Journal of Statistical Research *https://doi.org/10.3329/jsr.v55i2.58808*
2021, Vol. 55, No. 2, pp. 313-334 *ISSN 0256 - 422 X*

D-OPTIMAL DESIGNS FOR DOSE FINDING IN PHASE I CLINICAL TRIALS

HASAN MAHMUD SUJAN

Health Systems and Population Studies Division
International Centre for Diarrhoeal Disease Research, Bangladesh

Email: hasan.sujan@icddrb.org

M. IFTAKHAR ALAM*

Institute of Statistical Research and Training
University of Dhaka, Dhaka 1000, Bangladesh

Email: iftakhar@isrt.ac.bd

SUMMARY

Determining the maximum tolerated dose (MTD) is the main challenge of phase I clinical trials. There are many methods in the literature to determine the MTD. The D-optimal design can also be used to find the MTD. The D-optimal design depends on the Fisher information matrix (FIM), and it minimizes the generalized variance of the parameter estimates. However, the D-optimal design is yet to receive much attention from clinicians. Since a dose-response model is usually non-linear, the FIM depends on the unknown model parameters. To optimize the FIM through the D-criterion, values need to be assumed for the model parameters. This paper focuses on investigating four different D-optimal designs depending on parameter values: design based on posterior Bayes estimators, design based on maximum likelihood estimators, sequential Bayesian design and two-stage Bayesian design. Six plausible dose-response scenarios and a real scenario are investigated through a simulation study. Except for the D-optimal design that utilizes maximum likelihood estimates in FIM optimization, all other D-optimal designs are found very competitive for the correct MTD recommendation. The D-optimal designs are also compared with an A-optimal design. The performance of A-optimal design is not attractive as these designs. Because of its numerical simplicity compared to the others, the posterior-based D-optimal design is recommended for dose-finding in phase I clinical trials.

Keywords and phrases: Phase I clinical trials; maximum tolerated dose; adaptive design; D-optimal design; Bayesian D-optimal design

AMS Classification: 92B15

* Corresponding author

1 Introduction

Clinical trials explore whether a medical strategy, treatment, or device is safe and effective for humans. They are commonly divided into four phases: phase I, phase II, phase III, and phase IV. Phase I is the first for testing in humans and designed to assess the safety, tolerability, and pharmacokinetics of a drug. A phase I clinical trial tries to explore the dose-limiting toxicity (DLT) of a drug. The goal is to find the dose with a probability of DLT that patients can maximally bear and is known as the MTD. There are two classes of designs for phase I trials: rule-based and model-based. Standard rule-based designs assign new patients to dose levels according to prespecified rules. They do not stipulate any assumption regarding the dose-toxicity curve. An alternative dose escalation method uses statistical models that actively seek a dose level that produces a prespecified probability of DLT using information from all the enrolled patients. This method is conveniently carried out under the Bayesian framework. These model-based designs use all of the available data to model the dose-toxicity curve. Conventional model-based designs include the continual reassessment method (CRM), escalation with overdose control, and time-to-event monitoring CRM, etc.

The CRM is the first Bayesian model-based method proposed for phase I trials (O'Quigley et al., 1990). The initial estimate of parameters required for this method is generally elicited from the experts familiar with the preclinical data or who have experience with similar drugs, if any exists. Although this initial estimate may not be accurate, it guides dose escalation. In the original description of the CRM, all patients are treated at the dose thought to be closest to the MTD, which corresponds to the dose at the target toxicity level. The estimation of the probability of encountering a DLT is updated for each new patient who enters the study at any dose level until a prespecified condition is met, at which point the trial is stopped. The trial continues until a fixed sample size of n is achieved. The original CRM allowed for multiple-dose escalations and de-escalations. However, the original CRM version received considerable debate in the statistical literature since it starts with the initial MTD. Many patients are also likely to be exposed to high toxicity because of skipping a dose. As a result, many modifications have been proposed by O'Quigley and Chevret (1991), Korn et al. (1994), Faries (1994), Goodman et al. (1995), Møller (1995), Piantadosi et al. (1998), and Heyd and Carlin (1999).

The D-optimal design, a model-based design, can also be used to find the MTD. Chaloner and Verdinelli (1995) presented a unified view of Bayesian experimental design by putting experimental design in a decision-theoretic framework. This framework justifies many optimality criteria, including the D-optimality criterion, and opens new possibilities for implementing D-optimal designs. For example, Haines et al. (2003) presented a broad approach to the design of phase I trials for the efficient estimation of the MTD. They constructed constrained Bayesian c- and D-optimal designs. The imposed constraint incorporates the optimal design points and their weights, confirming that the probability that an administered dose exceeds the maximum acceptable dose is low. They considered both log doses on the real line and discrete doses in practical consideration. Their exploration was mainly in a sequential Bayesian D-optimal design discussed in Section 2.2.4. They found the design effective and efficient for estimating model parameters and protecting patients from highly toxic doses. Alam (2016) compared the continual reassessment method with the D-optimal design.

In this article, we engage D-optimal designs to find the MTD. The design maximizes the determinant of the FIM or equivalently minimizes the determinant of the inverse of the FIM. That is, it minimizes the generalized variance of the parameter estimates. The A-optimality minimizes the total variance of the parameter estimates. The c-optimality is different from the previous criteria because here interest is in the estimation of a linear combination of the model parameters with minimum variance. Since interest is on the variance of the parameter estimates only, we are using the D-optimality criterion. The Cramér-Rao inequality tells that the variance of an unbiased estimator is greater than or equal to the inverse of the FIM. Also, the maximum likelihood estimators' variance approaches the inverse of the FIM for large n. Therefore, by minimizing the inverse of the FIM's determinant, we minimize the asymptotic variance of the estimates of the model parameters. It is well known that the FIM for a non-linear model depends on the unknown model parameters (White, 1973). We can begin with a guess about the parameters, but with the advancement of a trial, the up-to-date estimates of the parameters can be used.

This paper looks at four different D-optimal designs based on parameter values. The first design we consider is a posterior mean-based D-optimal design. The second one is based on the maximum likelihood estimates (MLE) of the model parameters. Instead of using single values as in the previous two designs, the third and fourth ones use the prior distribution of the model parameters. The latter two designs are Bayesian D-optimal designs. An MLE-based A-optimal design is also incorporated to assess the D-optimal designs. The paper attempts to help the clinicians choose a D-optimal design out of many. It is organized as follows. Section 2 describes the designs for comparison. The simulation setup is presented in Section 3. Section 4 provides the numerical results. Implementation of the designs based on a real study is included in Section 5. Finally, a discussion appears in Section 6.

2 Methods

A principal goal of phase I trials is to establish a dose to be tested in phase II trials. The guiding principle for dose escalation in phase I trials is to avoid unnecessary patient exposure to subtherapeutic doses of a drug while preserving safety and maintaining rapid accrual. In the following sections, we describe the dose-response model and dose-optimization criteria used in this paper.

2.1 Model

Assume that there are d ordered doses $\mathcal{X} = \{x^{(1)}, x^{(2)}, \ldots, x^{(d)}\}$ for an experimental drug based on the preclinical studies. Responses are often continuous measurements in phase I trials, and they are dichotomized for the convenience of modeling. To characterize the dose-response relationship for such a binary response, the following logistic model is used

$$\psi(x, \boldsymbol{\vartheta}) = \frac{\exp\left(\vartheta_1 + \vartheta_2 x\right)}{1 + \exp\left(\vartheta_1 + \vartheta_2 x\right)}, \tag{2.1}$$

where $\boldsymbol{\vartheta} = (\vartheta_1, \vartheta_2)$ is the vector of dose response parameters and x is the dose given to a patient. For simplicity, we present $\psi(x, \boldsymbol{\vartheta})$ by ψ. Let us define a binary random variable

$$R = \begin{cases} 1, & \text{if the patient experiences DLT;} \\ 0, & \text{otherwise,} \end{cases}$$

with probabilities $\Pr(R = 1) = \psi$ and $\Pr(R = 0) = 1 - \psi$. This R has a Bernoulli distribution with probability function as

$$f(r) = \psi^r (1 - \psi)^{1-r},$$

Then the likelihood function for the model in (1) is

$$L(\boldsymbol{\vartheta}|x, r) = \psi^r (1 - \psi)^{1-r},$$

and the log-likelihood function is

$$l(\boldsymbol{\vartheta}|x, r) = r \log(\psi) + (1 - r) \log(1 - \psi).$$

The score functions can be obtained as

$$U_1 = \frac{\partial l}{\partial \vartheta_1} = r - \psi,$$

and

$$U_2 = \frac{\partial l}{\partial \vartheta_2} = x(r - \psi).$$

Using the scores, we can easily obtain the FIM for a single patient as

$$\boldsymbol{I}(x, \boldsymbol{\vartheta}) = \begin{bmatrix} \psi(1 - \psi) & x\psi(1 - \psi) \\ x\psi(1 - \psi) & x^2\psi(1 - \psi) \end{bmatrix}. \tag{2.2}$$

This FIM is used in the construction of various D-optimal designs that we discuss below.

2.2 Design criteria

The construction of D-optimal designs for non-linear models depends on the unknown model parameters $\boldsymbol{\vartheta}$. As the first patient usually receives the lowest dose in a trial, it is possible to utilize the parameter estimates based on the available data to determine the D-optimal dose for the second patient. The parameters are re-estimated each time in an adaptive trial, and hence the up-to-date estimates can be used in the dose-optimization criterion. Using the posterior means of the parameters or maximum likelihood estimates will result in two designs: posterior-based D-optimal design and MLE-based D-optimal design. An alternative is to adopt a Bayesian approach to design by optimizing the average function of the information matrix over a prior distribution placed on the unknown parameters. This approach can be applied in two ways: sequential Bayesian D-optimal design and two-stage Bayesian D-optimal design.

A D-optimal design tends to choose doses from the extremes of the design region. As a result, there is a chance of allocating a patient to the highly toxic dose, which may cause harm. Hence, a more straightforward constraint on dose escalation proposed by Goodman et al. (1995) is used in our designs, increasing the dose by only one pre-specified level at a time.

2.2.1 Posterior-based D-optimal design

Assume that we are at the kth step of a trial, which means k patients have been treated with different doses from $\mathcal{X}$. Then $\boldsymbol{x}$ be a $k \times 1$ dose vector with components x_l, and let $\boldsymbol{r}$ be a $k \times 1$ outcome vector with r_l as the lth row ($l = 1, \ldots, k$) representing the outcome obtained from a patient. The likelihood function for the kth step is written as

$$L_k(\boldsymbol{\vartheta}|\boldsymbol{x}, \boldsymbol{r}) \propto \prod_{l=1}^{k} \Big\{\psi(x_l, \boldsymbol{\vartheta})\Big\}^{r_l} \Big\{1 - \psi(x_l, \boldsymbol{\vartheta})\Big\}^{1-r_l}.$$

The posterior means of the components of $\boldsymbol{\vartheta}$ at the kth step are obtained as

$$\hat{\vartheta}_{ik} = \frac{\int_\Theta \vartheta_i \, p(\boldsymbol{\vartheta}) \, L_k\big(\boldsymbol{\vartheta}|\boldsymbol{x}, \boldsymbol{r}\big) \, d\boldsymbol{\vartheta}}{\int_\Theta \, p(\boldsymbol{\vartheta}) \, L_k\big(\boldsymbol{\vartheta}|\boldsymbol{x}, \boldsymbol{r}\big) \, d\boldsymbol{\vartheta}}, \;\; i = 1, 2,$$

where Θ is the parameter space and $p(\boldsymbol{\vartheta})$ is the prior distribution of the parameters. A bivariate uniform density can be assumed for the parameters. A choice of $u_1 < \vartheta_1 < u_2$ and $u_3 < \vartheta_2 < u_4$ gives a restricted parameter space as $\tilde{\Theta} = \big\{\boldsymbol{\vartheta} : u_1 < \vartheta_1 < u_2, \;\; u_3 < \vartheta_2 < u_4\big\}$ so that

$$p(\boldsymbol{\vartheta}) = \frac{1}{(u_2 - u_1)(u_4 - u_3)}, \;\; \boldsymbol{\vartheta} \in \tilde{\Theta}.$$

The doses received by the k successive patients from $\mathcal{X}$ are represented as $\boldsymbol{\xi}_k = \{x_1, x_2, \ldots, x_k\}$. Now let us define

$$\boldsymbol{M}\big(x|\boldsymbol{\xi}_k, \hat{\boldsymbol{\vartheta}}_k\big) = \frac{k}{k+1} \boldsymbol{M}\big(\boldsymbol{\xi}_k, \hat{\boldsymbol{\vartheta}}_k\big) + \frac{1}{k+1} \boldsymbol{I}\big(x, \hat{\boldsymbol{\vartheta}}_k\big),$$

where $\boldsymbol{M}(\boldsymbol{\xi}_k, \hat{\boldsymbol{\vartheta}}_k) = \sum_{i=1}^{k} \boldsymbol{I}(x_l, \hat{\boldsymbol{\vartheta}}_k)$ and $\boldsymbol{I}(x_l, \hat{\boldsymbol{\vartheta}}_k)$ is the FIM for a patient who received the dose x_l, as shown in (2.2).

According to the construction of optimal experimental designs by Atkinson et al. (2007), we can select the dose x_{k+1} for the next patient such that

$$x_{k+1} = \arg\max_{x \in \mathcal{X}} \, \phi_D\Big\{\boldsymbol{M}\big(x|\boldsymbol{\xi}_k, \hat{\boldsymbol{\vartheta}}_k\big)\Big\}, \tag{2.3}$$

where $\phi_D\{\boldsymbol{M}\} = |\boldsymbol{M}|$. A trial is continued until a fixed sample size n is achieved. We represent this design by D_1.

2.2.2 MLE-based D-optimal design

This design is similar to the posterior-based design apart from the fact that the dose-response parameters $\boldsymbol{\vartheta}$ are estimated through the maximum likelihood estimation procedure. It is well known that the maximum-likelihood estimates can be obtained by solving the iterative equation

$$\boldsymbol{I}^{(i-1)} \hat{\boldsymbol{\vartheta}}_k^{(i)} = \boldsymbol{I}^{(i-1)} \hat{\boldsymbol{\vartheta}}_k^{(i-1)} + \boldsymbol{U}^{(i-1)},$$

where $\boldsymbol{I}$ refers the FIM in (2.2), $\boldsymbol{U}$ is the score vectors, and the superscript (i) indicates the ith approximation. The obtained estimates $\hat{\boldsymbol{\vartheta}}_k$ are then used in the dose optimization criterion in (2.3) to find the appropriate dose for the next patient.

Getting nontoxic responses only means there is no heterogeneity in the information. This lack of sufficient information makes MLE unattainable (Silvapulle, 1981). Hence, Storer's up-and-down design can be used until the first DLT occurs. Storer (1989) recommended three designs to follow in phase I clinical trials. Among the recommended designs, we have followed the first design, where cohort size is kept one at each dose, and escalation is done to the next higher dose if a nontoxic outcome occurs. After the occurrence of a toxic outcome, the MLE-based design starts. As like the previous design, a trial is continued until a fixed number of patients n is achieved, and the design is referred to as D_2.

2.2.3 Sequential Bayesian D-optimal design

The sequential Bayesian D-optimal design at any step k, for any given dose x, can be constructed so that it maximizes the expectation

$$\begin{aligned}\phi_{D_{\text{seq}}}(\boldsymbol{\xi}_k) &= \mathrm{E}_{\boldsymbol{\vartheta}} \ln \left|\boldsymbol{M}(x|\boldsymbol{\xi}_k, \boldsymbol{\vartheta})\right| \\ &= \int_{\Theta} \ln \left|\boldsymbol{M}(x|\boldsymbol{\xi}_k, \boldsymbol{\vartheta})\right| p_k\left(\boldsymbol{\vartheta}|\boldsymbol{x}, \boldsymbol{r}\right) d\boldsymbol{\vartheta},\end{aligned} \tag{2.4}$$

where $\boldsymbol{\xi}_k$ has been defined earlier, $p_k\left(\boldsymbol{\vartheta}|\boldsymbol{x}, \boldsymbol{r}\right)$ is the posterior density for any step k, obtained by updating the prior $p(\boldsymbol{\vartheta})$, as

$$p_k\left(\boldsymbol{\vartheta}|\boldsymbol{x}, \boldsymbol{r}\right) = \frac{p(\boldsymbol{\vartheta})\, L_k\big(\boldsymbol{\vartheta}|\boldsymbol{x}, \boldsymbol{r}\big)\, d\boldsymbol{\vartheta}}{\int_{\Theta} p(\boldsymbol{\vartheta})\, L_k\big(\boldsymbol{\vartheta}|\boldsymbol{x}, \boldsymbol{r}\big)\, d\boldsymbol{\vartheta}}, \tag{2.5}$$

and for a specific dose x, $\boldsymbol{M}(x|\boldsymbol{\xi}_k, \boldsymbol{\vartheta})$ is defined as

$$\boldsymbol{M}\left(x|\boldsymbol{\xi}_k, \boldsymbol{\vartheta}\right) = \frac{k}{k+1}\boldsymbol{M}\left(\boldsymbol{\xi}_k, \boldsymbol{\vartheta}\right) + \frac{1}{k+1}\boldsymbol{I}(x, \boldsymbol{\vartheta}),$$

with $\boldsymbol{M}\left(\boldsymbol{\xi}_k, \boldsymbol{\vartheta}\right) = \sum_{i=1}^{k} \boldsymbol{I}\left(x_l, \boldsymbol{\vartheta}\right)$, and $\boldsymbol{I}\left(x, \boldsymbol{\vartheta}\right)$ is the FIM for a patient who received the dose x.

In this design, we can select the dose x_{k+1} for the next patient such that

$$x_{k+1} = \arg\max_{x \in \mathcal{X}} \phi_{D_{\text{seq}}}(\boldsymbol{\xi}_k).$$

As like the previous two designs, a trial is continued until a fixed number of patients n is achieved, and the design is presented as D_3.

2.2.4 Two-Stage Bayesian D-optimal design

This design is constructed based on the paper of Haines et al. (2003), which they called the *sequential design*. However, our design is slightly different from that of Haines et al. (2003). In both

stages of the *sequential design*, a constraint on the dose space using an upper value of the dose's distribution function is used. This constraint makes the computation of the design complex for any user. In contrast, we have used the dose-skipping constraint suggested by Goodman et al. (1995), as mentioned in Section 2.2. The stages of the design proceed as follows.

In the first stage, a Bayesian D-optimal design $\boldsymbol{\xi}_D^*$ is obtained by maximizing

$$\begin{aligned}\phi_{D_{\text{stage1}}}(\boldsymbol{\xi}) &= \mathrm{E}_{\boldsymbol{\vartheta}} \ln \left|\boldsymbol{M}(\boldsymbol{\xi}, \boldsymbol{\vartheta})\right| \\ &= \int_{\Theta} \ln \left|\boldsymbol{M}(\boldsymbol{\xi}, \boldsymbol{\vartheta})\right| \; p(\boldsymbol{\vartheta})\, d\boldsymbol{\vartheta},\end{aligned}$$

where

$$\boldsymbol{M}(\boldsymbol{\xi}, \boldsymbol{\vartheta}) = \sum_{i=1}^{d} w_i \boldsymbol{I}\big(x^{(i)}, \boldsymbol{\vartheta}\big), \tag{2.6}$$

is constructed over a finite set of d ordered doses $\mathcal{X} = \{x^{(1)}, x^{(2)}, \ldots, x^{(d)}\}$, and $p(\boldsymbol{\vartheta})$ is the prior distribution. Here, w_i's are the weights on the distinct points $x^{(i)}$ for $i = 1, 2, \ldots, d,\ \ w_i{>}0$, $\sum_{i=1}^{d} w_i = 1$ for a continuous design

$$\boldsymbol{\xi} = \begin{Bmatrix} x^{(1)} & x^{(2)} & \ldots & x^{(d)} \\ w_1 & w_2 & \ldots & w_d \end{Bmatrix},$$

and $\boldsymbol{I}(x^{(i)}, \boldsymbol{\vartheta})$ is the FIM defined in (2.2). After having an appropriate optimal design $\boldsymbol{\xi}_D^*$, initially, n_1 patients are allocated to the appropriate doses and their responses are recorded. The number of initial patients n_1, is kept as small as possible for ethical reasons, in particular, if the prior is not informative. However, at the same time, greater precision in starting the sequential procedure is achieved if n_1 is large. To achieve an integer allocation of the n_1 patients to the d doses, the algorithm in Pukelsheim (1993) has been used.

In the second stage of this procedure, patients are allocated to the appropriate optimal dose levels in a stepwise allocation. In this case, after having n_1 patients in the first stage, n_2 patients are considered in the second stage. Thus, in a total, $n_1 + n_2$ patients are considered in the whole design. Specifically, in the first step of the second stage, a single patient or a small cohort of patients is assigned to the dose x that maximizes

$$\phi_{D_{\text{stage2}}}(\boldsymbol{\xi}) = \int_{\Theta} \ln \left| n\boldsymbol{M}(\boldsymbol{\xi}_D^*, \boldsymbol{\vartheta}) + \boldsymbol{I}(x, \boldsymbol{\vartheta}) \right| \; p(\boldsymbol{\vartheta}|\boldsymbol{x}, \boldsymbol{r}) \; d\boldsymbol{\vartheta}, \tag{2.7}$$

where $n = n_1 + 1$, $\boldsymbol{M}(\boldsymbol{\xi}_D^*, \boldsymbol{\vartheta})$, as shown in (2.6), is the information matrix at the updated optimal design $\boldsymbol{\xi}_D^*$, $\boldsymbol{I}(x, \boldsymbol{\vartheta})$ is the FIM for a single dose x, and $p(\boldsymbol{\vartheta}|\boldsymbol{x}, \boldsymbol{r})$ is the posterior density defined in (2.5).

In subsequent steps and as the data accrue, the dose allocated at each step is chosen to maximize the above criterion, but with the number of the patients already in the study n, the optimal design $\boldsymbol{\xi}_D^*$, and the posterior density $p(\boldsymbol{\vartheta}|\boldsymbol{x}, \boldsymbol{r})$, appropriately updated. We can select the dose x_{k+1} for the next patient such that

$$x_{k+1} = \arg\max_{x \in \mathcal{X}} \phi_{D_{\text{stage2}}}(\boldsymbol{\xi}_k),$$

where $\phi_{D_{\text{stage2}}}(\boldsymbol{\xi}_k)$ is defined according to (2.7) for any step k. As like the other designs, a trial is continued until a fixed number of patients, the sum of n_1 and n_2, is achieved. This design is presented as D_4.

2.2.5 MLE-based A-optimal design

This design is similar to the MLE-based D-optimal design with the exception that A-optimality criterion selects the dose x_{k+1} as follows:

$$x_{k+1} = \arg\max_{x \in \mathcal{X}} \phi_A(\boldsymbol{\xi}_k),$$

where $\phi_A(\boldsymbol{\xi}_k) = \text{tr}\big\{\boldsymbol{M^{-1}}(\boldsymbol{\xi_k}, \hat{\boldsymbol{\vartheta}}_{\boldsymbol{k}})\big\}$ reserving all notations defined earlier (Atkinson et al., 2007). The design is referred as A in the paper.

3 Setup to Simulations

A simulation study is conducted to investigate the operating characteristics of the designs. Six plausible dose-response scenarios are considered, as shown in Figure 1. Each scenario has the set of six available doses as $\mathcal{X} = \{1, 3, 5, 7, 9, 11\}$. We use equally spaced doses, as it is a common practice in dose-finding studies. Moreover, we choose six doses since a typical dose-finding study uses 5-6 doses to determine the MTD. The scenarios differ only in terms of the shape of the toxicity curve. The target toxicity level of γ is assumed to be 0.33. The first four scenarios have the true MTDs as doses 3, 5, 7, and 11, respectively, and they all are available in the dose vector $\mathcal{X}$. These are the doses at which the probabilities of DLT are less than or equal to the target toxicity level γ. More specifically, the probabilities of DLT at these doses are 0.32, 0.32, 0.33, and 0.32. In the last two scenarios, the dose, at which the probability of DLT close to the target rate, is in the middle of two available doses from $\mathcal{X}$. The true MTD in Scenario 5 is 6, which lies between the available doses 5 and 7 with probabilities of DLT as 0.24 and 0.43, respectively. Since dose 6 is not available in $\mathcal{X}$, dose 5 is assumed to be the true MTD. In Scenario 6, we see that the true MTD is 10, which is not available in $\mathcal{X}$; it lies between the available doses 9 and 11 with probabilities of DLT as 0.28 and 0.39, respectively. Hence, dose 9 is assumed to be the true MTD. These two scenarios are considered, as in real trials, we may start with the set of doses, where none of them has the probability of DLT exactly or approximately equal to the target toxicity level.

For the Bayesian designs, a bivariate uniform distribution is assumed for the dose-response parameters $\boldsymbol{\vartheta}$. A single parameter space $\tilde{\Theta} = \big\{\boldsymbol{\vartheta} : -4.3 < \vartheta_1 < -2.3,\ 0 < \vartheta_2 < 1\big\}$ is considered for all the six scenarios. The parameter space has been found to allow a wide range of dose-response scenarios including the assumed ones for simulation study. The Bernoulli distribution is used to generate the response following the assignment of a dose to a patient.

Each trial starts with the lowest dose of 1 mg/kg-body weight applied to a patient for the first two designs. As described in Section 2, the dose-response parameters are estimated after receiving the outcomes. Then a dose is selected following the dose-optimization criterion associated with the particular design. The selected dose is applied to a new patient. After receiving the second patient's

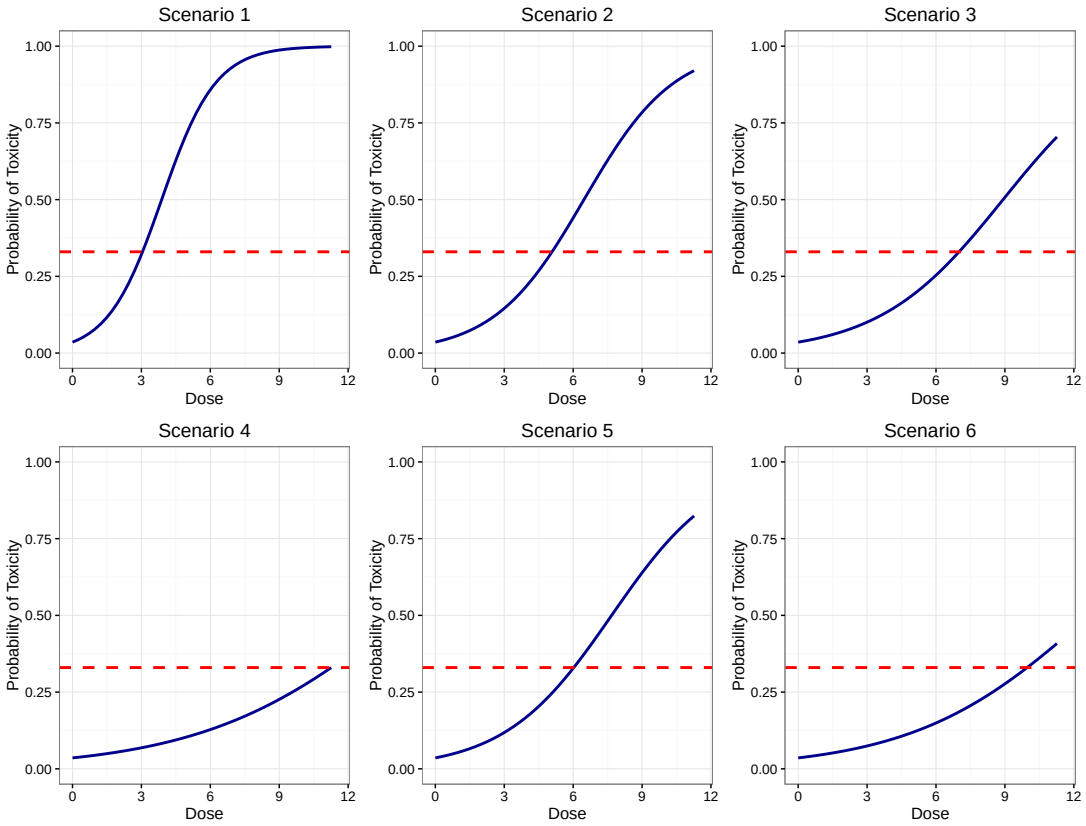

Figure 1: Dose-response scenarios for the simulation study with respective parameter values: Scenario 1, $\boldsymbol{\vartheta} = (-3.3, 0.85)$; Scenario 2, $\boldsymbol{\vartheta} = (-3.3, 0.51)$; Scenario 3, $\boldsymbol{\vartheta} = (-3.3, 0.37)$; Scenario 4, $\boldsymbol{\vartheta} = (-3.3, 0.23)$; Scenario 5, $\boldsymbol{\vartheta} = (-3.3, 0.43)$; Scenario 6, $\boldsymbol{\vartheta} = (-3.3, 0.26)$. The dashed horizontal line indicates the target toxicity level.

outcomes, the dose-response parameters are re-estimated, and a dose is chosen for the third patient. The process continues until the trial reaches the maximum number of patients n, after which the MTD is determined. The estimated dose-toxicity curve based on the parameter estimates at the last stage is used to find the MTD. The dose for which the absolute difference between the estimated probability of DLT and the target toxicity level is minimum is taken as the MTD. That is, we find the dose as

$$x_{n+1} = \arg\min_{x \in \mathcal{X}} \left| \psi(x, \hat{\boldsymbol{\vartheta}}_n) - \gamma \right|.$$

For the third design, the process is similar to the previous two designs. The difference is that dose-response parameter estimates have no use in dose selection for the successive patients. Instead, at each step we maximize the objective function in (2.4), which utilizes the prior distribution of the parameters $\boldsymbol{\vartheta}$. The prior is updated after receiving outcomes at each step, accordingly (2.5). Once

the trial reaches n, the dose-response parameters are estimated to obtain the MTD, as in the previous two designs.

As described in Section 2.2.4, in its first stage of the fourth design, we find an appropriate optimal design $\boldsymbol{\xi}_D^*$. The `R` package `Rsolnp` is used to obtain the weights w_i of that optimal design. Initially, $n_1 = 5$ patients are assigned to the corresponding doses in the optimal design, and then the responses are simulated. For an efficient rounding of the integer allocation of the 5 patients, the `R` function `efficient.rounding()` is utilized under the package `AlgDesign`. The function implements an efficient rounding procedure developed by Pukelsheim and Rieder (1992), to round approximate theory designs into replicated integer approximations. We obtain the response vector $\boldsymbol{r}$ for the first n_1 patients. Before starting the second stage, we update the prior $p(\boldsymbol{\vartheta})$ using the response vector $\boldsymbol{r}$ and the dose vector $\boldsymbol{x}$, accordingly (2.5), which is a function of the unknown parameters $\boldsymbol{\vartheta}$. In its first step of the second stage, the design searches for the best dose according to the design criterion in (2.7). Using the information obtained in that step, we update n, $\boldsymbol{\xi}_D^*$, and $p(\boldsymbol{\vartheta}|\boldsymbol{x}, \boldsymbol{r})$. The subsequent steps continue in a similar way until a pre-specified n is achieved. It is worth mentioning that weights w_i of the optimal design $\boldsymbol{\xi}_D^*$ are the percentages of dose allocation calculated in each step. Also, the determination of the MTD and the stopping rule are the same for all the four designs presented.

At the end of each simulated trial, we record the MTD, allocated doses to the patients, and estimate the probability toxicity at the MTD. The recorded estimates of toxicity over the simulations are averaged to find the bias in estimating the probability of toxicity at the recommended MTD. Varying n is considered for the scenarios, and it includes the values 15, 20, 25, and 30. Each case is investigated through 1000 simulations using a self-developed code in `R`. In a computer with a Core i7 processor and 8 GB RAM, the average processing times of 1000 simulations for the successive designs are 20 minutes, 10 minutes, 18 hours, and 25 hours, respectively.

4 Numerical Results

Scenario 1 has dose 3 as the true MTD. If $n = 15$, it is selected as the MTD in 95.8%, 76.5%, 96.5%, and 96.1% of the trials by the designs D_1, D_2, D_3, and D_4, respectively: see Table 1. As n increases, the identification of the MTD improves for all the designs. The percentage of patients that receive the true MTD during the trials is highest in D_2, followed by D_1, D_4 and D_3, respectively. For instance, when $n = 20$, 32.9%, 23.9%, 22.4%, and 18.5% of the patients are treated at the true MTD by the designs D_2, D_1, D_4 and D_3, respectively. The percentage of trials recommending a highly toxic dose as the MTD is highest for D_2. Except D_2, the other three designs perform almost in a similar way.

The true MTD in Scenario 2 is dose 5. If we engage 15 patients, as shown in Table 2, dose 5 is selected as the MTD in 75.4%, 52.2%, 66.7%, and 67.7% of the trials by D_1, D_2, D_3, and D_4, respectively. The designs become more able to identify the true MTD as n increases. The percentage of patients that receive the true MTD during the trials is highest at D_2, followed by D_4, D_3 and D_1, respectively. If $n = 15$, the percentage of patients treated at the true MTD by the successive designs are 23.5%, 19.1%, 13.7%, and 8.2%, respectively. The percentage of trials recommending a highly

toxic dose as the MTD is highest for D_2. Other than D_2, the designs perform very similarly.

Dose 7 is the right dose to be identified in Scenario 3. If $n = 15$, the successive designs D_1 to D_4 recommend this dose as the MTD in 46.6%, 42.4%, 46.7%, and 45.2% of the trials, respectively: see Table 3. This identification of the MTD improves for all the designs as n increases. The allocation of the true MTD during the trials is almost the same in all the designs. The percentage of trials recommending highly toxic doses and subtherapeutic doses as the MTD is highest for D_2. To conclude, except D_2, the other designs perform very similarly. Note that D_2 is the design obtained following the traditional approach to D-optimality. The maximum likelihood estimates of parameters are plugged in to obtain the next dose at each stage of a trial. Therefore, the behavior that D-optimum design can allocate doses from the extreme of the design region is quite evident in these scenarios.

Scenario 4 has dose 11 as the true MTD, and it is the highest available dose in $\mathcal{X}$. As indicated in Table 4, if n is 15, the true MTD is selected in 23.6%, 43.2%, 44.2%, and 24.8% of the trials by the successive designs. Except for D_4, the MTD identification does not always increase with the increase in n. The percentage of patients that receive the true MTD during the trials are very similar across the designs. Since dose 11 is the highest available dose and the probability of DLT at that dose is within the acceptable level, we do not have any trial with toxic dose as the MTD. In recommending subtherapeutic doses like 1 and 3, D_2 shows unsatisfactory performance, as many trials recommend these as the MTD. As a whole, D_2 performs better than the other designs.

Scenario 5 has the true MTD between doses 5 and 7. Since the true MTD is not available in $\mathcal{X}$ for this scenario, a dose-finding algorithm will tend to select these doses as the MTD. Since the DLT at dose 5 is closer to the target, it is expected to be selected as the MTD more often than the other dose 7. Table 5 reflects this expectation for all the designs except D_2. When $n = 15$, these two dose levels are selected as the MTD in 90.5%, 77.1%, 90.2%, and 86.8% of the trials by the successive designs. This performance improves for all the designs with an increase of n. The number of trials recommending a highly toxic dose as the MTD is highest for D_2. The percentage of patients that receive doses 5 and 7 during the trials is highest at D_2. Here again, except D_2, the other designs perform very similarly.

Like the previous scenario, the true MTD for Scenario 6 is not available in $\mathcal{X}$. Instead, it lies between the available doses 9 and 11. The target DLT is closer to the DLT at dose 9. As a result, all the designs select dose 9 as the MTD more often, as shown in Table 6. When n is 15, doses 9 and 11 are selected as the MTD in 57.8%, 68.3%, 62.9%, and 51.4% of the trials by the successive designs. D_2 allocates more patients to doses 9 and 11 during the trials than in other designs. To conclude, D_2 performs better than the other designs.

We observe a downward bias in estimating the probability of DLT for the scenarios in Table 7. These indicate that along with the correct identification, the designs tend to choose lower doses than the unacceptably toxic doses more often as the MTD. Generally, bias decreases for the designs as n increases. Other than D_2, all the designs experience relatively small bias in Scenarios 1, 2, 3, and 5. However, for Scenarios 4 and 6, the bias in D_2 is less compared to that of the other designs. The presented biases are in line with the selection of the correct MTD in the earlier tables.

The percentages of MTD selection and dose allocation for MLE-based A-optimal design are pre-

Table 1: The percentage of times the doses are identified as the MTD and the percentage of patients treated at these doses during the trials (in parenthesis) for Scenario 1. The dose level in bold is the MTD for this scenario.

n	Design	Dose (Probability of DLT)					
		1 (0.08)	**3 (0.32)**	5 (0.72)	7 (0.93)	9 (0.99)	11 (1.00)
15	D_1	0.1 (24.2)	95.8 (24.1)	4.1 (21.8)	0.0 (25.3)	0.0 (4.2)	0.0 (0.5)
	D_2	13.5 (19.3)	76.5 (33.9)	10.0 (31.4)	0.0 (11.9)	0.0 (2.5)	0.0 (0.9)
	D_3	0.0 (27.4)	96.5 (18.4)	3.5 (26.6)	0.0 (16.6)	0.0 (8.7)	0.0 (2.4)
	D_4	0.0 (24.3)	96.1 (23.6)	3.9 (24.8)	0.0 (14.8)	0.0 (6.7)	0.0 (5.7)
	A	31.8 (6.7)	44.6 (6.2)	18.0 (11.6)	5.3 (5.1)	0.3 (4.1)	0.0 (66.4)
20	D_1	0.4 (24.5)	97.5 (23.9)	2.1 (22.8)	0.0 (24.9)	0.0 (3.6)	0.0 (0.4)
	D_2	9.5 (18.1)	82.6 (32.9)	7.8 (33.5)	0.1 (12.2)	0.0 (2.6)	0.0 (0.8)
	D_3	0.4 (28.6)	97.3 (18.5)	2.3 (27.1)	0.0 (17.5)	0.0 (6.6)	0.0 (1.7)
	D_4	0.1 (26.0)	97.6 (22.4)	2.3 (26.3)	0.0 (16.0)	0.0 (5.0)	3.0 (4.3)
	A	28.9 (5.0)	46.7 (4.6)	19.2 (10.0)	5.2 (4.4)	0.0 (3.4)	0.0 (72.5)
25	D_1	0.4 (25.3)	98.9 (23.4)	0.7 (23.3)	0.0 (24.4)	0.0 (3.1)	0.0 (0.4)
	D_2	7.7 (19.3)	88.2 (32.5)	4.0 (33.7)	0.0 (11.7)	0.1 (2.3)	0.0 (0.6)
	D_3	0.0 (28.3)	98.7 (19.2)	1.3 (27.2)	0.0 (18.4)	0.0 (5.3)	0.0 (1.6)
	D_4	0.4 (27.1)	98.4 (21.9)	1.2 (26.7)	0.0 (16.8)	0.0 (4.0)	0.0 (3.5)
	A	31.1 (4.0)	46.5 (3.7)	17.1 (9.1)	5.3 (4.3)	0.0 (4.1)	0.0 (74.1)
30	D_1	0.5 (21.0)	99.3 (21.8)	0.2 (25.5)	0.0 (23.1)	0.0 (2.3)	0.0 (0.3)
	D_2	5.9 (19.1)	90.3 (31.9)	3.7 (34.3)	0.0 (12.2)	0.1 (2.1)	0.0 (0.5)
	D_3	0.5 (27.8)	98.9 (20.1)	0.6 (26.7)	0.0 (19.5)	0.0 (4.7)	0.0 (1.2)
	D_4	0.4 (27.4)	99.1 (21.5)	0.5 (27.1)	0.0 (17.5)	0.0 (3.6)	0.0 (3.0)
	A	32.8 (3.3)	44.8 (3.1)	17.1 (9.5)	5.3 (3.8)	0.0 (4.1)	0.0 (76.2)

sented in Tables 1-6. The percentage of correct MTD selection in Scenario 1 is the lowest compared to the D-optimal designs, and it does not improve with the increase of sample size remarkably. The tendency of dose allocation is completely different from those for the other designs. The A-optimal design generally allocates the highest toxic doses often and the true MTD rarely. This nature of MTD selection and dose allocation prevails in other scenarios as well. As seen in Table 7, the bias remains high for A-optimal design in most cases.

Table 2: The percentage of times the doses are identified as the MTD and the percentage of patients treated at these doses during the trials (in parenthesis) for Scenario 2. The dose level in bold is the MTD for this scenario.

n	Design	Dose (Probability of DLT)					
		1 (0.06)	3 (0.15)	**5 (0.32)**	7 (0.57)	9 (0.78)	11 (0.91)
15	D_1	0.0 (8.3)	16.6 (29.8)	75.4 (8.2)	7.8 (26.7)	0.2 (13.0)	0.0 (14.0)
	D_2	3.9 (11.1)	18.4 (17.1)	52.2 (23.5)	23.6 (23.8)	1.9 (14.9)	0.0 (9.5)
	D_3	0.0 (14.8)	20.6 (22.3)	66.7 (13.7)	12.3 (19.4)	0.4 (12.5)	0.0 (17.3)
	D_4	0.0 (8.7)	19.2 (24.5)	67.7 (19.1)	12.8 (19.9)	0.3 (11.5)	0.0 (16.2)
	A	10.3 (9.1)	24.0 (6.3)	42.2 (11.8)	19.1 (11.0)	1.9 (22.4)	2.6 (39.5)
20	D_1	0.0 (6.1)	14.4 (32.0)	76.3 (6.0)	9.1 (28.7)	0.2 (11.7)	0.0 (15.5)
	D_2	2.6 (9.8)	12.9 (15.6)	58.5 (22.9)	25.6 (24.0)	0.4 (17.4)	0.0 (10.3)
	D_3	0.0 (11.1)	13.0 (25.8)	77.4 (10.6)	9.4 (22.2)	0.2 (11.7)	0.0 (18.6)
	D_4	0.0 (7.3)	15.8 (27.0)	74.8 (15.2)	9.2 (22.8)	0.2 (10.6)	0.0 (17.1)
	A	9.4 (8.4)	26.6 (4.7)	44.4 (10.5)	16.1 (10.4)	1.1 (29.8)	2.4 (36.3)
25	D_1	0.0 (5.3)	10.9 (32.9)	81.5 (5.2)	7.6 (29.3)	0.0 (11.3)	0.0 (15.9)
	D_2	0.9 (9.3)	14.6 (15.8)	63.9 (23.4)	20.4 (23.2)	0.2 (17.8)	0.0 (10.6)
	D_3	0.0 (9.3)	11.0 (28.2)	82.7 (8.8)	6.3 (24.8)	0.0 (10.9)	0.0 (18.1)
	D_4	0.0 (6.0)	9.6 (28.5)	84.0 (12.7)	6.4 (24.2)	0.0 (10.2)	0.0 (18.5)
	A	8.4 (8.2)	22.4 (3.8)	47.6 (10.8)	17.4 (11.5)	1.2 (33.7)	3.0 (31.9)
30	D_1	0.0 (4.5)	7.4 (33.6)	86.4 (4.4)	6.2 (29.7)	0.0 (11.6)	0.0 (16.2)
	D_2	1.2 (9.4)	12.8 (16.8)	69.1 (22.2)	16.9 (22.4)	0.0 (18.2)	0.0 (10.9)
	D_3	0.0 (7.8)	9.3 (29.8)	85.4 (7.4)	5.3 (26.1)	0.0 (10.8)	0.0 (18.1)
	D_4	0.0 (5.6)	10.0 (30.0)	83.2 (11.1)	6.8 (25.5)	0.0 (10.2)	0.0 (18.0)
	A	9.7 (9.5)	24.1 (3.1)	48.6 (10.6)	13.6 (10.4)	1.5 (35.1)	3.0 (31.3)

Table 3: The percentage of times the doses are identified as the MTD and the percentage of patients treated at these doses during the trials (in parenthesis) for Scenario 3. The dose level in bold is the MTD for this scenario.

n	Design	Dose (Probability of DLT)					
		1 (0.05)	3 (0.10)	5 (0.19)	**7 (0.33)**	9 (0.51)	11 (0.68)
15	D_1	0.0 (10.6)	1.5 (19.9)	34.4 (10.0)	46.6 (17.5)	15.4 (12.7)	2.1 (29.3)
	D_2	4.7 (11.4)	3.5 (12.3)	22.5 (16.0)	42.4 (19.4)	24.1 (19.5)	3.3 (21.3)
	D_3	0.0 (16.2)	1.6 (15.2)	32.3 (13.0)	46.7 (12.9)	14.8 (12.4)	4.6 (30.3)
	D_4	0.0 (9.0)	0.8 (16.5)	38.4 (17.3)	45.2 (17.0)	12.8 (14.0)	2.8 (26.0)
	A	7.0 (11.6)	7.4 (6.3)	27.7 (12.8)	31.6 (17.3)	18.1 (36.3)	8.2 (15.8)
20	D_1	0.0 (8.6)	0.2 (21.8)	25.4 (8.3)	57.8 (18.6)	14.5 (10.8)	2.1 (32.1)
	D_2	2.3 (10.3)	2.2 (10.3)	19.9 (15.1)	48.2 (18.4)	25.7 (20.7)	1.7 (25.3)
	D_3	0.0 (12.8)	0.5 (18.5)	30.8 (10.5)	53.9 (15.5)	12.2 (10.4)	2.6 (32.3)
	D_4	0.0 (7.7)	0.6 (18.4)	27.6 (13.6)	56.1 (18.4)	13.3 (12.4)	2.4 (29.4)
	A	6.4 (11.3)	7.0 (4.8)	25.9 (11.8)	38.1 (16.1)	16.6 (41.9)	6.6 (14.2)
25	D_1	0.0 (7.4)	0.3 (22.6)	21.6 (7.0)	63.3 (20.2)	13.2 (9.2)	1.6 (33.7)
	D_2	2.3 (10.3)	1.9 (10.3)	21.7 (15.6)	47.5 (17.6)	25.6 (19.5)	1.0 (26.7)
	D_3	0.0 (10.9)	0.3 (19.9)	25.1 (9.0)	60.8 (17.2)	12.8 (9.1)	1.0 (33.9)
	D_4	0.0 (6.9)	0.0 (19.6)	23.2 (11.7)	63.6 (19.4)	11.6 (10.7)	1.6 (31.7)
	A	4.2 (11.7)	7.4 (3.8)	24.1 (10.6)	42.7 (16.5)	14.4 (45.3)	7.2 (12.1)
30	D_1	0.0 (6.4)	0.0 (23.6)	20.8 (6.3)	67.2 (20.9)	11.8 (8.2)	0.2 (34.6)
	D_2	1.3 (9.6)	1.6 (9.1)	21.6 (16.2)	53.1 (17.4)	21.7 (19.9)	0.7 (27.7)
	D_3	0.0 (9.5)	0.0 (21.0)	21.0 (8.0)	66.7 (18.1)	11.5 (8.0)	0.8 (35.4)
	D_4	0.0 (6.0)	0.0 (21.4)	23.2 (9.8)	65.0 (20.7)	10.8 (9.3)	1.0 (32.8)
	A	4.0 (11.8)	7.1 (3.1)	24.9 (10.3)	42.0 (14.6)	15.8 (49.2)	6.8 (10.9)

Table 4: The percentage of times the doses are identified as the MTD and the percentage of patients treated at these doses during the trials (in parenthesis) for Scenario 4. The dose level in bold is the MTD for this scenario.

n	Design	Dose (Probability of DLT)					
		1 (0.04)	3 (0.07)	5 (0.10)	7 (0.16)	9 (0.23)	**11 (0.32)**
15	D_1	0.0 (22.4)	0.1 (4.7)	22.6 (19.9)	10.1 (4.8)	43.6 (18.4)	23.6 (29.8)
	D_2	7.3 (12.0)	0.6 (10.2)	7.7 (13.0)	9.3 (11.2)	31.9 (22.3)	43.2 (31.4)
	D_3	0.0 (25.0)	0.0 (4.0)	25.1 (17.5)	9.4 (3.7)	65.6 (16.9)	44.2 (33.0)
	D_4	0.0 (15.2)	0.0 (7.5)	31.2 (21.6)	9.2 (7.7)	34.8 (20.3)	24.8 (27.7)
	A	12.1 (9.1)	2.7 (6.4)	11.1 (12.6)	5.8 (20.8)	27.0 (30.5)	41.8 (20.6)
20	D_1	0.0 (21.5)	0.0 (5.1)	16.9 (19.1)	10.4 (5.0)	33.3 (18.6)	39.4 (30.7)
	D_2	6.9 (12.5)	0.2 (8.9)	13.6 (12.9)	8.7 (11.3)	28.1 (21.7)	42.6 (32.8)
	D_3	0.0 (24.5)	0.0 (3.9)	36.8 (18.1)	8.5 (3.6)	27.1 (17.3)	27.6 (32.6)
	D_4	0.0 (17.0)	0.0 (6.3)	23.0 (21.3)	7.7 (6.5)	39.5 (20.7)	29.8 (28.3)
	A	9.8 (9.0)	1.8 (4.8)	11.8 (9.6)	4.8 (20.3)	20.6 (33.1)	51.3 (23.2)
25	D_1	0.0 (20.9)	0.0 (5.3)	16.6 (18.6)	7.7 (5.2)	32 (18.5)	43.7 (31.4)
	D_2	7.0 (13.6)	0.1 (9.0)	14.8 (12.7)	7.5 (11.2)	27 (20.3)	43.6 (33.2)
	D_3	0.0 (24.0)	0.0 (4.2)	28.2 (18.8)	6.3 (3.5)	36.8 (17.7)	28.7 (32.1)
	D_4	0.0 (17.6)	0.0 (5.6)	21.6 (20.9)	6.0 (5.7)	34.8 (20.1)	37.6 (29.9)
	A	9.4 (9.6)	1.8 (3.8)	10.2 (9.2)	5.3 (18.6)	21.9 (37.9)	52.0 (20.9)
30	D_1	0.0 (21.1)	0.0 (5.0)	18.0 (18.7)	5.5 (4.5)	33.9 (18.6)	42.6 (31.7)
	D_2	4.6 (13.6)	0.1 (7.6)	14.6 (12.7)	6.8 (11.7)	30.3 (20.3)	43.6 (34.0)
	D_3	0.0 (23.1)	0.0 (4.2)	23.7 (18.6)	5.5 (4.0)	37.7 (17.7)	33.1 (32.3)
	D_4	0.0 (18.4)	0.0 (5.4)	20.7 (20.5)	4.0 (5.2)	33.9 (20.5)	41.4 (30.1)
	A	10.1 (9.1)	2.8 (3.2)	9.8 (8.9)	5.5 (21.2)	21.5 (37.7)	51.9 (20.0)

Table 5: The percentage of times the doses are identified as the MTD and the percentage of patients treated at these doses during the trials (in parenthesis) for Scenario 5. The true MTD for this scenario lies between the dose levels in bold.

n	Design	Dose (Probability of DLT)					
		1 (0.05)	3 (0.12)	**5 (0.24)**	**7 (0.43)**	9 (0.64)	11 (0.81)
15	D_1	0.0 (8.4)	4.8 (25.3)	60.5 (8.2)	30.0 (22.5)	4.3 (12.3)	0.4 (23.3)
	D_2	4.8 (10.9)	7.8 (14.2)	36.4 (18.9)	40.7 (22.1)	10.2 (18.2)	0.6 (15.5)
	D_3	0.0 (14.4)	5.4 (19.6)	57.9 (12.6)	32.3 (16.4)	3.6 (12.2)	0.8 (24.9)
	D_4	0.0 (7.9)	6.6 (20.5)	53.4 (17.7)	33.4 (18.8)	6.0 (12.6)	0.6 (22.5)
	A	6.2 (11.0)	12.9 (6.3)	37.6 (12.6)	32.4 (15.4)	6.8 (31.6)	4.3 (23.0)
20	D_1	0.0 (6.4)	1.9 (27.4)	60.6 (6.3)	34.8 (23.5)	2.7 (10.8)	0.0 (25.8)
	D_2	2.8 (10.0)	4.6 (12.6)	39.3 (18.8)	45.7 (21.1)	7.4 (19.4)	0.2 (18.2)
	D_3	0.0 (11.1)	2.0 (22.5)	55.3 (9.7)	40.2 (18.7)	2.3 (10.37)	0.2 (27.7)
	D_4	0.0 (6.2)	3.3 (23.1)	54.0 (13.4)	39.8 (21.5)	2.7 (10.5)	0.2 (25.2)
	A	7.1 (11.3)	11.9 (4.7)	40.3 (11.3)	33.2 (13.6)	4.8 (39.0)	3.5 (20.1)
25	D_1	0.0 (5.3)	0.9 (27.9)	56.8 (5.2)	40.4 (24.4)	1.7 (9.5)	0.2 (27.6)
	D_2	2.0 (9.1)	3.5 (12.1)	40.9 (18.2)	49.0 (20.5)	4.6 (20.0)	0.0 (20.0)
	D_3	0.0 (8.9)	1.4 (24.4)	56.1 (8.0)	40.2 (20.8)	2.1 (8.8)	0.2 (29.2)
	D_4	0.0 (5.3)	1.6 (24.9)	58.4 (11.5)	38.8 (22.2)	1.2 (8.9)	0.0 (27.3)
	A	6.0 (10.7)	13.8 (3.8)	38.2 (11.3)	35.6 (14.9)	3.3 (40.7)	3.1 (18.6)
30	D_1	0.0 (4.5)	1.1 (28.7)	57.1 (4.5)	41.1 (24.8)	0.7 (8.6)	0.0 (28.9)
	D_2	0.9 (8.4)	3.3 (11.2)	42.2 (19.3)	49.5 (20.1)	4.0 (19.8)	0.1 (21.2)
	D_3	0.0 (7.6)	0.6 (25.7)	57.8 (6.9)	40.3 (21.8)	1.2 (8.1)	0.1 (29.9)
	D_4	0.0 (4.5)	0.9 (26.0)	58.2 (9.6)	39.7 (23.5)	1.1 (8.4)	0.0 (28.0)
	A	6.5 (10.6)	14.4 (3.2)	39.1 (11.0)	34.0 (12.3)	3.0 (47.7)	3.0 (15.3)

Table 6: The percentage of times the doses are identified as the MTD and the percentage of patients treated at these doses during the trials (in parenthesis) for Scenario 6. The true MTD for this scenario lies between the dose levels in bold.

n	Design	Dose (Probability of DLT)					
		1 (0.05)	3 (0.07)	5 (0.12)	7 (0.19)	**9 (0.28)**	**11 (0.39)**
15	D_1	0.0 (19.8)	0.0 (7.58)	23.1 (17.4)	19.1 (7.3)	40.0 (16.9)	17.8 (31.0)
	D_2	7.2 (11.8)	0.8 (10.5)	9.4 (13.2)	14.9 (13.5)	37.8 (21.4)	30.5 (29.6)
	D_3	0.0 (23.1)	0.0 (5.6)	18.4 (16.7)	18.7 (5.1)	24.3 (16.0)	38.6 (33.5)
	D_4	0.0 (14.3)	0.0 (8.7)	32.0 (21.0)	16.6 (9.1)	33.2 (19.2)	18.2 (27.7)
	A	11.1 (9.7)	2.4 (6.4)	12.9 (11.7)	11.1 (21.8)	30.7 (33.3)	32.6 (17.1)
20	D_1	0.0 (18.9)	0.0 (8.1)	14.4 (16.8)	20.7 (7.5)	36.7 (16.7)	28.2 (32.1)
	D_2	7.2 (12.9)	0.5 (9.9)	12.3 (12.8)	14.5 (12.2)	34.8 (20.2)	30.7 (32.0)
	D_3	0.0 (21.6)	0.0 (6.8)	25.1 (16.3)	19.7 (6.0)	29.2 (15.6)	26.0 (33.7)
	D_4	0.0 (15.0)	0.1 (8.4)	22.5 (19.6)	17.6 (8.6)	38.3 (19.1)	21.5 (29.4)
	A	8.0 (10.3)	2.6 (4.8)	13.7 (11.2)	10.1 (19.2)	26.0 (37.6)	40.3 (17.0)
25	D_1	0.0 (18.8)	0.0 (7.9)	14.9 (16.7)	14.8 (7.5)	39.4 (16.7)	30.9 (32.5)
	D_2	5.2 (13.1)	0.2 (8.5)	12.5 (12.9)	10.8 (12.5)	38.3 (19.9)	33 (32.9)
	D_3	0.0 (21.7)	0.0 (6.3)	26.2 (16.9)	15.3 (5.7)	37.6 (16.0)	20.9 (33.4)
	D_4	0.0 (15.7)	0.0 (8.0)	16.2 (19.0)	14.8 (7.8)	40.6 (18.6)	28.4 (30.8)
	A	6.9 (10.1)	3.0 (3.8)	12.5 (9.8)	8.5 (20.1)	31.6 (40.0)	37.5 (16.2)
30	D_1	0.0 (18.3)	0.0 (8.3)	17.4 (16.1)	15.9 (7.9)	35.7 (16.2)	31.0 (33.2)
	D_2	0.0 (13.1)	0.2 (8.3)	14.8 (13.4)	10.3 (12.6)	42.5 (19.2)	28.9 (33.5)
	D_3	0.0 (20.6)	0.0 (6.8)	22.8 (16.7)	12.6 (6.3)	41.5 (15.9)	23.1 (33.7)
	D_4	0.0 (16.3)	0.1 (7.6)	17.9 (18.6)	13.1 (7.5)	39.1 (18.6)	29.8 (31.5)
	A	8.3 (8.8)	2.4 (3.2)	9.1 (8.9)	5.2 (18.7)	36.0 (42.8)	39.2 (17.6)

Table 7: Bias in the estimation of probability of DLT at the recommended MTD.

Scenario	Design	n			
		15	20	25	30
1	D_1	-0.00833	-0.00715	-0.00722	-0.00622
	D_2	-0.05136	-0.04028	-0.03834	-0.03471
	D_3	-0.01115	-0.00548	-0.00055	-0.00343
	D_4	-0.00981	-0.00925	-0.00599	-0.00751
	A	-0.28452	-0.28721	-0.27908	-0.07964
2	D_1	-0.00894	-0.00668	-0.00515	-0.00450
	D_2	-0.04431	-0.02234	-0.02207	-0.01928
	D_3	-0.01346	-0.00434	-0.00596	-0.00101
	D_4	-0.00727	-0.00745	-0.00850	-0.00789
	A	-0.10877	-0.07403	-0.05342	-0.07964
3	D_1	-0.02341	-0.01762	-0.00988	-0.01190
	D_2	-0.05741	-0.03441	-0.02305	-0.02297
	D_3	-0.01702	-0.01240	-0.01312	-0.01766
	D_4	-0.02198	-0.01784	-0.01281	-0.01242
	A	-0.04934	-0.03567	-0.03558	-0.02655
4	D_1	-0.13927	-0.11797	-0.10883	-0.10519
	D_2	-0.06487	-0.07678	-0.08177	-0.07942
	D_3	-0.12773	-0.10255	-0.10536	-0.10739
	D_4	-0.13468	-0.13041	-0.11659	-0.11399
	A	-0.07604	-0.07085	-0.06957	-0.06650
5	D_1	-0.01412	-0.01345	-0.00674	-0.01145
	D_2	-0.04977	-0.02669	-0.01717	-0.01832
	D_3	-0.00827	-0.01136	-0.00527	-0.01276
	D_4	-0.01065	-0.00690	-0.00960	-0.00898
	A	-0.05672	-0.03932	-0.03145	-0.02277
6	D_1	-0.10929	-0.08851	-0.07779	-0.08082
	D_2	-0.06510	-0.06191	-0.05381	-0.06573
	D_3	-0.10094	-0.07102	-0.07514	-0.07514
	D_4	-0.12207	-0.09883	-0.08409	-0.07964
	A	-0.05709	-0.05372	-0.05371	-0.04400

5 An example

Now, dose-response data from Karp et al. (2001) is used to implement the designs considered in this paper. It was a phase I clinical trial conducted at the University of Maryland School of Medicine, Baltimore, involving 34 acute leukemia patients. A conventional up-and-down design with a chemotherapeutic agent was employed. The dose space selected was $\mathcal{X} = \{100, 300, 600, 900, 1200\}$, where the unit is in mg. The administration of these doses to a group of 34 patients resulted in the following toxicities: 0/6 at 100 mg, 0/5 at 300 mg, 3/8 at 600 mg, 6/11 at 900 mg and 3/4 at 1200 mg. We fit the logistic regression model of Section 2 to these data and obtain the MLE estimates of parameters as $\vartheta_1 = -3.80$ and $\vartheta_2 = 0.0045$. The fitted dose-toxicity curve is shown in Figure 2.

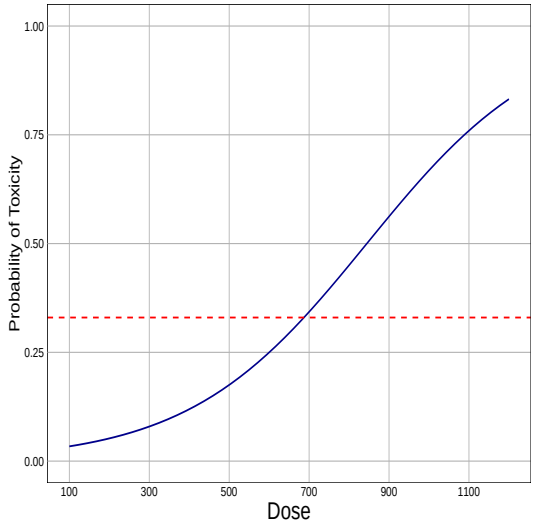

Figure 2: Dose-response scenario with parameter values $\boldsymbol{\vartheta} = (-3.80, 0.0045)$. The dashed horizontal line indicates the target toxicity level.

The target toxicity rate is $\gamma = 0.33$. The prior for Bayesian designs is assumed as $\tilde{\Theta} = \{\boldsymbol{\vartheta} : -4.3 < \vartheta_1 < -2.3,\ 0.0 < \vartheta_2 < 0.01\}$. The true MTD for this scenario is dose level 3, which is 600 mg. We consider $n = 15$ patients in each trial and 1000 simulated trials are explored for each design. Table 8 summarizes the performance of the D-optimal designs for this scenario. The selection of dose 600 as the MTD for the designs D_1, D_2, D_3, D_4 and A are 75.0%, 54.8%, 89.6%, 90.2% and 57.6%, respectively. The percentage of patients that receive the true MTD is highest in D_2, followed by A, D_4, D_1, and D_3. The design D_1 has the least bias, followed by D_4, D_3, A and D_2. Overall, D_4 is the best design in terms of the correct MTD identification, while D_2 and A are the two low-performing designs.

Table 8: The percentage of times the doses are identified as the MTD, the percentage of patients treated at these doses during the trials (in parenthesis) and the bias in the estimation of probability of DLT at the recommended MTD for the scenario in Figure 2.

Design	Dose (Probability of DLT)					Bias
	100 (0.03)	300 (0.08)	**600 (0.25)**	900 (0.56)	1200 (0.83)	
D_1	0.0 (7.8)	7.6 (30.04)	75.0 (10.3)	16.9 (29.0)	0.5 (22.9)	-0.00667
D_2	0.7 (11.2)	14.8 (15.3)	54.8 (26.1)	29.0 (27.9)	0.7 (19.5)	-0.07451
D_3	0.0 (4.1)	1.2 (34.9)	89.6 (5.1)	9.2 (30.3)	0.0 (25.7)	-0.01592
D_4	0.0 (3.8)	1.0 (32.5)	90.2 (11.1)	8.8 (28.8)	0.0 (23.8)	-0.01568
A	5.2 (7.5)	15.2 (6.4)	57.6 (15.3)	17.7 (32.4)	4.3 (38.3)	-0.06302

6 Discussion

This paper has investigated six dose-response scenarios to explore the performance of different D-optimal designs. Except for D_2, all the designs have shown competitive performance in Scenarios 1, 2, 3, and 5 in identifying the correct MTD. Notably, the D_1 is always entirely satisfactory in these scenarios, with the increase in n. The design D_2 performs well in Scenarios 4 and 6, where the MTD lies towards the upper end of the dose region. The MLE-based D-optimal design works well when the correct MTD is in the upper end. If we consider allocating patients to the correct MTD, D_2 has been found to allocate the highest percentage of patients to the correct MTD compared to the other designs in almost all the scenarios. Scenarios 1, 2, 3, and 5 have doses bearing the probability of DLT above the acceptable level. The tendency to select these toxic doses as the MTD is highest in D_2 compared to that of the other designs. The recommendation of a subtherapeutic dose as the MTD is also highest at D_2. As mentioned earlier, the MTD selection in various designs is well supported by the estimates of bias. Comparing the A-optimal design with the D-optimal designs shows that A performs poorly most of the time. The paper also includes an example extracted from a published study. The results there again assure the performance of the designs.

No design has been found to perform uniformly best in all situations. Although D_2 performs well in terms of dose allocation, except Scenarios 4 and 6, the dose recommendation is not attractive. Implementation of D_2 requires an up-and-down design to consider until the first toxicity occurs. The D_3 and D_4 require a complicated function to optimize. The D_1 is numerically much simpler than the D_3 and D_4. Considering all these, D_1 may be preferred over the other designs presented. The methodologies presented here are quite general and can be applied to other dose-response models in phase I trials. In general, an optimal design tries to identify the support points and its best distribution of the number of trials over those points. In this study, we fix the trial points and only consider the second part of the problem. We had to make this choice, as clinical trials are usually conducted at the discrete choice of pre-specified doses. However, we would like to mention this as a limitation of

the work.

Acknowledgements

The authors would like to thank the anonymous reviewers for their helpful suggestions to improve the paper. Also, the first author would like to acknowledge the Ministry of Science and Technology, Government of Bangladesh, for providing him the National Science and Technology fellowship during this work.

References

Alam, M. I. (2016), "A comparison between the continual reassessment method and D-optimum design for dose finding in phase I clinical trials," *Biometrical Letters*, 53, 69–82.

Atkinson, A. C., Donev, A. N., and Tobias, R. D. (2007), *Optimum Experimental Designs, with SAS*, Oxford University Press, Oxford.

Chaloner, K. and Verdinelli, I. (1995), "Bayesian experimental design: A review," *Statistical Science*, 10, 273–304.

Faries, D. (1994), "Practical modifications of the continual reassessment method for phase I cancer clinical trials," *Journal of Biopharmaceutical Statistics*, 4, 147–164.

Goodman, S. N., Zahurak, M. L., and Piantadosi, S. (1995), "Some practical improvements in the continual reassessment method for phase I studies," *Statistics in Medicine*, 14, 1149–1161.

Haines, L. M., Perevozskaya, I., and Rosenberger, W. F. (2003), "Bayesian optimal designs for phase I clinical trials," *Biometrics*, 59, 591–600.

Heyd, J. M. and Carlin, B. P. (1999), "Adaptive design improvements in the continual reassessment method for phase I studies," *Statistics in Medicine*, 18, 1307–1321.

Karp, J. E., Lancet, J. E., Kaufmann, S. H., End, D. W., Wright, J. J., Bol, K., Horak, I., Tidwell, M. L., Liesveld, J., Kottke, T. J., et al. (2001), "Clinical and biologic activity of the farnesyltransferase inhibitor R115777 in adults with refractory and relapsed acute leukemias: a phase 1 clinical-laboratory correlative trial," *Blood, The Journal of the American Society of Hematology*, 97, 3361–3369.

Korn, E. L., Midthune, D., Chen, T. T., Rubinstein, L. V., Christian, M. C., and Simon, R. M. (1994), "A comparison of two phase I trial designs," *Statistics in Medicine*, 13, 1799–1806.

Møller, S. (1995), "An extension of the continual reassessment methods using a preliminary up-and-down design in a dose finding study in cancer patients, in order to investigate a greater range of doses," *Statistics in Medicine*, 14, 911–922.

O'Quigley, J. and Chevret, S. (1991), "Methods for dose finding studies in cancer clinical trials: A review and results of a Monte Carlo study," *Statistics in Medicine*, 10, 1647–1664.

O'Quigley, J., Pepe, M., and Fisher, L. (1990), "Continual reassessment method: A practical design for phase I clinical trials in cancer," *Biometrics*, 46, 33–48.

Piantadosi, S., Fisher, J., and Grossman, S. (1998), "Practical implementation of a modified continual reassessment method for dose-finding trials," *Cancer Chemotherapy and Pharmacology*, 41, 429–436.

Pukelsheim, F. (1993), *Optimal Design of Experiments. 1993*, Wiley, New York.

Pukelsheim, F. and Rieder, S. (1992), "Efficient rounding of approximate designs," *Biometrika*, 79, 763–770.

Silvapulle, M. J. (1981), "On the existence of maximum likelihood estimators for the binomial response models," *Journal of the Royal Statistical Society. Series B (Methodological)*, 310–313.

Storer, B. E. (1989), "Design and analysis of phase I clinical trials," *Biometrics*, 45, 925–937.

White, L. V. (1973), "An extension of the general equivalence theorem to nonlinear models," *Biometrika*, 60, 345–348.

Received: October 25, 2020

Accepted: May 21, 2021