions of x and $\boldsymbol{\theta}$; x may be a dose or a combination of doses selected from a set $\mathscr{X}$ and $\boldsymbol{\theta}$ is a $m-$dimensional vector of unknown parameters.

In studies such as those with both efficacy and toxicity, $\mathbf{y}$ and $\boldsymbol{\eta}$ will be vectors (e.g., Gooley, et al. [13], Li, Durham, and Flournoy [4], Fan and Chaloner [14], Rabie and Flournoy [15], Thall and Cook [5], and Dragalin and Fedorov [16]). For instance, correlated binary measures of efficacy and toxicity yield four possible outcomes, and a vector of expected responses can be defined as

$$\boldsymbol{\eta}^{\mathrm{T}}(x, \boldsymbol{\theta}) = (p_{10}(x, \boldsymbol{\theta}), p_{11}(x, \boldsymbol{\theta}), p_{01}(x, \boldsymbol{\theta}), p_{00}(x, \boldsymbol{\theta})),$$

where $p_{10}(x, \boldsymbol{\theta})$ is the probability of efficacy and no toxicity; $p_{11}(x, \boldsymbol{\theta})$ is the probability of efficacy with toxicity, etc.

A practitioner typically is concerned, not with $\boldsymbol{\eta}(x, \boldsymbol{\theta})$ itself, but with a utility function $\zeta(x, \boldsymbol{\theta})$ that describes the potential benefits associated with treatment at a particular dose. Reasonable utility functions in the study of efficacy and toxicity include $\zeta(x, \boldsymbol{\theta}) = p_{10}(x, \boldsymbol{\theta})$ and $\zeta(x, \boldsymbol{\theta}) = p_{10}(x, \boldsymbol{\theta}) + p_{11}(x, \boldsymbol{\theta})$. Another example can be build in the following way: let $p_{1 \cdot}(x, \boldsymbol{\theta})$ and $p_{\cdot 1}(x, \boldsymbol{\theta})$ denote the marginal probabilities of efficacy and toxicity, respectively, and let $p_{1 \cdot}^{*}$ and $p_{\cdot 1}^{*}$ denote "desirable" values of these probabilities; utility functions can be defined as $\zeta(x, \boldsymbol{\theta}) = -\pi(x, \boldsymbol{\theta})^{\mathrm{T}} \mathbf{W} \pi(x, \boldsymbol{\theta})$

4

where $\pi^{\mathrm{T}}(x, \boldsymbol{\theta}) = [p_{1\cdot}(x, \boldsymbol{\theta}) - p_{1\cdot}^*, p_{\cdot 1}(x, \boldsymbol{\theta}) - p_{\cdot 1}^*]$ measures the discrepancy between probabilities at x and the best dose. If $\mathbf{W}$ is the identity matrix, toxicity and efficacy are given equal weight and the utility function is $\zeta(x, \boldsymbol{\theta}) = (p_{1\cdot}(x, \boldsymbol{\theta}) - p_{1\cdot}^*)^2 + (p_{\cdot 1}(x, \boldsymbol{\theta}) - p_{\cdot 1}^*)^2$. See also Dragalin, Fedorov and Wu [17], Gooley et al. [13] and Thall and Cook [5], and . This type of utility functions is similar to what in engineering and economics is called the desirability function [18].

The two most common problems in dose finding studies are, respectively, the search for the dose in which the utility function equals a prespecified value ζ^* and the dose that maximizes the utility function. Denoting the best dose being sought by $x^*(\boldsymbol{\theta})$, we can define these two problems as

Type I :
$$x^*(\boldsymbol{\theta}) = \arg \min_{x \in \mathscr{X}} \mid \zeta(x, \boldsymbol{\theta}) - \zeta^* \mid ; \tag{1}$$

Type II :
$$x^*(\boldsymbol{\theta}) = \arg \max_{x \in \mathscr{X}} \zeta(x, \boldsymbol{\theta}). \tag{2}$$

One can argue that Type I problem can be viewed as a special case of Type II with utility function $\zeta' = - \mid \zeta - \zeta^* \mid$ but we prefer to consider them separately to emphasize the very different properties of BI designs for each of them. Note that in general $\mathscr{X}$ can be either continuous or discrete. Everywhere, except in Section 4, $\mathscr{X}$ is assumed to be continuous and and if it is not stated differently $\mathscr{X} = [-1, +1]$. When the toxicity rate is increasing, designs aiming to identify a dose having a prescribed toxicity rate will be said to be of Type I. This dose is often called the maximum tolerated dose (MTD); see the example in Figure 1. If the probability of efficacy increases with dose, identifying the least dose that is effective $100p\%$ of the time, the EDp, is mathematically the same problem as identifying the MTD. However, search procedures may differ. For example, in searching for the MTD, trials typically start with low doses and escalate avoiding overdosing, whereas, in searching for the EDp, trials typically start with a high dose and it is assumed that is not as harmful as over dosing. Early examples of Type I dose–finding designs are given by Derman [19], Dixon and Mood [20], Durham and Flournoy [21], O'Quigley et al. [3], Robbins and Monro [22], von Békésy [23], Wetherill [24] and Wu [25]; others can be found in [12] and the references therein.

In the toxicity and efficacy curves depicted in Figure 1, the probability of efficacy with no toxicity increases as doses become more efficacious and then decrease when doses become more toxic. In this context, a typical goal is to identify the dose having maximum probability of efficacy without toxicity, and this is a Type II problem. Examples of Type II designs are found in Dragalin and Fedorov [16], Dragalin, Fedorov and Wu [26], Durham, Flournoy and Li [27], Fedorov and Mueller [28], Fedorov and Wu [29], Gooley, et al. [13], Hardwick and Stout [30], Kiefer and Wolfowitz [31], Kpamegan and Flournoy [32], Li, Durham and Flournoy [4], Pronzato [33] and Thall and Cook [5].

To estimate $x^*(\boldsymbol{\theta})$ with the least possible uncertainty (e.g., variance or standard deviation) and with the least harm to patients in the study are objectives that often "pull" designs in different directions. The phrases "treatment versus experimentation", "treatment versus learning", etc. can be found in numerous publications on dose–finding studies. To quantify the problem we have to introduce a measure of losses, or harm, or cost, or potential penalty for a wrong prediction of x^*, etc.. The situation is similar to what is traditionally used in decision theory and associated with risk, loss or regret function functions, cf. Le Cam [34] and Pratt, Raiffa and Schlaife [35]. In dose–finding setting, Lai and Robbins [2] were probably the first to use a penalty function to measure the quality of the proposed design. Dragalin, Fedorov, Wu [26] and Pronzato [33] regularly use penalized optimal designs in dose–finding experiments.

In drug development, there are a few players that roughly can be described as the targeted population, sampled patients, a specific patient, a sponsor (e.g. pharmaceutical company) and various regulating agencies. For instance, for the n-th in-trial patient, Lai and Robbins use the quadratic penalty $\phi(x_n) = (x_n - x^*)^2$, where x_n is the dose he(she) is allocated. As for the targeted population the potential loss (harm, penalty) is determined by uncertainty of the recommended (predicted) best dose x^*, i.e. $\mathrm{E}[(\hat{x}_N^* - x^*)^2]$; for the sampled patients the total average penalty $N^{-1} \sum_{n=1}^N \phi(x_n)$, or its expected value, might be a sound measure of harm. Fedorov and Wu added the cost of an observation and used $\phi(x_n) = (x_n - x^*)^2 + c$ for a patient and its obvious extensions for the sampled and targeted populations. In this paper, we use the same penalty function in comparing designs. More penalty functions can be found in already cited [26], [29] and [33].

6

3 Best Intention and Adaptive $\mathscr{D}$–Optimal Designs

Sections 3.1 and 3.2 discuss the use of specific BI and penalized PAD designs with continuous response variables for linear and quadratic response models and with best doses of Type I and Type II, respectively. In Section 3.3, we consider probit models that are the dichotomized versions of two previous cases. All findings for continuous responses hold for the dichotomized ones, except that one has to perform significantly more observations to obtain comparable results, cf. Fedorov and Liu [36], [37].

3.1 Type I problem: $x^*(\boldsymbol{\theta}) = \arg\min_{x \in \mathscr{X}} |\zeta(x, \boldsymbol{\theta}) - \zeta^*|$.

To examine the simplest Type I dose–finding problem, consider

$$y_i = \eta(x, \boldsymbol{\theta}) + \epsilon, \quad \zeta(x, \boldsymbol{\theta}) = \eta(x, \boldsymbol{\theta}) = \theta_1 + \theta_2 x, \qquad (3)$$
$$\mathrm{E}[\epsilon] = 0 \quad \mathrm{Var}[y|x] = \sigma^2, \quad x \in \mathscr{X} = [-1, 1]. \qquad (4)$$

Let $\tilde{x}(\boldsymbol{\theta}) = (\zeta^* - \theta_1)/\theta_2$. The best dose is defined as

$$x^*(\boldsymbol{\theta}) = \arg\min_{x \in \mathscr{X}} |\zeta(x, \boldsymbol{\theta}) - \zeta^*| = \left\{ \begin{array}{ll} \tilde{x}(\boldsymbol{\theta}), & |\tilde{x}(\boldsymbol{\theta})| \leq 1 \\ \pm 1, & \pm\tilde{x}(\boldsymbol{\theta}) > 1 \end{array} \right. . \qquad (5)$$

While model (3) looks too simple to be useful in dose–finding studies, where logistic, probit and other nonlinear models dominate, it provides a local approximation for any dose–finding problem with a "smooth" utility function. Therefore, the asymptotic properties of other BI and PAD designs that are valid for this model will be very much the same as for any other smooth (twice differentiable) response model.

In 1951 Robbins and Monro [22] proposed to allocate n-th subject at

$$x_{n+1} = x_n + \alpha_n(y_n - \zeta^*), \qquad (6)$$

where $\{\alpha_n\}_{n=1}^{\infty}$ is any deterministic sequence such that $\sum_{n=1}^{\infty} \alpha_n = \infty$ and $\sum_{n=1}^{\infty} \alpha_n^2 \leq \infty$. This procedure has the best convergence rate if $\alpha_n = \theta_2/n$; see Sacks [38]. One can also verify (cf. Wetherill [1], Anbar [39]) that (6) with $\alpha_n = \theta_2/n$ is equivalent to

$$x_{n+1} = \bar{x}_n - \theta_2^{-1} \left(\bar{y}_n - \zeta^* \right), \qquad (7)$$

so that while (6) looks like a short memory procedure, the opposite is true! On an intuitive level, one can say that x_n carries cumulating prior information on the experiment. Unfortunately, knowledge of the true value of the slope θ_2 is essential for equating (6) with (7).

In practice it seems natural to replace θ_2 by its least square or maximum likelihood estimator $\hat{\theta}_{2n}$:

$$x_{n+1} = \bar{x}_n - \hat{\theta}_{2n}^{-1} \left(\bar{y}_n - \zeta^*\right). \tag{8}$$

Although motivated by the Newton–Raphson Method (see Ypma [40] and Lai [41]), Robbins and Munro proposed using a sequence of constants in place of $\hat{\theta}_{2n}/n$ in (6) because they already knew that, in spite of the seeming simplicity of (8), such plug in prediction estimates caused problems. Lai and Robbins [6] noted that, with a non zero probability, x_n sticks to a boundary point from which it does not move away even if $n \to \infty$. A few corrective measures were proposed to make $\{x_n\}$ converge to x^*, and to be asymptotically normally distributed. Sequences of predicted best doses converge to the wrong doses not only for the simple linear response model, but also for more general models (see, for instance, Wu [25] and [9]) when

$$x_{n+1} = \arg\min_{x \in \mathscr{X}} \left|\zeta(x, \hat{\boldsymbol{\theta}}_n) - \zeta^*\right|. \tag{9}$$

For (8) and (9), Wu [25] called this specific BI design the Adaptive Robins–Monro procedure (ARM). The corrective measures are rather simple: one has to bound absolute value of the estimated slope or select a sign (but correct one) to insure convergence and asymptotical normality [6],[39]. Wu [25] also noted that asymptotically (6), (7) and (8) coincide. More details will be discussed in the examples that follow. We use the term "naive" ARM when either (8) or (9) is used directly, i.e. without any adjustment. Otherwise we use "tuned" ARM.

To illustrate the potential lack of consistency in the allocation sequence $\{x_n\}$, we resort to Monte–Carlo simulations with independently normally distributed observations; and $\theta_1 = 0.0$, $\theta_2 = 1.0$, $\sigma^2 = 1.0$, $\zeta^* = 0.0$; i.e., $x^* = 0.0$; in addition, doses were restricted to lie in $[-1, 1]$. Figure 2(a) shows exemplary sequences of $\{x_1, \ldots, x_{400}\}$ from (8) with an initial cohort of two subjects, one each at ± 1. Most BI designers use the last x_n to estimate x^*, and we do so here, so $\{x_1, \ldots, x_{400}\}$ is also the sequence of predicted best doses. In this and all other examples, 10,000 Monte–Carol simulations
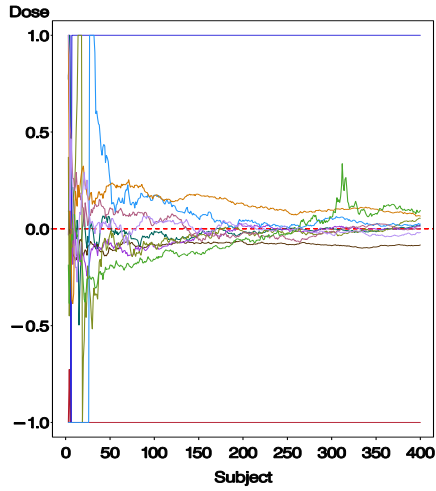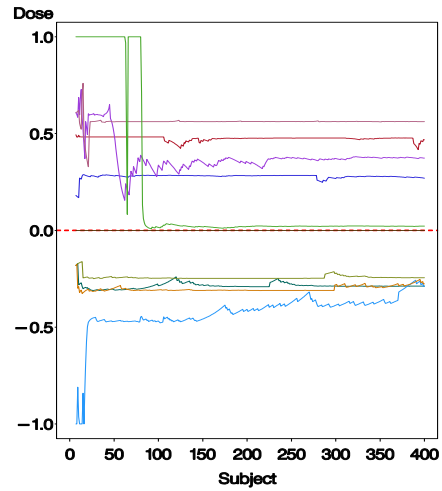
(a) Type I ARM predicted best dose sequences under the continuous linear model.
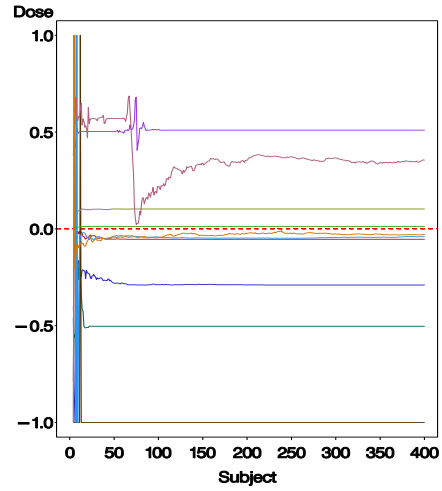
(b) Type II BI predicted best dose sequences under the continuous quadratic model.
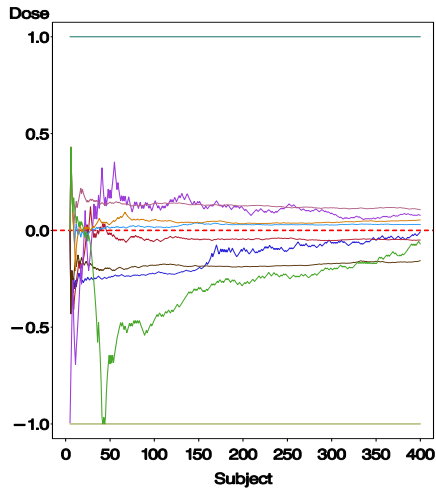
(c) Type I BI predicted best dose sequences under the probit model with $F(\theta_1 + \theta_1 x)$.

(d) Type II BI predicted best dose sequences under the probit model with $F(\theta_1 + \theta_2 x + \theta_3 x^2)$.

Figure 2: Sample dose allocations sequences $\{x_1, \ldots, x_{400}\}$ with best doses $x^* = 0$

9

were performed. Note that some sequences show early wide variation that simmers down after about 25 trials; convergence then occurs approximately at rate $n^{-1/2}$. Note the lines at ± 1 indicate that some sequences "stick" to boundaries of the design space, and there is no evidence that these sequences will ever converge to the best dose $x^* = 0.0$. Table 1 shows that 1.9 and 1.8 percent of predicted best dose sequences stuck to the boundaries for $n = 100$ and 400, that is, increasing the sample size four fold did not reduce the likelihood of this problem when using ARM with the continuous linear model.

In contrast, the PAD design allocates the next subject to the dose

$$x_{n+1} = \arg \max_{x \in \mathscr{X}} \text{Var} \left[\eta(x, \hat{\boldsymbol{\theta}}_n) \right], \tag{10}$$

where the variance of the predicted response is the largest, i.e. our knowledge about response function is the worst. Observation at this point maximizes the decrements of the determinant of the variance–covariance matrix of $\hat{\boldsymbol{\theta}}_n$; see Fedorov [42]. For a simple linear regression, the sequence $\{x_n\}$ consists alternating $+1$ and -1. With penalty functions (see appendix), the selection of x_n involves both $\text{Var} \left[\eta(x, \hat{\boldsymbol{\theta}}_n) \right]$ and the penalty function $\phi(x, \hat{\boldsymbol{\theta}}_n)$ and

$$x_{n+1} = \arg \max_{x \in \mathscr{X}} \left\{ \text{Var} \left[\eta(x, \hat{\boldsymbol{\theta}}_n) \right] - \frac{m \phi(x, \hat{\boldsymbol{\theta}}_n)}{\sum_{i=1}^n \phi(x_i, \hat{\boldsymbol{\theta}}_n)} \right\}, \tag{11}$$

where in general m is the number of unknown parameters in $\boldsymbol{\eta}(x, \boldsymbol{\theta})$ and in the considered case $m = 2$. For PAD designs, typically least squares or maximum likelihood methods are used to estimate the best dose.

From Table 1, one can see that no PAD predicted best dose sequences stuck to the boundaries in 10,000 simulations for $n = 100$ and 400, as compared with about 2 percent using ARM. More is said about comparing estimators later.

Figures 3(a) and 3(b) show panels of histograms of predicted best doses for ARM and PAD, respectively, that is, histograms of x_{100} and x_{400} are shown from ARM and histograms of least squares estimates, $\hat{x}_{100}^*$ and $\hat{x}_{400}^*$, are shown from the PAD design. Histograms from each procedure were simulated with start-up sample sizes of 2 and 4, and these are shown in the first and second rows, respectively. As expected for each procedure, comparing columns of histograms for $n = 100$ and 400, one sees the histograms are much more

Table 1: Percent of Predicted Best Dose Sequences Stuck on the Boundaries.

MODEL	Start-up Sample Size	Total Sample $n = 100$		Total Sample $n = 400$	
		ARM/BI	PAD with $c = 0.10$	ARM/BI	PAD with $c = 0.10$
Continuous linear	2	1.9	0.0	1.8	0.0
	8	0.0	0.0	0.0	0.0
Continuous quadratic	3	11.3	0.0	10.9	0.0
	12	1.8	0.0	1.5	0.0
Probit $F(\theta_1 + \theta_2 x)$	4	0.7	0.4	0.9	0.0
	16	0.0	0.0	0.0	0.0
Probit $F(\theta_1 + \theta_2 x + \theta_2 x^2)$	6	6.5	0.1	5.8	0.0
	24	0.9	0.0	0.6	0.0

dispersed for $n = 100$ than for $n = 400$. Standard deviations for each set of data give the same message.

However, the frequency of observations on the boundaries decreases from 1.9 percent for ARM with a start-up sample size of 2 to 0.0 with a start–up sample size of 8 (Table 1) and comparing the rows of histograms in Figures 3(a) and 3(b), significant improvement can be seen in having a larger fixed start–up cohort. With start–up cohorts of size 2, the standard deviation of ARM predicted best doses reduces from 0.17 for $n = 100$ to 0.14 for $n = 400$ but it halves from 0.10 reducing to 0.05 with start–up cohorts of size of 8. This observation leads to the obvious recommendation that the size and allocation of the start–up cohorts for ARM should be selected carefully and should be at least large enough to avoid estimators that stick to boundaries. In procedure (8), $x_{n+1} = -\hat{\theta}_{1n}/\hat{\theta}_{2n}$, which for start–up designs is the ratio of two normally distributed random variables, which may have a distribution of quite exotic shape; for instance, it may be multi–modal; cf. Hinkley [43].

In contrast with ARM, PAD predicted best doses sequences did not stick to the boundaries (Table 1) and their standard deviation (Figures 3(a) and 3(b)) halves as expected from 0.10 reducing to 0.05 when n increases from 100 to $n = 400$ regardless of start–up cohorts size. Note that with the larger initial cohort size, histograms of ARM predicted best doses are visually similar to those from PAD designs, and both of them are asymptotically

11
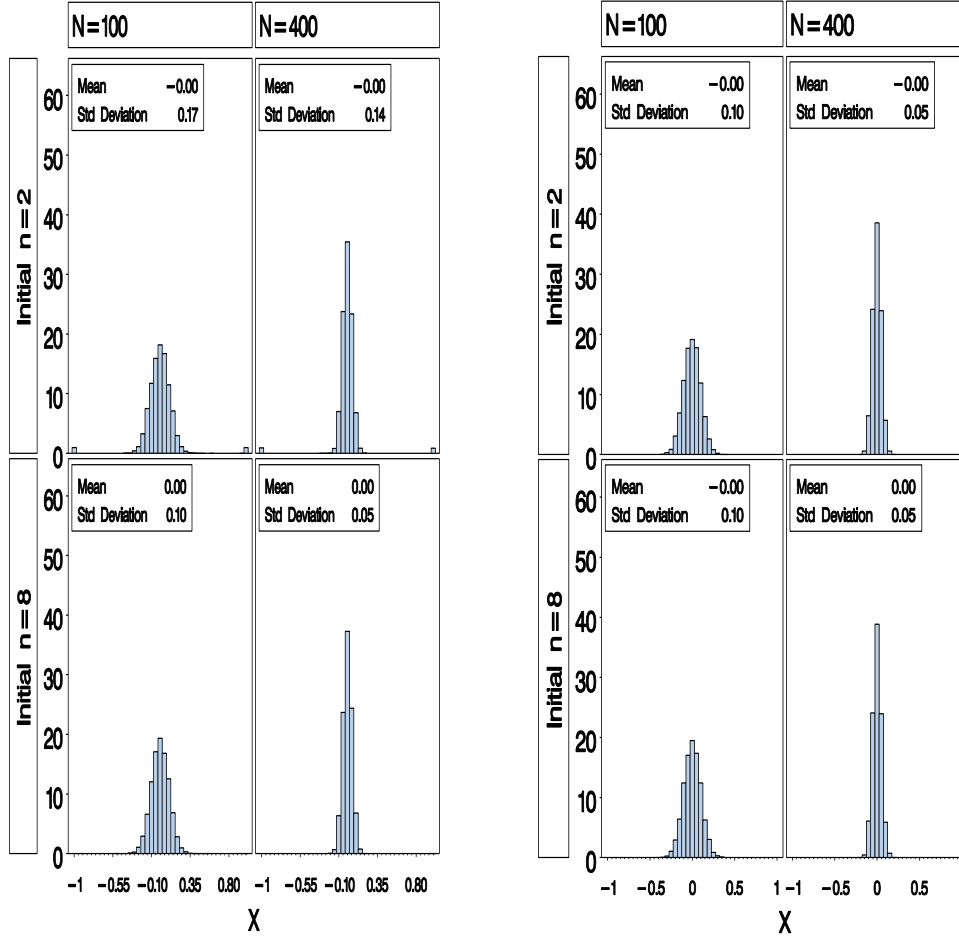
(a) Predicted best doses from the naive ARM design. Columns show frequencies of x_{100} and x_{400}, respectively.

(b) Predicted best doses from PAD designs with cost $c = 0.10$. Columns show frequencies of least squares estimates $\hat{x}^*_{100}$ and $\hat{x}^*_{400}$, respectively.

Figure 3: Predicted best doses under the continuous linear model (3) with $\theta_1 = 0.0$ and $\theta_2 = 1.0$; $x^* = 0.0$. Histograms in the top and bottom rows had start–up cohorts sizes of 2 and 8, respectively, equally allocated to ± 1.

identical to the distribution of x_n from the tuned ARM (the slope θ_2 is bounded and estimates $\hat{\theta}_{2n}$ are correspondingly restricted. More about tuning of ARM may be found in Lai's survey papers [41] and with Bartroff [44]).

Again, we emphasize that the initial design and its sample size are critical to ARM performance, but have no noticeable effect on PAD designs.

Table 2 compares risks based on $\phi(x, \boldsymbol{\theta}) = (x - x^*)^2$ from the tuned ARM and the locally $\mathscr{D}-$optimal design for the targeted population, sampled patients, a specific patient and a sponsor. Note that with these more idealized procedures that the tuned ARM is equally good (in that it gains the same information about x^*) as the $\mathscr{D}-$optimal design, but superior in having smaller risk. The tuned ARM has reduced risk for two types of customers, the sample in-study patients and an individual patient, mainly due to the selection of a very specific loss function. As soon as one adds a per-patient cost c to obtain $\phi(x, \boldsymbol{\theta}) = (x - x^*)^2 + c$, the slow logarithmic growth of the ARM risk to a patient sample is replaced with linear growth, putting its convergence rate on par with locally $\mathscr{D}-$optimal designs.

Table 2: Risks for Different Customers

Customers	Risk	Tuned ARM	Locally $\mathscr{D}-$optimal
Targeted Population	$\mathrm{E}\left(x_N^* - x^*\right)^2$	$\sim (\sigma/\theta_2)^2\, N^{-1}$	$(\sigma/\theta_2)^2\, N^{-1}$
Patient Sample	$\mathrm{E}\left[\sum_{i=1}^{N} \left(x_i^* - x^*\right)^2\right]$	$\sim (\sigma/\theta_2)^2\, ln\, N$	N
nth Patient	$\mathrm{E}\left(x_n^* - x^*\right)^2$	$\sim (\sigma/\theta_2)^2\, n^{-1}$	1
Sponsor	N	$Q + qN$	$Q + qN$

$\sim$ denotes "asymptotically", or more loosely, for large N; Q is the cost of a trial initiation; q is the cost of a patient enrollment.

Total penalties, $\sum_{i=1}^{n} \phi(x_i, \boldsymbol{\theta}) = (x_i - x^*)^2 + nc$, with per-subject cost $c = 0.10$, are shown in Table tab:totalpenalty for ARM and PAD designs with $n = 400$. With the continuous linear model, total penalties using ARM are seen to be skewed and to depend heavily on the start–up cohort size (increasing start–up size from 2 to 8, the mean drops from 13.1 to 4.2, which is the mean total penalty using the PAD design for both cohorts). In comparison, the PAD penalties have relatively little skew and are independent of start–up cohort size. ARM and PAD procedures have similar penalties with the larger

13

start–up sample.

Table 3: Total Penalty $\sum_{i=1}^{n} (x_i - x^*)^2 + nc$ for $c = 0.10$ and $n = 400$.

MODEL	Start-up Size	ARM/BI Mean	ARM/BI Median	PAD Mean	PAD Median
Continuous	2	13.1	4.8	4.2	3.0
linear	8	4.2	3.1	4.2	3.0
Continuous	3	71.1	18.3	7.2	4.7
quadratic	12	29.1	7.2	4.9	2.9
Binary	4	17.3	9.9	9.1	6.6
$F(\theta_1 + \theta_2 x)$	16	8.1	5.7	5.9	4.3
Binary	6	57.0	21.1	9.8	4.5
$F(\theta_1 + \theta_2 x + \theta_2 x^2)$	24	24.2	6.1	4.2	2.2

Interestingly, for simple linear regression with penalty $(x - x^*)^2 + c$, there are infinitely many designs ξ^* with the same penalized information matrix, $M(\xi^*, \theta) / \Phi(\xi^*, \theta)$; see the Appendix for definitions of the normalized information matrix, the penalized adaptive optimal design and the relation between Type II dose-finding and regularized optimization. In particular, all designs that are symmetrical with respect to the x^* and satisfy the simple condition on its second moment,

$$\int_{x^*-1}^{x^*+1} (x - x^*)^2 \, \xi(dx) = c, \tag{12}$$

are locally $\mathscr{D}$–optimal designs. So there is considerable flexibility in selecting doses. For instance, at the beginning of the study, one can allocate subjects at the boundaries of an interval that has high probability of covering the unknown x^* in order to insure the ability of the linear model to locally describe the true response function. After some knowledge about the location of x^* is obtained, symmetric allocations may be made closer to x^*. The constraint (12) insures that the rate of convergence of $\hat{\theta}_{2n}$ to the true slope θ_2 is of order $n^{-1/2}$ for PAD procedures. One penalized $\mathscr{D}$–optimal procedure distributes subjects among three doses such that

$$x_1 = x^* - \sqrt{\frac{c}{2w_n}} \quad x_2 = x^* \quad x_3 = x^* + \sqrt{\frac{c}{2w_n}}, \tag{13}$$

14

where w_n is the proportion of subjects in the nth cohort that are allocated to doses x_1 and x_3, and $1 - 2w_n$ subjects are allocated to dose x_2. Alternatively, if subjects are treated sequentially, randomize a subject to doses x_1, x_2 and x_3 with probability w_n, $1 - 2w_n$ and w_n, respectively. Probably the reader has noted that unlike $\mathscr{D}$−optimal designs that can be build for the linear regression model a priori, penalized $\mathscr{D}$−optimal designs depend on x^* and that is why we need to use adaptive designs.

While knowledge of x^* and penalties/risks are important components of clinical trial design, the purpose of most trials also is to learn (at least) about the local behavior of the dose–response function around x^*. Figure 4 plots $\hat{\theta}_{1n}$ by $\hat{\theta}_{2n}$ from the naive ARM and PAD designs after $n = 100$ and 400 observations. Estimates from PAD designs are tightly packed, showing small variation, relative to estimates from ARM deigns. With the naive ARM, variation is significantly greater, not ellipsoid in shape, and distinct clusters of estimates corresponding to sequences stuck on the boundaries are especially evident for designs with small start–up cohorts. We have observed (data not shown) that a tuned ARM usually eliminates predicted best dose sequences from the boundaries, but does not change the general shape of the central "cloud" of parameter estimates, that is, the distribution of parameter estimates from ARM is much more variable and very much less ellipsoid in shape than from PAD designs.

Figure 4 also shows that even though x_n from an ARM with a sufficient initial cohort size will estimate x^* well, the corresponding estimators for $\hat{\theta}_{1n}$ and $\hat{\theta}_{2n}$ behave rather poorly relative to PAD designs. The following observations may help to explain this fact. For polynomial regression of order m centered at $x^* = 0.0$, the values of the elements of the information matrix,

$$M_{\alpha\beta}(n) = \sum_{n'=1}^{n} x_{n'}^{\alpha+\beta-2}, \quad \alpha, \beta = 0, \ldots, m,$$

determine the behavior of $\hat{\theta}_n$, see Lai [45] and with Wei [46]. As was reported in Lai [45] for the tuned ARM, the allocations x_n converge to x^* at rate $n^{-1/2}$. In this case, the element $M_{22}(n)$ determines the asymptotic variance of the slope $\hat{\theta}_{2n}$ which still grows to infinity but very slowly:

$$M_{22}(n) = \sum_{n'=1}^{n} x_{n'}^2 \sim \sum_{n'=1}^{n} \frac{1}{n'} \sim \mathcal{C} + ln\ n \to \infty,$$

15
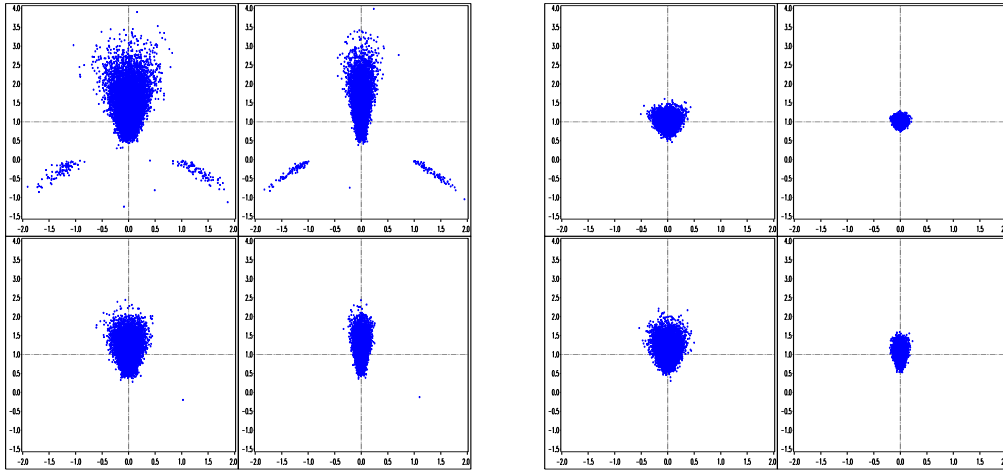
(a) $\hat{\theta}_{2n}$ by $\hat{\theta}_{1n}$ from naive ARM. Columns show estimates at $n = 100$ and 400, respectively.

(b) $\hat{\theta}_{2n}$ by $\hat{\theta}_{1n}$ from PAD designs with cost $c = 0.10$. Columns show estimates at $n = 100$ and 400, respectively.

Figure 4: Plots of $\hat{\theta}_{2n}$ by $\hat{\theta}_{1n}$ under the continuous linear model (3) with $\theta_1 = 0.0$ and $\theta_2 = 1.0$; $x^* = 0.0$. The top and bottom rows had initial cohorts of size of 2 and 8, respectively, equally allocated to ± 1.

where $\mathcal{C} = 0.577\ldots$ is Euler's constant, while $M_{00} \sim n$ and that insures the faster $(1/\sqrt{n})$ convergence rate of $\hat{\theta}_{2n}$. For PAD designs, $M_{22}(n) \sim n$ provides much faster growth than ARM designs with $M_{22}(n) \sim \sqrt{ln\ n}$. This fact was noted by Anderson and Taylor [47].

To conclude this section, we summarize the discussed results:

- For Type I problem the naive ARM, which can be viewed as the best intention design approach, may fail.

- Simple tuning, like bounding the slope and its estimates, or adding extra variability to x_n, makes the ARM a good method for estimating x^* and allocating subjects near x^*. Details on various improvements/tunings of ARM are given by Lai [41] and [45].

- ARM leads to lower (comparatively to $\mathscr{D}-$optimal designs) quadratic risks. However, the penalized adaptive $\mathscr{D}-$optimal designs is a strong alternative.

- ARM gives very poor estimation of the local (in the vicinity of x^*)

behavior of the dose–response curve. Penalized adaptive $\mathscr{D}-$optimal designs do much better job.

- One can verify that the adaptive penalized c–optimal designs, i.e. designs that minimize $\phi(x,\boldsymbol{\theta})\mathrm{Var}\left[\hat{x}^*\right]$, where $\hat{x}^* = -\hat{\theta}_1/\hat{\theta}_2$ and $\phi(x,\boldsymbol{\theta}) = (x - x^*)^2 + c$, coincide with naive ARM, and therefore, what is true for ARM is relevant here as well.

3.2   Type II problem: $x^*(\boldsymbol{\theta}) = \arg\max_{x \in \mathscr{X}} \left[\zeta(x,\boldsymbol{\theta})\right]$.

As in the previous section, we examine the simplest model for Type II dose–finding experiments:

$$y = \eta(x,\boldsymbol{\theta}) + \epsilon, \quad \zeta(x,\boldsymbol{\theta}) = \eta(x,\boldsymbol{\theta}) = \theta_1 + \theta_2 x + \theta_3 x^2,$$
$$\mathrm{E}[\epsilon] = 0, \quad \mathrm{Var}[y|x] = \sigma^2, \quad x \in \mathscr{X} = [-1, 1]. \tag{14}$$

The dose that maximizes the quadratic function is considered to be best, i.e., $x^*(\boldsymbol{\theta}) = -\theta_{2n}/2\theta_{3n}$. While (14) is simple, it provides a good approximation to every "smooth" function in the vicinity of x^*.

For (14), the naive BI procedure is

$$x_{n+1}(\boldsymbol{\theta}) = \arg\max_{x \in \mathscr{X}}\left[\zeta(x,\boldsymbol{\theta})\right] = \begin{cases} \tilde{x}_n & \text{if } |\tilde{x}_n| \leq 1, \\[2mm] \pm 1 & \text{if } \tilde{x}_n > 1 \text{ or } < -1, \text{ respectively,} \end{cases} \tag{15}$$

where $\tilde{x}_n = -\hat{\theta}_{2n}/2\hat{\theta}_{3n}$.

Independent observations were simulated from the normal distribution $\epsilon \sim N(\boldsymbol{\theta}, \sigma^2)$ with $\theta_1 = 1.0$, $\theta_2 = 0.0$, $\theta_3 = -1.0$ and $\sigma^2 = 1.0$, so again the best dose is $x^*(\boldsymbol{\theta}) = 0.0$. Each simulated sequence begins with a fixed initial cohort of either 3 or 12 patients equally allocated to -1, 0 and 1. Simulations of PAD designs include the quadratic penalty with cost $c = 0.10$. Details about this PAD design and some alternatives are in the appendix.

Table 1 shows that 11.3 percent of BI predicted best dose sequences were stuck to the boundaries when $n = 100$, and this percentage only reduced to 10.9 when n increases to 400. Thus the boundary problem for BI sequences is significantly worse for nonlinear models. In contrast, again, 0.0% of the PAD predicted best dose sequences were stuck on the boundaries.

Most alarming for BI designs, however, is that all predicted best dose sequences are seen in Figure 2(b) converge, but to values that are different

from x^*! This phenomenon was noted long ago and a short survey with a number of major references can be found in [48]. Of course after disappointing Monte-Carlo simulations, no theoretical proof is needed for this negative result. However, more subtle theoretical results, like the existence of the attractors different from x^* can be found, for instance, in [7].

Histograms of predicted best doses are presented, i.e, x_{100} and x_{400} from the naive BI design in Figure 5(a), and x^*_{100} and x^*_{400} from the PAD design in In Figure 5(b). The top and bottom rows of histograms result from fixed start–up cohorts of size $n = 3$ and $n = 12$, respectively. But unlike the ARM, BI histograms for x_{100} and x_{400} are (almost) identical! There is no improvement after 300 observation were added. Adding 300 observations, the standard deviation for the ARM best dose predictions reduces from 0.42 to 0.41 for start–up size 3 and from 0.27 to 0.26 for start–up size 12. Similarly to the ARM under linear models, increasing the start–up sample sizes visually "improves" the histograms, but still no sequences converge to the true target dose! Only due to the start–up design, and without gaining substantial information from the BI procedure, are the sequences getting closer to the origin.

The lack of improvement with increased sample sizes for BI designs under the quadratic model is also seen in the plots of parameter estimates in Figure 6(a); and as for ARM with the linear model, Figure 6(a) shows that parameter estimates for the BI design with the quadratic model are clustered. For corresponding PAD designs, marked improvement is seen with increased sample size and clusters are not evident among the parameter estimates (Figure 6(b)).

To illustrate why BI designs lead to very poor estimates of $x^*(\boldsymbol{\theta})$, let us look at the behavior of the least squares (or maximum likelihood) estimates $\hat{\boldsymbol{\theta}}_n$. Let assume that this estimate is at least as good as a (non-adaptive) estimate generated by some fixed regular design. For the latter, the rate of convergence is $O(n^{-1/2})$. Similar to the ARM for the linear model,

$$M_{22}(n) = \sum_{n'=1}^{n} x_{n'} \sim \mathcal{C} \ + \ ln \ n \to \infty, \tag{16}$$

and it gives a hope for consistency of $\hat{\theta}_{2n}$. But (see, for example [49], Ch.1.2),

$$M_{33}(n) \ \sim \ \sum_{n'=1}^{n} x_{n'}^4 \sim \sum_{n'=1}^{n} 1/n'^2 \longrightarrow \pi^2/6 \tag{17}$$

18
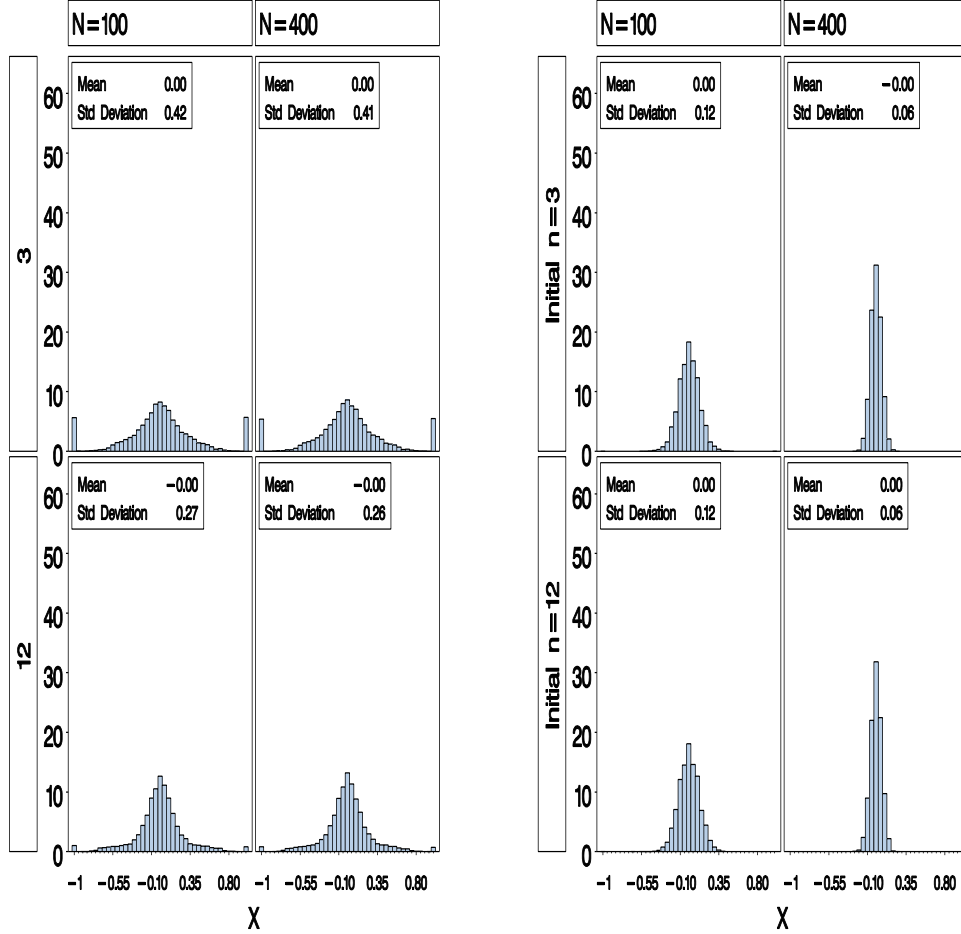
(a) Predicted best dose frequencies from BI designs. Columns show histograms of x_{100} and x_{400}, respectively.

(b) Predicted best dose frequencies from PAD designs with cost $c = 0.10$. Columns show histograms of least squares estimates $\hat{x}^*_{100}$ and $\hat{x}^*_{400}$.

Figure 5: Predicted best dose predictions under the continuous quadratic model (14) with $\theta_1 = 1.0$, $\theta_2 = 0.0$ and $\theta_3 = -1.0$; $x^* = 0.0$. Histograms in the top and bottom rows had start–up cohort sizes of 3 and 12, respectively, equally allocated to 0 and ± 1

and the convergence of the least squares estimator of $\hat{\theta}_{3n}$ and subsequently of $\hat{x}^*_n(\boldsymbol{\theta}) = -\hat{\theta}_{2n}/2\hat{\theta}_{3n}$ to the true value $x^*(\boldsymbol{\theta})$ cannot be guaranteed, cf. Lai
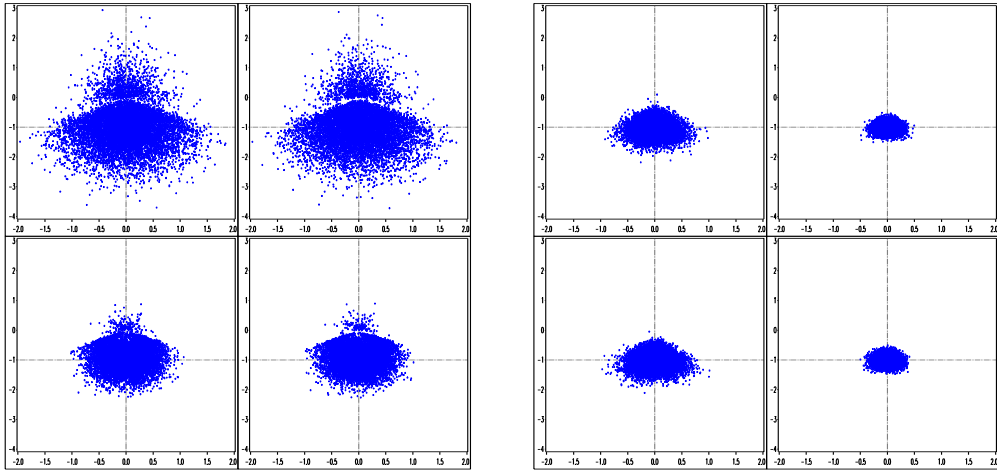
(a) $\hat{\theta}_{2n}$ by $\hat{\theta}_{1n}$ from BI designs. Columns show estimates at $n = 100$ and $n = 400$, respectively.

(b) $\hat{\theta}_{2n}$ by $\hat{\theta}_{1n}$ from PAD designs with cost $c = 0.10$. Columns show estimates at $n = 100$ and $n = 400$, respectively.

Figure 6: Plots of $\hat{\theta}_{2n}$ by $\hat{\theta}_{1n}$ under the continuous quadratic model (14) with $\theta_1 = 1.0$, $\theta_2 = 0.0$ and $\theta_3 = -1.0$; $x^* = 0.0$. Plots in the top and bottom rows had start–up cohort sizes of 3 and 12, respectively, equally allocated to 0 and $-1, 0, 1$.

and Wei [46]. This provides an explanation (but not a proof) of the observed behavior of $\{\hat{x}_n^*\}$.

The existence of attractors for $\{\hat{x}_n^*\}$ was established by Bozin and Zarrop [7] for the Kalman–Bucy type of estimators, which are simpler to analyze than are the least squares estimators. Our Monte Carlo simulations verify the existence of attractors for BI estimates based on least squares method, and this is enough to assert that, in general, BI sequences $\{\hat{x}_n^*(\boldsymbol{\theta})\}$ do not converge to the best dose. Actually, the existence of those attractors may exacerbate problems with applications if an unaware practitioner, observing a "convergence" to the wrong dose, will make a prediction with unfounded confidence.

The introduction of forced perturbations in dose allocation is a popular remedy for BI procedures, cf. Bozin and Zarrop [7] and Pronzato [48]. For instance, for allocations

$$x_{n+1} = \hat{x}_n^* + z_n n^{-1/4},$$

20

where z can be any random variable symmetrically distributed around zero, $M_{33}(n) \sim \mathcal{C} + \ln n$ and consistency of $\hat{\tilde{\theta}}_{3n}$, and hence convergence of $\{\hat{x}_n^*(\boldsymbol{\theta})\}$, is secured; see Lai [41]. At the same time the sequence $\{x_n\}$ stays close to x^* keeping risk at relatively low level. However, we recommend the use of penalized adaptive designs to permit the experimenter to quantify needs and ethical/cost constraints.

In conclusion we would like to emphasize that for Type II problem:

- The naive BI designs do not work. Their convergence to wrong doses may lead to the false optimism of a practitioner that can result in false prediction of the best dose.

- The empirical tuning of the BI designs leads to the estimators that are inferior to what can be found for the penalized adaptive optimal designs.

- The penalized adaptive design machinery provides routine tools for quantifying trial objectives (through criteria of optimality) and allows for ethical and/or cost constraints through selection of risk/penalty functions.

- We add a rather counterintuitive remark: in many cases fully adaptive designs (updating after each observation) perform worse than two or three stage designs. See Dragalin, Fedorov and Wu [17], Fedorov, Wu and Zhang [50] and Hardwick and Stout [51] for comments and explanations of this phenomenon.

3.3   Binary Responses

The use of binary responses is a popular choice in clinical dose–response studies, even when continuous variables must be dichotomized to create them. To make the forthcoming exposition comparable with the what has been discussed so far, we introduce dichotomized versions of models (3) and (14). So Y is defined by one of these models and

$$Z = \begin{cases} 1 & \text{if } Y \geq c, \\ 0 & \text{if } Y < c. \end{cases} \tag{18}$$

For Type I and Type II problems, respectively, let the response functions be

$$\eta(x, \boldsymbol{\theta}) = P(Z = 1|x) = F(\theta_1 + \theta_2 x), \tag{19}$$

21

and

$$\eta(x, \boldsymbol{\theta}) = P(Z = 1|x) = F(\theta_1 + \theta_2 x + \theta_3 x^2), \tag{20}$$

where F is the cumulative distribution function of Y.

In our simulations, we assumed that F is the cumulative normal distribution with $\theta_1 = 0$ and $\theta_2 = 1$ and thus (19) and (20) are probit models. All parameters $\boldsymbol{\theta}$ coincide with the corresponding parameters from the previous section. ARM was defined by the sequence

$$x_{n+1} = \arg \min_{x \in \mathscr{X}} \left| F\left(\hat{\theta}_{1n} + \hat{\theta}_{2n} x\right) - 0.5 \right|. \tag{21}$$

The solution of (21) is $x_{n+1} = -\hat{\theta}_{1n}/\hat{\theta}_{2n}$ or one of the boundary points for continuous $\mathscr{X}$. For $\hat{\theta}_{1n}$ and $\hat{\theta}_{2n}$, we used regularized maximum likelihood estimators. Regularized maximum likelihood estimators were used to be able to use all runs including simulations producing a likelihood function with no unique minimum; see, for example, [29]. The regularization consisted of subtracting $\gamma \sum_{\alpha=1}^{m} (\theta_\alpha - \theta_{\alpha 0})^2$ from the loglikelihood, where $m = 2$ or 3 parameters when using models (19) and (20), respectively. In our simulations, we set $\theta_{\alpha 0}$ equal to the true value of θ_α which, of course, cannot be done in practice; the constant γ was set to to 0.0001 to mitigate the influence of regularization. The role of regularization was negligible when n was above $20 - 30$. We used the same penalty function as with continuous responses.

ARM (21) targets estimation of ED50 ($x^* = 0$), which is the easiest quantile to estimate. The cut off level c in (18) corresponds to the median response, where the loss of information due to dichotomization is the least; see [52]. For the probit model, the loss of information for any other quantile is greater than $1 - 2/\pi$.

Simulations summarized in Figures 7 and 8 for Type I best doses are analogous to those in sections (3.1) and (3.2). Sample sequences of BI predicted best doses (Figure 2(c)) show the persistence of sequences sticking to the boundaries. In Figure 7 one sees that even with the initial start–up cohort sizes doubled from what was used with the linear model, the problem of BI sequences sticking to the boundaries persists. Observe that the variability of BI predicted best doses is increased compared with all corresponding continuous scenarios. Increased variability in the BI parameter estimates is apparent in Figure 8. Increased variability is an unfortunate consequence of dichotomization; see Fedorov, Mannino and Zhang [52].
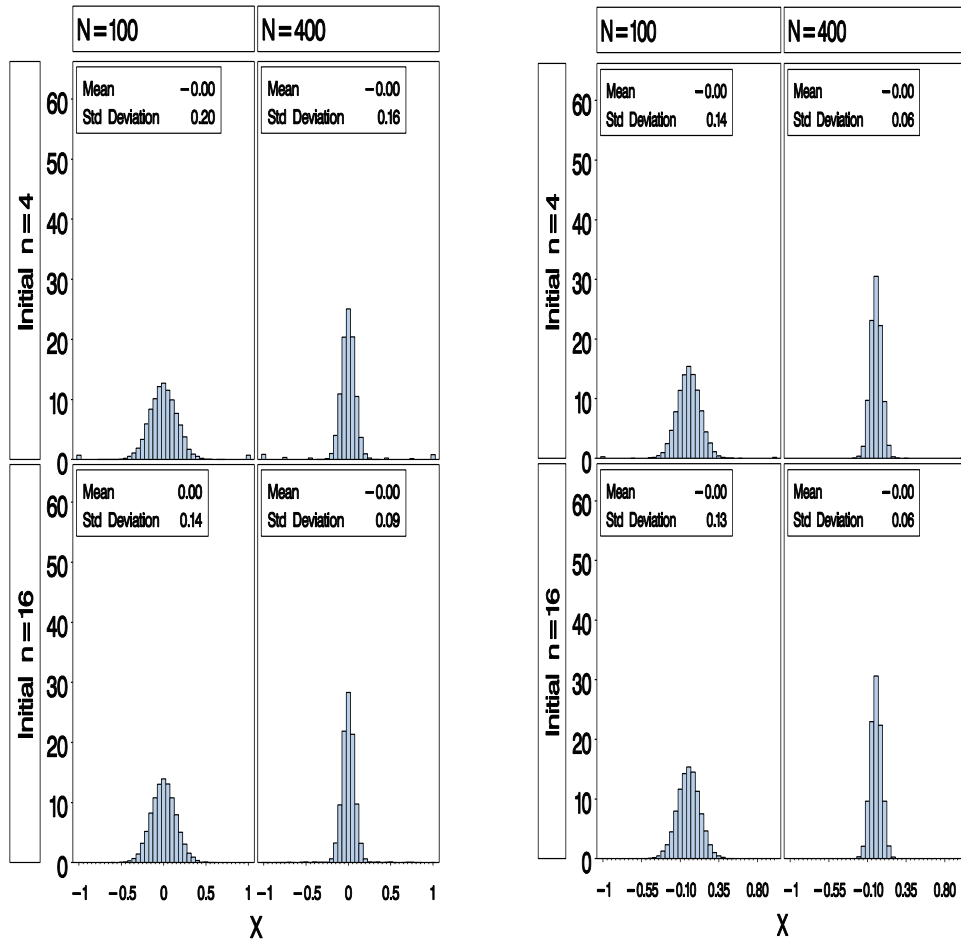
(a) Predicted best doses from BI designs. Columns show histograms of x_{100} and x_{400}, respectively,

(b) Predicted best doses from PAD designs with cost $c = 0.10$. Columns show histograms of least squares estimates $\hat{x}_{100}^*$ and $\hat{x}_{400}^*$, respectively.

Figure 7: Predicted best doses under the probit model with mean response function $F(\theta_1 + \theta_2 x)$ (18) with $\theta_1 = 0.0$ and $\theta_2 = 1.0$; $x^* = 0.0$. Histograms in the top and bottom rows had initial start–up cohorts of size 4 and 16, respectively, allocated equally to ± 1.

The BI procedure for seeking Type II best doses with mean function given
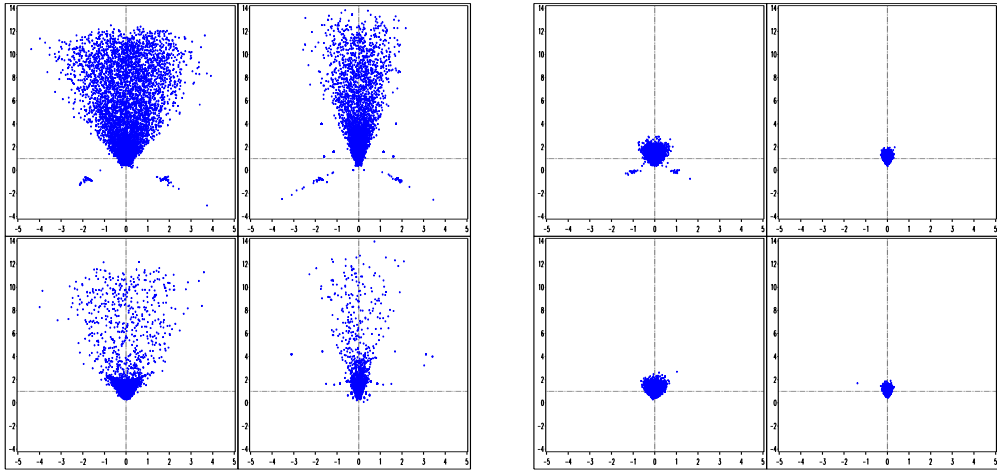
23

(a) $\hat{\theta}_{2n}$ by $\hat{\theta}_{1n}$ from BI designs. Columns show plots for at $n = 100$ and 400, respectively.
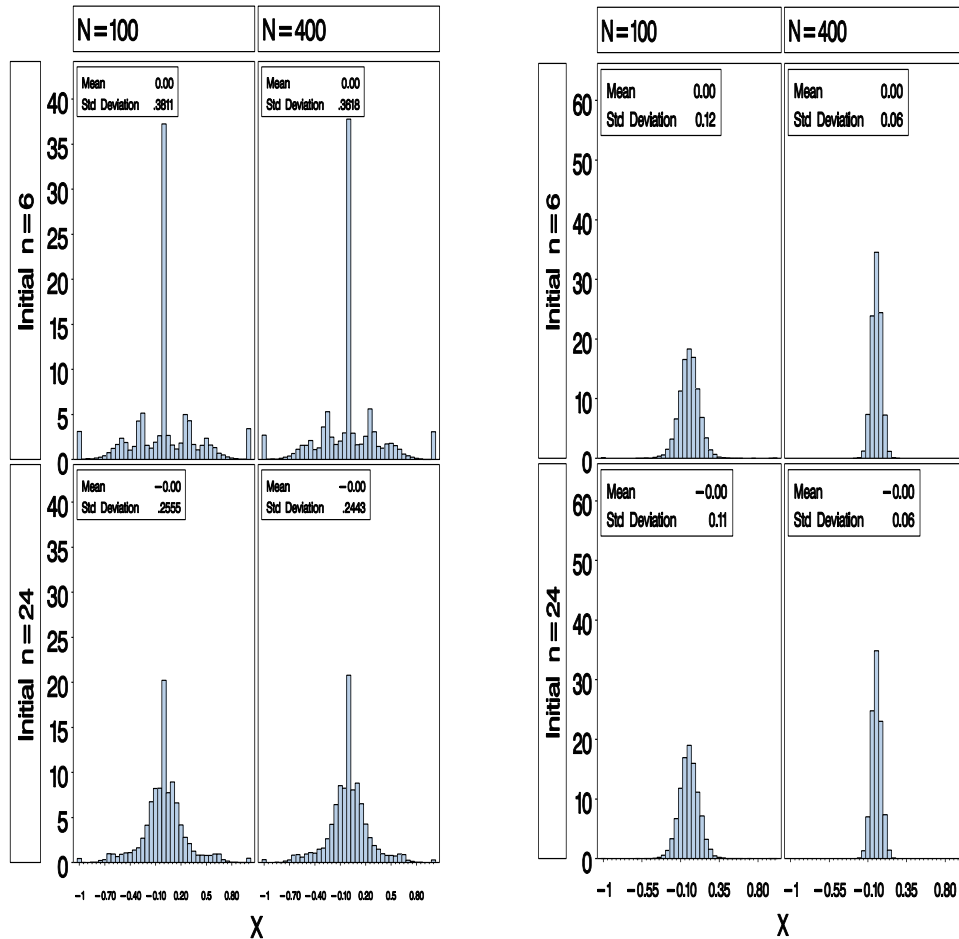
(b) $\hat{\theta}_{2n}$ by $\hat{\theta}_{1n}$ from PAD designs with cost $c = 0.10$ Columns show plots for at $n = 100$ and 400, respectively.

Figure 8: Plots of $\hat{\theta}_{2n}$ by $\hat{\theta}_{1n}$ under the probit model with mean response function $F(\theta_1 + \theta_2 x)$ (19, 18) with $\theta_1 = 0.0$ and $\theta_2 = 1.0$; $x^* = 0.0$. Rows have start–up cohort sizes of 3 and 12, respectively.

by (20) is

$$x_{n+1}(\boldsymbol{\theta}) = \arg\max_{x \in \mathscr{X}} F(\theta_1 + \theta_2 x + \theta_3 x^2) = \begin{cases} \tilde{x}_n & \text{if } |\tilde{x}_n| \leq 1, \\ \\ \pm 1 & \text{if } \tilde{x}_n > 1 \text{ or } < -1, \text{ respectively,} \end{cases}$$
(22)

where $\tilde{x}_n = -\hat{\theta}_{2n}/2\hat{\theta}_{3n}$. So results can be compared with those in section (3.2), F is the cumulative normal distribution; $c = 0.10$, $\theta_1 = 1.0$, $\theta_2 = 0.0$ and $\theta_3 = -1.0$. Sample BI predicted best dose sequences are shown in Figure 2(d). As under the continuous quadratic model, the persistent problem of BI sequences converging to the wrong values is evident. Simulations for Type II best doses analogous to those described in Section 3.2 are summarized in Figures 9 and 10. In the histograms of BI predicted best dose sequences in Figure 9(a), one sees a sizable proportion of sequences at x^* for $n = 100$ with minimal improvement when n increases to 400; the standard deviation with start–up sample sizes of 6 decreases from 0.38 for $\{x_{100}\}$ to 0.36 for $\{x_{400}\}$. In contrast, considerable reduction in variability is seen for PAD predicted best dose sequences in Figure 9(b) with the standard de-

24

(a) Predicted best doses of x_{100} and x_{400}, respectively, from BI designs.

(b) Predicted best doses from PAD designs with cost $c = 0.10$. Columns show histograms of least squares estimates $\hat{x}^*_{100}$ and $\hat{x}^*_{400}$

Figure 9: Predicted best doses under the probit model with mean response function $F(\theta_1 + \theta_2 x + \theta_3 x^2)$ (18, 20) with $\theta_1 = 1.0$, $\theta_2 = 0.0$ and $\theta_3 = -1.0$; $x^* = 0.0$. Histograms in the top and bottom rows had start–up cohorts of size 4 and 16, respectively, equally allocated to $-1, 0, 1$.

viation halving from 0.12 to 0.06 as n increases from 100 to 400. Similar comparisons hold for start–up sample sizes of 24. The changes in standard

(a) $\hat{\theta}_{2n}$ by $\hat{\theta}_{1n}$ from BI designs. Columns show plots for $n = 100$ and 400, respectively.

(b) $\hat{\theta}_{2n}$ by $\hat{\theta}_{1n}$ from PAD designs with cost $c = 0.10$. Columns show plots for $n = 100$ and 400, respectively.
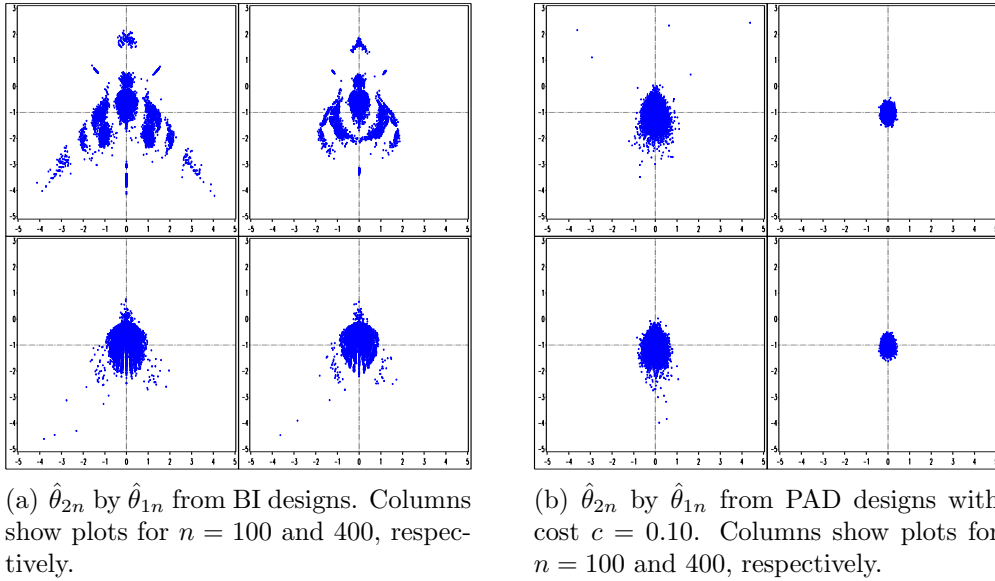
Figure 10: Parameter estimates $\hat{\theta}_{2n}$ by $\hat{\theta}_{1n}$ under the probit model with mean response function $F(\theta_1 + \theta_2 x + \theta_3 x^2)$ (18, 20) with $\theta_1 = 1.0$, $\theta_2 = 0.0$ and $\theta_3 = -1.0$; $x^* = 0.0$. Rows have initial start–up cohort sizes of 3 and 12.

deviations of predicted best doses under the continuous quadratic model and the dichotomized quadratic model are similar for various scenarios for both BI and ARM designs. However, the shapes of the histograms of predicted best doses for BI designs are very different under the continuous quadratic model and the dichotomized quadratic model, while they are similar for PAD designs.

The distribution of parameter estimates from BI designs shown in Figure 10(a) has many clusters caused from dichotomizing responses, particularly for $n = 100$. For all scenarios, distributions from BI designs are rather exotic compared with distributions from PAD designs, which are shown in Figure 10(b).

We emphasize that the simulations look worse with dichotomized response than for continuous responses, mainly, for two reasons due to the dichotomization: (1) loss of information and (2) discreteness of parameter estimators. The discreteness of parameter estimators is mainly noticeable for small sample sizes; All patterns of "bad" behavior of naive ARM in the linear case and the extremely "bad" behavior for BI designs in the quadratic

case persist with dichotomization. Unfortunately, nothing is better for more practical models than it is for the simplest ones.

4 Discrete design region $\mathscr{X}$

In clinical dose–finding studies, allocation rules that were adapted from numeric optimization, optimal control or optimal design usually must be modified to operate on a lattice of doses. For instance, in the one dimensional case, $\mathscr{X} = \{x_1, \ldots, x_K\}$. With a discrete design region $\mathscr{X}$, it is useful to distinguish the design region from the prediction space $\tilde{\mathscr{X}}$, because typically finding a best dose will require interpolation between the doses in $\mathscr{X}$. For example, allocations may be sampled from the given doses of currently manufactured pills. Alternatively, one may only be able to study a few "doses" of some drug, yet seeks the best dose over a continuous range that includes these doses. Usually $\tilde{\mathscr{X}}$ will include $\mathscr{X}$.

4.1 Type I Dose–finding

Almost immediately after the pioneering papers on the Robbins–Munro procedure appeared, it was noted that some modifications are needed to make it work for discrete $\mathscr{X}$. The introduction of additional randomization, intrapolation/extrapolation from $\mathscr{X}$ to $\tilde{\mathscr{X}}$, and changing doses not farther than one step away from the current allocation were found useful in transporting major Robbins–Monro ideas to discrete sample spaces. Various amendments and improvements were extensively discussed in the dose–finding community, and many references can be found in [12].

With regard to naive Type I BI designs, Azriel, Mandel and Rinott [53] recently proved that there exists no such design that converges to x^* for all increasing response functions. Our simulations confirm the need of at least some corrective measures with BI procedures. Indeed, all patterns seen for continuous $\mathscr{X}$ are observed for the straightforward generalization of ARM (compare with (9)):

$$x_{n+1} = \arg\min_{x \in \mathscr{X}} |\zeta(x, \hat{\boldsymbol{\theta}}_n) - \zeta^*|. \tag{23}$$

We make a few recommendations to improve the discrete version of ARM:

- Note that in the case of the sparse grid, instead of using the final allocation x_N as a recommended dose, one can use

$$\tilde{x}_N = \arg \min_{x \in \tilde{\mathscr{X}}} |\zeta(x, \hat{\boldsymbol{\theta}}_N) - \zeta^*|. \tag{24}$$

  In practice, when the validity of $\zeta(x, \boldsymbol{\theta})$ is in doubt for the whole set $\tilde{\mathscr{X}}$, local approximations with any simple function (e.g., linear, sigmoid shaped or any of functions used in the numerous CRM papers) will work; compare with Dupa and Herkenrath [54]. For studies with small samples sizes, Stylianou and Flournoy [55] recommend using isotonic regression with linear interpolation from $\mathscr{X}$ to $\tilde{\mathscr{X}}$

- Introduce extra randomization, for instance, as in Derman [19], Durham and Flournoy [21],[56] or Dupa and Herkenrath [54]. It is crucial to avoid long "runs" at same dose that can falsely be viewed as convergence. This is especially important for binary responses. This also can be achieved by treating subjects in cohorts (Gezmu and Flournoy [57]) or by sampling at a dose until some k successive toxicities are observed (Oron and Hoff [58]). These procedures can also be used to start–up parametric procedures.

- New allocations should be to one of the neighboring points of the current x_n. This stabilizes the sequence in that only the sign of $\hat{\theta}_n$ is important. This restriction also is usually desired by clinical researchers, but our recommendation follows directly from analysis of the procedures.

- We recommend using PAD to get information about local behavior of dose-response function near x^*. While using PAD, one may restrict allocations to doses neighboring the current one. We do not recommend c–optimal adaptive design, even though it coincides with ARM for the linear model and also with typically recommended local models like the CRM, because it does not provide good information about local behavior of dose-response function near x^*.

4.2   Type II Dose–finding

As for Type I dose finding, all patterns observed for continuous $\mathscr{X}$ hold for Type II objectives. Defining dose–finding problems in terms of the two sets $\mathscr{X}$ and $\tilde{\mathscr{X}}$ is even more beneficial here than for Type I objectives. It allows

a better understanding of the structure of $\zeta(x, \boldsymbol{\theta})$ near the optimal dose. All concluding recommendations from Section 4.1 hold, but more pronounced emphasis on extra randomization is required.

5   Conclusions

This paper is an attempt to attract the attention of statisticians, who design and analyze dose–finding trials, to potential flaws in some gaining popularity adaptive designs, which we dubbed here as "best intention" designs. We considered two types of them: the search of dose with a response equal or close to some predetermine value (Type I problem, the search of MTD is the most frequent case) and the determination of the dose with maximal value of a utility function (Type II problem, maximizing the probability of cure). In both cases, best intention designs allocate the next subject (cohort) to the dose that is the best one accordingly to the current knowledge. Our extensive Monte Carlo simulations, together with some unfortunately partially neglected theoretical results developed in the framework of optimal control theory, show that the reckless application of ethically very sound and attractive idea may lead to very dismal results, especially for Type II problem. While various remedies are readily available for Type I dose-finding and they lead to the procedures that are very close to statistically optimal, the situation with Type II dose-finding is much worse.

We found that the systematic use of the machinery of the optimal design theory and in particular penalized adaptive (multistage) designs leads to designs that for Type I are either equivalent to the popular BI designs (like ARM or its younger sibling CRM) or superior to them. For Type II all simulation confirm superiority of adaptive penalized (either $\mathscr{D}-$ or c–) optimal designs.

6   Appendix: Some Notations from Optimal Design

Likelihood estimators. The exposition is done for the maximum likelihood estimators. For the least squares method, almost all formulae look identical with the the moment matrix (aka information matrix) replacing the Fisher

information matrix. We use notations mainly following Fedorov and Hackl [59].

Let
$$\mathbf{Y} \ \sim \ p\,(\mathbf{Y}|\mathbf{x}, \boldsymbol{\theta})$$

be our working model. The sequence $\xi_N = \{\mathbf{x}_i\}_1^N$ is called a design (of an experiment). When there is a need to emphasize that at K points $\mathbf{x}_i$ there are n_i replicated observations, we write $\xi_N = \{\mathbf{x}_i, w_i\}_1^K, \ \ w_i = n_i/N, \ \ N = \sum_{i=1}^K n_i$
. It will be clear from the context which is used. In optimal design theory, the major results are derived for "continuous" designs, i.e. designs where weight w_i could be any value in $[0, 1]$. In this case, we use $\xi = \{\mathbf{x}_i, w_i\}_1^K, \ \sum_{i=1}^K = 1$. More generally ξ can be any probability measure defined on $\mathscr{X}$. We associate with every design a (normalized) penalty function $\Phi(\xi)$:

$$\Phi(\xi) = \sum_{i=1}^K w_i \phi(\mathbf{x}_i, \boldsymbol{\theta}) = N^{-1} \sum_{i=1}^K n_i \phi(\mathbf{x}_i, \boldsymbol{\theta}).$$

If design ξ does not depend on the observed values of the response Y, then under rather mild conditions, the maximum likelihood estimator

$$\hat{\boldsymbol{\theta}}_N = \arg\max_{\boldsymbol{\theta} \in \Omega} \prod_{i=1}^K \prod_{j=1}^{n_i} p(\mathbf{y}_{ij}|\mathbf{x}_i, \boldsymbol{\theta})$$

is asymptotically normally distributed, i.e.

$$\sqrt{N}(\hat{\boldsymbol{\theta}}_N - \boldsymbol{\theta}) \sim \mathcal{N}(0, D(\xi, \boldsymbol{\theta})) \ , \tag{25}$$

where
$$D(\xi, \boldsymbol{\theta}) = M^{-1}(\xi, \boldsymbol{\theta})$$

and
$$NM(\xi, \boldsymbol{\theta}) = \sum_{i=1}^K n_i \mu(\mathbf{x}_i, \boldsymbol{\theta}) = N \sum_{i=1}^K w_i \mu(\mathbf{x}_i, \boldsymbol{\theta}).$$

The matrix

$$\mu(\mathbf{x}, \boldsymbol{\theta}) = \mathrm{E}\{\ell(\mathbf{Y}|\mathbf{x}, \boldsymbol{\theta})\ell^T(\mathbf{Y}|\mathbf{x}, \theta)\}, \quad \ell(\mathbf{Y}|x, \boldsymbol{\theta}) = \frac{\partial}{\partial \boldsymbol{\theta}} \log p(\mathbf{Y}|x, \boldsymbol{\theta})$$

describes the information that can be gained if an observation is performed at $\mathbf{x}$. For the asymptotical validity of (25) in adaptive design setting, stronger

assumptions are needed; compare with [41]. But necessary assumptions hold for cases addressed in this article.

A penalized optimal design. A penalized optimal design can be defined as

$$\xi^* = \arg \min_{\xi \in \Xi(\mathscr{X})} \Psi(NM(\xi, \theta)), \quad \text{s.t. } N(\xi)\Phi(\xi) \leq C,$$

where $\Xi(\mathscr{X})$ is a set of all possible designs on $\mathscr{X}$. For continuous designs this problem is equivalent to

$$\xi^* = \arg \min_{\xi \in \Xi(\mathscr{X})} \Psi\left(\frac{M(\xi, \theta)}{\Phi(\xi)}\right). \tag{26}$$

In this article $\phi(x, \boldsymbol{\theta}) = (x - x^*)^2 + c$. The examples of optimality criteria Ψ can be found in Fedorov and Hackl [59], Atkinson, Donev and Tobias [60], and many other works on optimal design. For the sake of simplicity, we stay with a popular $\mathscr{D}$–criterion in which case a necessary and sufficient condition for optimality is

$$\text{tr}\left[\boldsymbol{\mu}(\mathbf{x}, \boldsymbol{\theta})M^{-1}(\xi^*, \boldsymbol{\theta})\right] \leq m\varphi(\mathbf{x})/\Phi(\xi^*).$$

For the "normal" linear regression $\eta(\mathbf{x}, \boldsymbol{\theta}) = f^T(\mathbf{x})\boldsymbol{\theta}$, the latter becomes

$$\sigma^{-2}\text{Var}\left[\eta(\mathbf{x}, \hat{\boldsymbol{\theta}})\right] \leq m\phi(\mathbf{x})/N\Phi(\xi^*).$$

An adaptive penalized optimal design. Adaptive design, which is needed only for optimal designs when (26) or the utility function involves $\boldsymbol{\theta}$, is generated by

$$\mathbf{x}_{n+1} = \arg \max_{\mathbf{x} \in \mathscr{X}} \left\{\text{tr}\left[\boldsymbol{\mu}(\mathbf{x}, \boldsymbol{\theta})M^{-1}(\xi_n, \hat{\boldsymbol{\theta}}_n)\right] - m\phi(\mathbf{x})/\Phi(\xi_n, \hat{\boldsymbol{\theta}}_n)\right\}. \tag{27}$$

The latter can be viewed as a "forward" first order numerical procedure of building locally optimal designs with the true $\boldsymbol{\theta}$ replaced by the current MLE estimator $\hat{\boldsymbol{\theta}}_n$, compare with [26],[50].

Relation with regularized optimization. Note that for the Type II dose–finding and for quadratic regression, (26) considered in all simulations is equivalent to the design problem with $\phi(x, \boldsymbol{\theta}) = \zeta^* - \zeta(\mathbf{x}, \boldsymbol{\theta}) + c'$. It is different from

31

what Pronzato [8] proposed to address the same problem. In the framework of convex optimal design theory, the closest adaptive design can be derived using the following modified $\mathscr{D}$–criterion:

$$\Psi(\xi) = \int_{\mathscr{X}} \zeta(\mathbf{x}, \boldsymbol{\theta}) \xi(dx) + \gamma \left| M(\xi, \boldsymbol{\theta}) \right|^{1/m}. \tag{28}$$

Note that $\max_\xi \int_{\mathscr{X}} \zeta(\mathbf{x}, \boldsymbol{\theta}) \xi(dx) = \max_\xi \zeta(\mathbf{x}, \boldsymbol{\theta})$. One can verify that (28) generates the following adaptive design:

$$\mathbf{x}_{n+1} = \arg \max_{x \in \mathscr{X}} \left[\zeta(\mathbf{x}, \hat{\theta}_n) + \alpha(\xi_n) \left(n^{-1} \mathrm{Var}(\eta(\mathbf{x}, \hat{\boldsymbol{\theta}}_n)) - m \right) \right],$$

where $\alpha(\xi_n) = \gamma m^{-1} |M(\xi_n, \hat{\boldsymbol{\theta}}_n)|^{1/m}$.

References

[1] Wetherill G. Sequential estimation of quantal response curves. *Royal Statistical Society B* 1963; **25**:1–48.

[2] Lai TL, Robbins H. Adaptive design in regression and control. *Proceedings of the National Academy of Sciences of the United States of America* 1978; **75**:586–587.

[3] O'Quigley J, Pepe M, Fisher L. Continual reassessment method: A practical design for Phase 1 clinical trials in cancer. *Biometrics* 1990; **46**:33–48.

[4] Li Z, Durham SD, Flournoy N. An adaptive design for maximization of a contingent binary response. *Adaptive Designs*, Flournoy N, Rosenberger WF (eds.), Institute of Mathematical Statistics, 1995; 179–196.

[5] Thall PF, Cook JD. Dose-finding based on efficacy?Toxicity trade-offs. *Biometrics* 2004; **60**(3):684–693.

[6] Lai TL, Robbins H. Iterated least squares in multiperiod control. *Advances in Applied Mathematics* 1982; **3**(1):50–73.

[7] Bozin A, Zarrop M. Self-tuning extremum optimizer convergence and robustness. *ECC* 1991; **91**:672–677.

[8] Pronzato L. Adaptive optimization and D-optimum experimental design. *The Annals of Statistics* 2000; **28**(6):1743–1761.

[9] Chang HH, Ying Z. Nonlinear sequential designs for logistic item response theory models with applications to computerized adaptive tests. *The Annals of Statistics* 2009; **37**(3):1466–1488.

[10] Oron AP, Azriel D, Hoff PD. Dose–finding designs: The role of convergence properties. *The International Journal of Biostatistics* in press; **7**(1). URL `http://www.bepress.com/ijb/vol7/iss1/39`.

[11] Azriel D. A note on the robustness of the continual reassessment method. *preprint* 2011; .

[12] Chevret S ((ed.)). *Statistical Methods for Dose-Finding Experiments.* John Wiley & Sons, Ltd., 2006.

[13] Gooley TA, Martin PJ, Lloyd DF, Pettinger M. Simulation as a design tool for phase i/ii clinical trials: An example from bone marrow transplantation. *Controlled Clinical Trials* 1994; **15**:450–460.

[14] Fan SK, Chaloner K. Optimal designs and limiting optimal designs for a trinomial response. *Journal of Statistical Planning and Inference* 2004; **126**(1):347–360.

[15] Rabie H, Flournoy N. *mODa 7: Advances in Model–Oriented Design and Analysis*, chap. Optimal designs for contingent responses models, pp. 133–142. Physical–Verlag, Heidelberg. eds. n Di Bucchianico, A.; L 2004.

[16] Dragalin V, Fedorov V. Adaptive designs for dose-finding based on efficacy–toxicity response. *Journal of Statistical Planning and Inference* 2006; **136**:1800–1823.

[17] Dragalin V, Fedorov V, Wu Y. Two–stage design for dose–finding that accounts for both efficacy and safety. *Statistsics in Medicine* 2008; **27**:5156–5176.

[18] Fuller D, Scherer W. The desirability function: underlying assumptions and application implications. *IEEE* 1998; :4016–4021.

[19] Derman C. Non-parametric up-and-down experimentation. *Annals of Mathematical Statistics* 1957; **28**:795–798.

[20] Dixon WJ, Mood A. A method for obtaining and analyzing sensitivity data. *Journal of the American Statistical Association* 1948; **13**:109–126.

[21] Durham SD, Flournoy N. *Statistical Decision Theory and Related Topics, V*, chap. Random Walks for Quantile Estimation, 467–476. Springer-Verlag Inc., eds. Gupta, Shanti S. and Berger, James O., 1994.

[22] Robbins H, Monro S. A stochastic approximation method. *Ann. Math. Statistics* 1951; **22**:400–407.

[23] von Békésy G. A new audiometer. *Archives of Otolaryngology* 1947; **35**:411–422.

[24] Wetherill GB, Glazebrook KD. *Sequential Methods in Statistics*. Chapman & Hall Ltd, 1986.

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

[26] Dragalin V, Fedorov V, Wu Y. Adaptive designs for selecting drug combinations based on efficacy-toxicity response. *Journal of Statistical Planning and Inference* 2008; **2**:352–373.

[27] Durham SD, Flournoy N, Li W. A sequential design for maximizing the probability of a favourable response. *Canad. J. Statist.* 1998; **26**(3):479–495.

[28] Fedorov V, Mueller W. Comparison of two approaches in the optimal design of an observation network. *Statistics* 1989; **20**(3):339–351.

[29] Fedorov V, Wu Y. Dose finding designs for continuous responses and binary utility. *Journal of Biopharmaceutical Statistics* 2007; **17**:1085–1096.

[30] Hardwick J, Stout QF. *Optimum Design 2000*, chap. Optimizing a unimodal response function for binary variables. Kluwer, eds. A. Atkinson and B. Bogacka and A. Zhigljavsky, 2001; 195–208.

[31] Kiefer J, Wolfowitz J. Stochastic estimation of the maximum of a regression function. *The Annals of Mathematical Statistics* 1952; **23**(3):462–466.

[32] Kpamegan E, Flournoy N. *Optimum Design 2000*, chap. An optimizing up-and-down design, pp. 211–224. Kluwer Academic Publishers: Dordrecht. eds. A. Atkinson, B. Bogacka, A. Zhigljavsky, 2001.

[33] Pronzato L. Penalized optimal designs for dose-finding. *Journal of Statistical Planning and Inference* 2010; **140**(1):283–296.

[34] Le Cam LM. *Asymptotic methods in statistical decision theory.* Springer–Verlag, 1986.

[35] Pratt JW, Raffa H, Schlaife R. *Introduction to statistical decision theory.* The MIT Press, 1995.

[36] Fedorov VV, Liu T. *Encyclopedia of Clinical Trials (D'Agostino R, Editor in Chief)*, chap. Enrichment Design. Wiley, 2008.

[37] Fedorov V, Liu T. Randomized discontinuation trials: Design and efficiency. *GSK BDS Technical Report* 2005; **2005**.

[38] Sacks J. Asymptotic distribution of stochastic approximation procedures. *The Annals of Mathematical Statistics* 1958; **29**:373–405.

[39] Anbar D. A stochastic Newton–Raphson method. *Journal of Statistical Planning and Inference* 1978; **2**:153–163.

[40] Ypma TJ. Historical development of the Newton-Raphson method. *SIAM Review* 1995; **37**:531–551.

[41] Lai TL. Sequential analysis: Some classical problems and new challenges (Pkg: P303-408). *Statistica Sinica* 2001; **11**(2):303–351.

[42] Fedorov V. *Theory of Optimal Experiments.* Academic Press: New York, 1972.

[43] Hinkley DV. On the ratio of two correlated normal random variables. *Biometrika* 1969; **56**(3):635–639.

[44] Bartroff J, Lai TL. Approximate dynamic programming and its application to the design of phase i cancer trials. *Statistical Science* 2010; **25**(2):245–257.

[45] Lai TL. Stochastic approximation. *The Annals of Statistics* 2003; **31**(2):391–406.

[46] Lai TL, Wei CZ. Least squares estimates in stochastic regression models with applications to identification and control of dynamical systems. *The Annals of Statistics* 1982; **10**:154–166.

[47] Anderson TW, Taylor JB. Some experimental results on the statistical properties of least squares estimates in control problems. *Econometrica* 1976; **44**(6):1289–1302.

[48] Pronzato L. Optimal experimental design and some related control problems. *Automatica* 2008; **44**:303–325.

[49] Andrews GE, Askey R, Roy R. *Special functions*. Encyclopedia of Mathematics and Its Applications, The University Press, Cambridge, 1999.

[50] Fedorov V, Wu Y, Zhang R. Dose-finding experiments with discrete and continous responses. *Statistics in Medicine* 2011; **in press**.

[51] Hardwick J, Stout QF. Optimal few-stage designs. *Journal of Statistical Planning and Inference* 2002; **104**(1):121–145.

[52] Fedorov V, Mannino F, Zhang R. Consequences of dichotomization. *Pharmaceutical Statistics* 2009; **8**:50–61.

[53] Azriel D, Mandel M, Rinott Y. The treatment versus experimentation dilemma in dose finding studies. *J. Statist. Plann. Inference* 2011; **141**(8):2759–2768.

[54] Dupa V, Herkenrath U. Stochastic approximation on a discrete set and the multi–armed bandit problem. *Communications in Statistics. Part C: Sequential Analysis* 2007; **1**(1):1–25.

[55] Stylianou M, Flournoy N. Dose finding using the biased coin up-and-down design and isotonic regression. *Biometrics* 2002; **58**(1):171–177.

[56] Durham SD, Flournoy N. Up-and-down designs I: Stationary treatment distributions. *Adaptive Designs*, Flournoy N, Rosenberger WF (eds.), Institute of Mathematical Statistics, 1995; 139–157.

[57] Gezmu M, Flournoy N. Group up-and-down designs for dose-finding. *Statistical Planning Inference* 2006; **136**(6):1749–1764.

[58] Oron AP, Hoff PD. The k-in-a-row up-and-down design, revisited. *Statistics in Medicine* 2009; **28**:1805–1820.

[59] Fedorov VV, Hackl P. *Model-oriented Design of Experiments*. Springer-Verlag Inc, 1997.

[60] Atkinson AC, Donev AN, Tobias R. *Optimum Experimental Designs with SAS*. Oxford Statistical Science Series, 34, Oxford University Press, 2007.