Phase I/II dose-finding design for molecularly targeted agent: Plateau determination using adaptive randomization

Marie-Karelle Riviere, Ying Yuan, Jacques-Henri Jourdan, Frédéric Dubois, Sarah Zohar

Phase I/II Dose-Finding Design for Molecularly Targeted Agent: Plateau Determination using Adaptive Randomization

Marie-Karelle Riviere[a,b], Ying Yuan[c], Jacques-Henri Jourdan[d], Frédéric Dubois[b] and Sarah Zohar[a]

[a] *INSERM, U1138, Equipe 22, Centre de Recherche des Cordeliers, Université Paris 5, Université Paris 6, Paris, France*

[b] *IRIS (Institut de Recherches Internationales Servier), Suresnes, France*

[c] *Department of Biostatistics, The University of Texas MD Anderson Cancer Center, Houston, TX 77030, U.S.A.*

[d] *INRIA Paris-Rocquencourt, France*

marie-karelle.riviere@inserm.fr

Abstract

Conventionally, phase I dose-finding trials aim to determine the maximum tolerated

dose (MTD) of a new drug under the assumption that both toxicity and efficacy monotonically increase with the dose. This paradigm, however, is not suitable for some molecularly targeted agents (MTAs), such as monoclonal antibodies, for which efficacy often increases initially with the dose and then plateaus. For MTAs, the goal is to find the optimal dose, defined as the lowest safe dose that achieves the highest efficacy. We develop a Bayesian phase I/II dose-finding design to find the optimal dose. We employ a logistic model with a plateau parameter to capture the increasing-then-plateau feature of the dose-efficacy relationship. We take the weighted likelihood approach to accommodate for the case where efficacy is possibly late-onset. Based on observed data, we continuously update the posterior estimates of toxicity and efficacy probabilities and adaptively assign patients to the optimal dose. The simulation studies show that the proposed design has good operating characteristics. This method is going to be applied in more than two phase I clinical trials as no other method is available for this specific setting. We also provide an R package *dfmta* that can be download from CRAN website.

Keywords: Dose-finding; Molecularly targeted agents; Oncology; Phase I; Phase II.

1 Introduction

Traditionally, phase I dose-finding trials aim to determine the maximum tolerated dose (MTD) of a new drug that will be further investigated for efficacy in phase II. This paradigm is built upon the assumption that both toxicity and efficacy monotonically

increase with the dose, which is typically true for conventional cytotoxic agents. Recently, molecularly targeted agents (MTAs) have emerged as a new therapeutic option in oncology that has changed the practice of cancer patient care. For many MTAs, e.g., monoclonal antibodies, the monotonicity assumption may be violated for efficacy although it typically holds for toxicity. For example, the FDA guidance points out "cancer vaccine trials have used the '3 + 3 design' and the results show that, except in very rare situations, an MTD for a cancer vaccine may not be identified. In these trials, the dose-toxicity curve may be so flat that the highest dose that can be administered is limited by manufacturing or anatomic issues rather than toxicity" [1]. As another example, the efficacy of PTK/ZK (an orally active inhibitor of vascular endothelial growth factor receptor tyrosine kinases) virtually does not change with the dose once it reaches the threshold (or plateau) of 1000mg, which is below the MTD [2, 3]. Further increasing the dose of PTK/ZK to the MTD does not improve its efficacy. As a result, traditional dose-finding methods for cytotoxic agents are not always suitable for MTAs because the MTAs do not necessarily need to be administered at their MTDs to achieve maximal efficacy. Depending on the structure of the molecule and its type, some MTAs can reach a plateau of efficacy. Therefore, for MTAs, we are interested in finding the biological optimal dose, which is defined as the lowest safe dose that achieves the highest efficacy, i.e., the dose corresponding to the plateau changing point in the dose-efficacy curve while satisfying certain toxicity requirements (see Figure 1).

A limited number of dose-finding methods have been proposed to handle the case

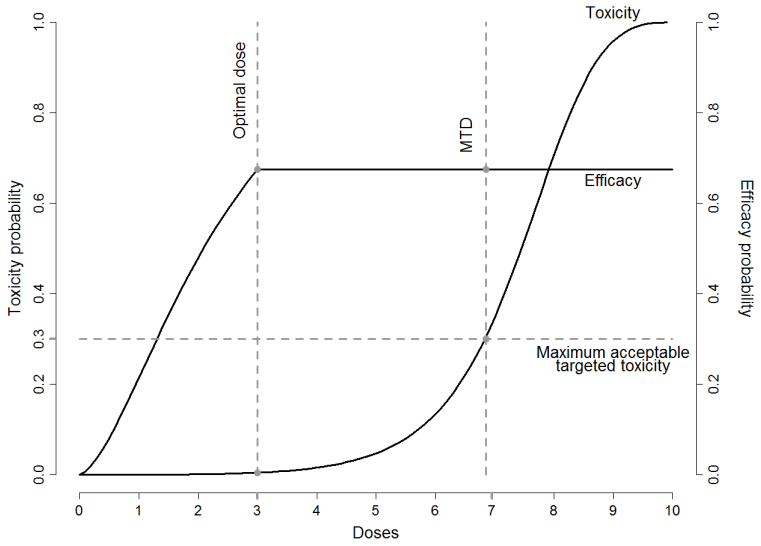

Figure 1: Illustration of dose-toxicity and dose-efficacy relationships for MTAs.

where the dose-response curve first increases with the dose and then reaches a plateau. Hunsberger et al [4] proposed a phase II dose-escalation design, assuming a linear regression model for the dose-response curve on the last three or four doses explored. If the estimated slope is null or negative, the trial is terminated and the dose with the highest response rate is selected as the optimal dose. Hirakawa [5] proposed another dose-finding design by jointly modeling a binary toxicity outcome with a continuous efficacy outcome, in which Mahalanobis distance was used to measure the desirability of the dose for dose assignment. Cai et al [6] proposed a Bayesian phase I/II to handle drug combination trials involving two MTAs when the dose-response curve plateaus at high doses. Ivanova and Xiao [7] developed a phase II design to determine the

minimum effective dose (MED), which does not necessarily correspond to the biological optimal dose. More recently, Zang, Lee and Yuan [8] proposed nonparametric and semiparametric approaches to find the biological optimal dose for MTAs with unimodal or plateaued dose-response curves.

This work is motivated by a phase I clinical trial initiated by the French Sarcoma Group, for which some authors served as statistical collaborators. The goal of the trial is to find the biological optimal dose of propranolol for treating patients with locally advanced or metastatic angiosarcoma. Stiles et al. [9] demonstrated that beta-adrenergic inhibition blocks cell proliferation and induces apoptosis in a dose dependent manner. Specifically, using in vivo tumor models, they demonstrated that propranolol, non-specific beta-adrenergic receptor inhibitor, shows remarkable efficacy in reducing the growth of angiosarcoma tumors. Physicians are expecting an efficacy plateau for this agent due to its action mechanism. Therefore, traditional MTD based dose finding methods cannot be used for this trial.

The aim of this paper was to propose a practical phase I/II dose-finding design for MTAs. We employ a logistic model with a plateau parameter to capture the increasing-then-plateau feature of the dose-efficacy relationship. We take the weighted likelihood approach to accommodate for the possibility that efficacy is late-onset in the sense that efficacy takes a relatively long time to be assessed compared to toxicity (with respect to the accrual rate), such that when the next new patient arrives, patients who have enrolled into the trial have not completed their efficacy assessment yet [10, 11, 12]. We

assume that toxicity monotonically increases with the dose and model it using a logistic model. Based on observed data, we continuously update the posterior estimates of toxicity and efficacy probabilities and adaptively assign patients to the optimal dose. We conduct extensive simulation to examine the operating characteristics of the proposed design.

This paper differs from our previous work on the combination of a cytotoxic agent with a MTA [13] in several important aspects. In our previous work, we employed the conventional greedy dose-assignment rule that always assigns new patients to the currently estimated optimal dose. Such greedy myopic algorithm works reasonably well when the number of investigational doses for the MTA is small. For example, in the combination trial of a cytotoxic agent with a MTA, in order to control the total number of investigational dose combinations within a practical range (e.g., < 12), the number of investigational dose levels for the MTA is often small, e.g., combining 3 doses of the MTA with a cytotoxic agent with 3 doses results in a total of 9 dose combinations. Therefore, the greedy dose-assignment rule often works well in that setting. However, we have observed that the percentage of correct recommendation of dose combinations at the end of the trial by our design were highly decreasing with the number of dose levels for the MTA. Even in the single agent dose-finding trial we focus here, when there are many (e.g., ≥ 6) doses for the MTA, the greedy dose-assignment rule becomes problematic and often causes the dose finding to get stuck in local maxima (more details are provided in Section 2.4). In this paper, we propose several robust, well

performing dose assignment rules to address this issue. The second difference between this paper and our previous work is that herein we propose a weighting based approach to acknowledge and allow the use of partial information available to make the decision of dose assignment for new patients. We also provide an R package *dfmta*, available from CRAN website, to implement the proposed design.

2 Methods

2.1 Toxicity model

Consider a trial involving K doses of an MTA. We model the toxicity probability of the k^{th} dose level, denoted as $\psi(k, \beta_0, \beta_1) = \psi_k$, using a logistic model as follows:

$$\text{logit}(\psi_k) = \beta_0 + \beta_1 u_k, \tag{1}$$

where β_0, β_1 are unknown parameters, and u_k is the "effective" dose associated with dose level k, which typically differs from the actual clinical dosage. In (1), we require $\beta_1 > 0$ so that toxicity monotonically increases with dose levels, as is often the case in practice. Let $\tilde{\psi}_k$ denote the prior guess of toxicity probability (i.e. working model, or skeleton) for dose level k, and $\tilde{\beta}_0$ and $\tilde{\beta}_1$ denote the prior estimates of β_0 and β_1 (for instance means of the prior distributions). The "effective" dose u_k is determined by

back-solving the dose-toxicity model, i.e.,

$$u_k = \left\{ \log\left(\frac{\tilde{\psi}_k}{1-\tilde{\psi}_k} \right) - \tilde{\beta}_0 \right\} / \tilde{\beta}_1.$$

The reason that we use the "effective" dose, rather than the actual clinical dosage, in the logistic regression is to regularize the estimate of ψ_k to a practically reasonable range (i.e., *a priori*, the estimate of ψ_k matches our prior guess $\tilde{\psi}_k$), thereby improving the stability of the trial design. This is important because the small sample size of phase I trials often results in very unreliable model fitting, especially at the beginning of the trial when there are only a few observations. Using the "effective" dose to fit regression models has been adopted previously in dose-finding studies [14, 15, 16].

Let y_i denote the binary toxicity outcome (1 = toxicity, and 0 = no toxicity) for patient i treated at the dose level $x_i \in \{1, \ldots, K\}$, where $i = 1, \ldots, N$. After the first I patients are enrolled into the trial, the likelihood of the toxicity data $\mathcal{D}_{\text{tox}} = \{(x_1, y_1), \ldots, (x_I, y_I)\}$ is:

$$L(\mathcal{D}_{\text{tox}}|\beta_0, \beta_1) = \prod_{i=1}^{I} \psi_{x_i}^{y_i} (1 - \psi_{x_i})^{1-y_i}.$$

Letting $f(\beta_0, \beta_1)$ denote the prior distribution of β_0 and β_1, the posterior is then given by:

$$f(\beta_0, \beta_1|\mathcal{D}_{\text{tox}}) \propto L(\mathcal{D}_{\text{tox}}|\beta_0, \beta_1) f(\beta_0, \beta_1) \tag{2}$$

We assume prior distributions are independent and take a vague normal prior $N(0, 100)$ for the intercept β_0, and following Chevret [15], we assign the slope β_1 an exponential distribution with a rate parameter of 1, i.e. $\beta_1 \sim Exp(1)$.

2.2 Efficacy model

An important feature of MTAs that distinguishes them from conventional cytotoxic agents is that the dose-efficacy curves of MTAs do not necessarily increase with the dose. For MTAs, efficacy is often expected to monotonically increase with the dose and then plateau after reaching the level of saturation. Let $\phi(k, \gamma_0, \gamma_1, \tau) = \phi_k$ denote the efficacy probability for dose level k. We model efficacy using a logistic model with plateau parameter τ, as follows:

$$\text{logit}(\phi_k) = \gamma_0 + \gamma_1(v_k \mathbb{1}(k < \tau) + v_\tau \mathbb{1}(k \geq \tau)), \tag{3}$$

where γ_0 and $\gamma_1 > 0$ are unknown parameters. The plateau parameter τ is an integer between 1 and K that indicates at which dose level the dose-efficacy curve reaches the plateau. When the dose level is lower than τ, efficacy monotonically increases with the dose, and when the dose level is equal to or higher than τ, efficacy plateaus with a constant dose effect $\gamma_1 v_\tau$. In model (3), the "effective" dose v_k is determined in a similar way as before. That is, we first elicit prior estimates of parameters $\tilde{\phi}_k$, $\tilde{\gamma_0}$, $\tilde{\gamma_1}$ and $\tilde{\tau}$ from physicians, and then obtain the value of v_k by back-solving the dose-efficacy

model as follows, $v_k = \left\{\log\left(\frac{\tilde{\phi}_k}{1-\tilde{\phi}_k}\right) - \tilde{\gamma}_0\right\}/\tilde{\gamma}_1$. In order to make v_k identifiable, we require $\tilde{\phi}_1 < \ldots < \tilde{\phi}_K$ (note that this does not mean that the true value of ϕ_k is monotonic). $\mathbb{1}(\mathcal{C})$ denote the indicator function, which takes a value of 1 if $\mathcal{C}$ is true and 0 if $\mathcal{C}$ is false.

Let z_i denote the binary efficacy outcome (1 = efficacy, and 0 = no efficacy) for patients i treated at the dose level $x_i \in \{1, \ldots, K\}$, where $i = 1, \ldots, N$. After the first I patients are enrolled into the trial, the likelihood of the efficacy data $\mathcal{D}_{\text{eff}} = \{(x_1, z_1), \ldots, (x_I, z_I)\}$ is:

$$L(\mathcal{D}_{\text{eff}}|\gamma_0, \gamma_1, \tau) = \prod_{i=1}^{I} \phi_{x_i}^{z_i}(1 - \phi_{x_i})^{1-z_i}.$$

Letting $f(\gamma_0, \gamma_1, \tau)$ denote the prior distribution of γ_0, γ_1, and τ, the posterior is then given by:

$$f(\gamma_0, \gamma_1, \tau|\mathcal{D}_{\text{eff}}) \propto L(\mathcal{D}_{\text{eff}}|\gamma_0, \gamma_1, \tau) f(\gamma_0, \gamma_1, \tau). \tag{4}$$

We assume prior distributions are independent and took a vague normal prior $N(0, 100)$ for the intercept γ_0 and an exponential distributions with a rate parameter of 1 for γ_1, i.e. $\gamma_1 \sim Exp(1)$. For the plateau parameter, τ, we assign a discrete prior distribution $\mathrm{P}(\tau = k) = p_k$, $k = 1, \ldots, K$, with $\sum_{k=1}^{K} p_k = 1$ and $\forall k, p_k \geq 0$. When there is prior information on the location of τ, e.g., we know the saturation dosage of the MTA from pharmacokinetic and pharmacodynamic studies, we can choose a set of p_k's to reflect the likelihood of each dose level being the plateau point. When no information is available

on the plateau location, the uniform prior is recommended with $p_1 = \ldots = p_K = 1/K$. After specifying the prior distributions, the posterior distribution is sampled using the Gibbs sampler.

2.3 Accommodating for delayed efficacy outcome

Unlike toxicity, efficacy often takes a longer follow-up time to assess in practice. Such a "delayed" outcome causes a logistic issue for implementing adaptive designs: when a new patient is enrolled and ready for dose assignment, the patients who have been treated in the trial have not finished their efficacy assessment yet and thus their efficacy outcomes are not available to apply the adaptive rule to assign a dose for the newly accrued patient. Designs in the literature that do not enable to account for time-to-event outcome are forced to "wait" during the efficacy evaluation window. Note that we cannot ignore the data information from these partially assessed (or followed) patients, otherwise the resulting estimates are biased [17]. Recently, Liu et al [10] and Jin et al [12] proposed a systematical approach to handle delayed outcomes for early phase clinical trials based on Bayesian data augmentation, which enjoys attractive theoretical and practical properties. For simplicity, we here take the approach of Cheung and Chappell [18] by weighting the observed data likelihood with the follow-up time. Specifically, let T be a fixed time window for evaluating efficacy, and t_i denote the time-to-efficacy of the i^{th} patient. Let $C_{i,I}$ denote the follow-up time for patient i prior to the entry of the $(I+1)$-th patient (ranging between 0 and T; i.e. $C_{i,I} =$

$\min(\text{time arrival of patient}(I+1) - \text{time arrival of patient} i, T))$, and $z_{i,I} = \mathbb{1}(t_i < C_{i,I})$ denote the response indicator of patient i prior to the entry of the $(I+1)$-th patient. Before the dose assignment of the $(I+1)$-th patient (corresponding to a new cohort), the weighted likelihood of the efficacy data $\mathcal{D}_{\text{eff}} = \{(x_1, z_{1,I}, w_{1,I}), \ldots, (x_I, z_{I,I}, w_{I,I})\}$ obtained from the first I patients is given by

$$L(\mathcal{D}_{\text{eff}}|\gamma_0, \gamma_1, \tau) = \prod_{i=1}^{I} \left(w_{i,I}\phi_{x_i}\right)^{z_{i,I}} \left(1 - w_{i,I}\phi_{x_i}\right)^{1-z_{i,I}},$$

where $w_{i,I}$ takes the form of "adaptive" weights proposed by Cheung and Thall[19], given by

$$w_{i,I} = \begin{cases} 1 & \text{if } t_i \leq C_{i,I} \\ \dfrac{\#\left\{j : t_j \leq C_{i,I} \text{ and } C_{j,I} = T\right\} + C_{i,I}/T}{\#\left\{j : t_j \leq T \text{ and } C_{j,I} = T\right\} + 1} & \text{if } t_i > C_{i,I}. \end{cases}$$

where $\#\left\{j : t_j \leq T \text{ and } C_{j,I} = T\right\}$ is the number of patients who experienced toxicity (i.e., $t_j \leq T$) and completed the followup (i.e., $C_{j,I} = T$) before the entry of the $(I+1)$-th patient; and $C_{i,I}/T$ is the proportion of the time that patient i was followed compared to the full follow-up time T before the entry of the $(I+1)$-th patient. Under this weight function, the data (i.e. $z_{i,I}$) from patients whose efficacy outcomes have been observed receive a full weight of $w_{i,I} = 1$ (i.e., the first line of the weights). For patients whose efficacy outcome has not been observed by the entry of the $(I+1)$-th patient, their weight is the ratio between the number of patients that have experienced

efficacy at some point before the follow up of patient i plus the fraction of time of the window T that these patients have been followed up and the total the number of patients that have experienced efficacy within the window $T + 1$ (i.e., the second line of the weights). If many patients experience efficacy early on in the window T (before the point up to which patient i has been followed up), then the weight of a patient i whose efficacy outcome has not been observed will be higher than if many patients at this dose experience efficacy later on. If, however, efficacy occurs later on in the window T then higher weight is only given the longer a patient has been observed. In general, $w_{i,I}$ monotonically increases with the follow-up time $C_{i,I}$ from 0 to 1. That is, the longer we follow a patient, the more confidence we have about that patient's current efficacy outcome.

2.4 Plateau estimation

Using the Gibbs sampler, we have estimated the posterior probability, π_k, of the k^{th} dose being the plateau point, i.e. $\mathrm{P}(\tau = k|\text{data})$, given by

$$\pi_k = \frac{p_k \iint L(\gamma_0, \gamma_1|k, \mathcal{D}_{\text{eff}}) f(\gamma_0, \gamma_1) d\gamma_1 d\gamma_0}{\sum_{\tau=1}^{K} p_\tau \iint L(\gamma_0, \gamma_1|\tau, \mathcal{D}_{\text{eff}}) f(\gamma_0, \gamma_1) d\gamma_1 d\gamma_0}.$$

The estimation of the plateau is difficult because it poses an ethical dilemma in the sense that there is a conflict between estimating correctly the optimal dose (which

requires allocating patients to most or all doses) and allocating more patients to the currently estimated best dose level (which is, in a sense, being greedy). If the dose allocation is greedy too early, then it is more likely that the recommended dose level at the end of the trial will be incorrect. This is because the greedy algorithms that choose each successive action by myopically optimizing a decision criterion can get stuck at a suboptimal action [20, 21, 22]. In this paper, being myopic would be to select the plateau location at the dose level corresponding to the highest posterior probability (as in [13]). The algorithm, by repeatedly selecting the sub-optimal action, fails to obtain enough data to identify the true optimal dose at the end of the trial. On the other hand, if the dose allocation is greedy too late, then many patients will be exposed to wrong dose levels. This problem is sometimes known as the "exploitation versus exploration" dilemma, which has been recognized in dose-finding clinical trials [23, 24, 25] and other contexts, e.g., phase II-III trials [26, 27, 28].

To address the above issue, we proposed two new dose-assignment rules as follows,

- **MTA-RA - Adaptive randomization:** Randomize the plateau to a dose location based on π_k's. Specifically, let $\mathcal{R}$ denote the set of doses whose posterior probabilities of being the plateau point was close to the largest one with a difference less than a positive threshold s_1, i.e.

$$\mathcal{R} = \left\{ j : \left| \max_{1 \le k \le K} (\pi_k) - \pi_j \right| \le s_1; \quad 1 \le j \le K \right\}.$$

In other words, $\mathcal{R}$ contains a set of doses that are most likely to be the plateau point. We sample the plateau estimate to dose $k \in \mathcal{R}$ with a re-normalized probability $\pi_k \Big/ \sum_{j \in \mathcal{R}} \pi_j$. The value of the threshold s_1 should be calibrated by simulation studies. In practice, this can be done as follows: first define a set of representative dose-toxicity scenarios that may be encountered in the trial, and then conduct simulation under different values of s_1 to evaluate the performance of the design. This is a trial-and-error process and may involve repeatedly tuning the values of s_1 based on the simulation results. The goal is to find the values of s_1 that yield good overall performance across different scenarios (e.g., the percentage of correct selection, the number of patients exposed to over-toxic doses or under-toxic doses). Such a calibration based approach has been widely used in clinical trial designs [29, 30, 31, 32, 33]. One version of the threshold we found that generally works well in our simulation study is $s_1 = 0.20\left(1 - \frac{I}{N}\right)$. By letting s_1 depend on the current sample size I, the threshold is more liberal at the beginning of the trial when we have high uncertainty on model estimates, and the threshold becomes more stringent toward the end of the trial when we have more data to estimate the model. The posterior efficacy probabilities, ϕ_k, are then estimated as the mean of the efficacy probabilities calculated using only the MCMC samples including the estimated plateau parameter. Whatever the chosen threshold, at the end of the trial the plateau location is determined at the dose with the highest posterior probability so that there is no more randomization for

dose recommendation.

- **MTA-PM - Posterior efficacy probabilities:** Assign plateau to the highest dose where we see a big drop on the estimates of efficacy probabilities. More precisely, fix the plateau location at each possible dose level and estimate posterior efficacy probabilities. Then perform Bayesian model averaging on the estimated posterior efficacy probabilities

$$\phi_k = \phi(d_k, \gamma_0, \gamma_1, \tau) = \sum_{\tau=1}^{K} \phi(d_k, \gamma_0, \gamma_1|\tau)\tilde{\pi}_\tau(\tau)$$

  The plateau is determined at dose

$$k = \max\left\{j; 1 \leq j \leq K, \hat{\phi}_j - \hat{\phi}_{j-1} \geq s_2\right\},$$

  where cutoff s_2 can be interpreted as the minimal efficacy difference of practical importance. The value of s_2, which is a constant, should be calibrated by simulation to ensure good design operating characteristics.

Efficacy probabilities, ϕ_k, are then re-evaluated so that efficacy at doses higher than the plateau are set equal to the efficacy posterior probability estimated at the plateau point.

We also tried other model selection criteria for plateau determination such as the DIC (Deviance Information Criterion) or the PPL (Posterior Predictive Loss) which

we chose not to present in this paper as results were not satisfying. Notice, the two approaches presented enable determination of the plateau location at each step. They are expressed in terms of plateau parameter randomization. Equivalently we could have estimated posterior probabilities and expressed them in terms of patients randomization at a dose. We choose this way of presenting the method as the description is more readable and intuitive.

2.5 Dose-finding algorithm

At the beginning of the trial, the posterior estimates of toxicity and efficacy probabilities typically are not reliable due to the limited amount of data [34, 35, 36, 37, 33]. To gather enough information for estimating model parameters, we implement the following start-up phase. Taking a cohort size of 3, we treat the first cohort of patients at the lowest dose level 1, and if no toxicity is observed, we escalate to dose level 2 for treating the second cohort. We continue this dose escalation until we encounter the first toxicity. Once a toxicity is observed or the highest dose level is reached, the start-up phase ends and we switch to the model-based dose-finding phase, where patients are treated in cohort size of c. If $c = 1$, patients are treated one by one.

Let θ and ξ denote the pre-specified toxicity upper bound and efficacy lower bound, respectively. Let $n_{k,I}$ denote the number of patients treated at dose level k. We shall consider that dose level k is *admissible* if it satisfies the safety requirement

$$P(\psi_k > \theta) < L_{\mathrm{T}} \tag{5}$$

and also the efficacy requirement

$$P(\phi_k > \xi) \geq L_{\mathrm{E}} \mathbb{1}(n_{k,I} > \max(c, 3)), \tag{6}$$

where L_{T} and L_{E} are the respective probability thresholds for toxicity and efficacy. Note that the efficacy requirement (6) at one dose takes effect only when more than one cohort (or 3 patients) is treated at that dose, as controlled by the indicator function $\mathbb{1}(n_{k,I} > \max(c, 3))$. This is to ensure that we have some data to reliably evaluate the efficacy criterion, given that the efficacy model is relatively complicated.

Let k denote the current dose level and h denote the highest dose level that has been used previously to treat patients, prior to the entry of the new cohort of patients. We use $\mathcal{B} = \{k'; 1 \leq k' \leq \max(\min(k+1, K), h)\}$ to denote the set of doses that are not one level higher than the current dose k or more than h, and $\mathcal{A}$ to denote the set of admissible doses in $\mathcal{B}$. To assign a dose to the incoming cohort, we fit the proposed model using the data collected from the first I patients enrolled into the trial. The next incoming $(I+1)$-th patient (or the new cohort) is assigned to the dose level with the highest efficacy in $\mathcal{A}$:

$$d^{\mathrm{next}} = \min\left(\underset{j \in \mathcal{A}}{\operatorname{argmax}}\left(\hat{\phi}_j\right)\right)$$

If several dose levels are possible, it means that the efficacy plateau is reached and

among them the dose with the lowest toxicity is selected. Note that although our dose assignment rule seems greedy (i.e., always assign the incoming patient to the dose level with the highest estimate of efficacy), it actually is not. This is because, as described previously, our estimate of the plateau (i.e., the optimal dose with the highest efficacy) is randomized among several promising doses. As a result, under our dose assignment rule, the new patient has chances to be assigned to the doses whose estimate of efficacy are not, but close to, the highest, thereby overcoming the drawback of the conventional greedy algorithm that tends to get stuck in some local suboptimal doses.

We continue the above dose assignment processes until the maximum sample size is reached. At the end of the trial, we select the optimal dose as the lowest dose level that is admissible and has the highest estimate of efficacy among all doses tested during the trial. At any time during the model-based dose-finding phase, if all doses are not admissible, we terminate the trial to protect patients from overly toxic or futile doses.

To summarize, after a start-up phase, the two proposed rules estimate the dose of the plateau location and then assign the next cohort to the admissible dose with the highest estimated efficacy and among them with the lowest toxicity according to plateau location.

3 Numerical Studies

3.1 Simulation

We simulated 2000 independent phase I/II trials to evaluate the operating characteristics of the proposed design. We assumed 6 dose levels and considered 10 scenarios (Table 1) with different locations of the true optimal dose. These scenarios cover a wide range of dose-toxicity and dose-efficacy relationships we may encounter in practice. The pre-specified toxicity upper bound and efficacy lower bound was fixed at $\theta = 0.35$ and $\xi = 0.20$, respectively. The maximum sample size was $N = 60$ and the cohort size was $c = 3$ patients. The trial started at the lowest dose d_1. We took the initial guesses of toxicity and efficacy probabilities as (0.02, 0.06, 0.12, 0.20, 0.30, 0.40) and (0.12, 0.20, 0.30, 0.40, 0.50, 0.59), respectively, to obtain the "effective" doses u_k and v_k used in the toxicity and efficacy models. First, we have considered toxicity and efficacy independently, ignoring possible correlation.

We set the toxicity threshold as $L_{\mathrm{T}} = 0.90$ and the efficacy threshold as $L_{\mathrm{E}} = 0.40$. We assumed that the patient accrual followed a Poisson process with the rate of 0.28 patients per week, i.e. approximately one patient every 3.5 weeks. We assumed that toxicity was quickly ascertainable, as the majority of phase I trial designs, such that when the next new patient arrives, the toxicity outcome has been observed for all patients who have enrolled into the trial. We supposed that the evaluation of efficacy required 7 weeks, i.e. $T = 7$ weeks. Under each scenario, we assumed that at each dose,

the time to efficacy followed an exponential distribution $Exp(\lambda)$. The parameter λ was chosen such that at the end of the follow-up, the efficacy rate of each dose matched those displayed in Table 1, i.e., $\lambda = -\log(1-\phi_k)/T$. As a result, the value of λ had to vary across doses.

For convenience, we refer to the proposed design with the first plateau determination procedure based on randomization as the MTA-RA, and the proposed design with the second plateau determination procedure based on posterior mean of efficacy probabilities as the MTA-PM.

3.2 Compared designs

We compared the proposed design with the methods proposed by Hunsberger et al [4] (denoted as HRDK) and the method proposed by Thall and Cook [38] (denoted as TC). As HRDK and TC designs assume that the efficacy endpoint is binary, quickly ascertainable, when implementing these two designs, we waited for the efficacy response of treated patients to be completely observed before enrolling a new cohort of patients. As a result, the comparison inherently favors the HRDK and TC designs because these designs are based on full information (i.e., all patients' outcomes are fully observed), while our approach is based on partial information. For the same reason, our method leads to substantially shorter trial durations because it does not require waiting for the efficacy outcome and treats patients in real time.

HRDK

HRDK is placed in the same context of efficacy plateau than our method. However, HRDK design considers efficacy as the only endpoint for dose finding and assumes that the efficacy is binary and immediately ascertainable. Toxicity is assumed acceptable and not taken into account. Therefore, for the simulation study, the results comparison can only be made when toxicity is acceptable at all dose levels or when the plateau is reached before the toxicity upper bound. For HRDK, we took the cohort size as 10 such that the total sample size matches that of the proposed design (i.e., $N = 60$) when all six dose levels are explored. We implemented this approach in R using "lm" function.

TC

TC is a more complex phase I/II Bayesian dose-finding design based on trade-offs between the probabilities of efficacy and toxicity. It enables to select the optimal dose in terms of both toxicity and efficacy. For TC, as for the proposed method, the same maximum sample size for 60 patients and cohort size of 3 patients were used. The toxicity and efficacy upper and lower limits were specified as $\overline{\pi}_T = 0.35$ and $\underline{\pi}_E = 0.20$, and p_T and p_E are fixed probability cutoffs both chosen equal to 0.10 as chosen in the original paper [38]. Simulations were performed using the free "EffTox" software available on the MD Anderson Biostatistics software website (`https://biostatistics.mdanderson.org/SoftwareDownload`). As no data are available from real clinical trials, we let the default parameters in the software for other settings.

3.3 Scenarios

In practice, when a dose-finding design missed the optimal dose, but selected a safe dose that has the same efficacy as the optimal dose (although it may have higher dose level than the optimal dose), it is still of interest to physicians. We refer to the set of admissible doses that have the same efficacy as the optimal dose as "correct" doses. For example, in scenario 1, dose level 6 is a "correct" dose, because it has the same efficacy probability of 0.8 as the optimal dose (i.e., dose level 5).

Table 1: Selection percentage and number of patients (shown in parentheses) allocated to each dose under the HRDK design, TC design, and the proposed MTA-RA and MTA-PM designs. In bold are given the optimal doses, and correct doses are underlined.

Design	Dose levels 1		2		3		4		5		6		Early stop
Scenario 1													
(ψ_k, ϕ_k)	(0.005,0.01)		(0.01,0.10)		(0.02,0.30)		(0.05,0.50)		**(0.10,0.80)**		(0.15,0.80)		
HRDK	0.1	(10.0)	0.6	(10.0)	1.3	(10.0)	1.3	(9.9)	**38.2**	(9.8)	58.4	(9.8)	0.0
TC	0.0	(3.2)	0.0	(3.0)	0.0	(3.2)	4.5	(6.9)	**45.0**	(23.5)	50.5	(20.2)	0.0
MTA-RA	0.0	(3.3)	0.1	(3.7)	0.6	(5.0)	4.6	(8.5)	**58.5**	(21.4)	35.7	(17.8)	0.7
MTA-PM	0.2	(3.9)	0.1	(3.7)	7.3	(6.9)	12.3	(9.3)	**68.7**	(29.8)	10.1	(6.0)	1.4
Scenario 2													
(ψ_k, ϕ_k)	**(0.01,0.40)**		(0.05,0.40)		(0.10,0.40)		(0.25,0.40)		(0.50,0.40)		(0.70,0.40)		
HRDK	**25.3**	(10.0)	26.4	(10.0)	23.3	(10.0)	12.8	(5.8)	7.7	(3.5)	4.5	(1.5)	0.0
TC	**79.5**	(48.3)	1.5	(1.3)	11.5	(4.8)	7.5	(4.9)	0.0	(0.5)	0.0	(0.1)	0.0
MTA-RA	**63.5**	(17.0)	19.3	(13.0)	8.3	(10.9)	5.4	(11.7)	2.7	(6.4)	0.4	(0.8)	0.5
MTA-PM	**70.2**	(34.4)	18.6	(13.3)	8.8	(8.0)	1.4	(3.0)	0.2	(1.0)	0.0	(0.1)	0.9
Scenario 3													
(ψ_k, ϕ_k)	(0.01,0.25)		(0.02,0.45)		**(0.05,0.65)**		(0.10,0.65)		(0.20,0.65)		(0.30,0.65)		
HRDK	0.7	(10.0)	5.6	(10.0)	**28.9**	(10.0)	26.4	(9.8)	21.5	(8.4)	16.8	(4.7)	0.0
TC	37.0	(27.3)	1.0	(2.1)	**4.5**	(4.2)	38.5	(17.6)	18.5	(7.3)	1.5	(1.4)	0.0
MTA-RA	2.0	(6.8)	14.5	(10.1)	**48.3**	(14.6)	20.1	(12.8)	8.7	(10.7)	6.5	(5.1)	0.0
MTA-PM	8.9	(8.3)	23.5	(14.2)	**49.8**	(23.8)	15.6	(9.8)	1.8	(2.8)	0.0	(0.9)	0.5
Scenario 4													
(ψ_k, ϕ_k)	(0.01,0.05)		(0.02,0.25)		(0.05,0.45)		**(0.10,0.70)**		(0.25,0.70)		(0.50,0.70)		
HRDK	0.1	(10.0)	0.9	(10.0)	3.0	(10.0)	**27.0**	(9.9)	31.6	(9.8)	37.3	(8.9)	0.0
TC	0.0	(4.5)	0.0	(3.2)	1.5	(4.5)	**62.0**	(30.0)	35.5	(16.2)	1.0	(1.5)	0.0
MTA-RA	0.0	(4.0)	1.0	(5.7)	8.5	(10.0)	**53.9**	(17.4)	27.5	(15.7)	8.9	(7.1)	0.4
MTA-PM	0.3	(4.3)	3.5	(5.6)	22.6	(13.4)	**47.0**	(22.0)	24.8	(12.9)	0.2	(1.0)	1.7
Scenario 5													
(ψ_k, ϕ_k)	(0.01,0.10)		(0.05,0.35)		**(0.15,0.60)**		(0.20,0.60)		(0.45,0.60)		(0.60,0.60)		
HRDK	0.0	(10.0)	2.6	(10.0)	**27.5**	(10.0)	26.6	(10.0)	25.1	(9.1)	18.0	(5.3)	0.0
TC	2.0	(7.6)	4.0	(5.3)	**30.5**	(15.9)	61.5	(27.6)	3.0	(3.3)	0.0	(0.3)	0.0
MTA-RA	0.1	(5.5)	8.7	(9.1)	**55.4**	(17.4)	26.4	(15.7)	8.4	(10.2)	1.0	(2.0)	0.2
MTA-PM	1.5	(5.8)	12.8	(10.1)	**53.4**	(25.6)	27.5	(14.7)	2.0	(2.7)	0.0	(0.1)	2.9
Scenario 6													
(ψ_k, ϕ_k)	(0.01,0.05)		(0.05,0.10)		(0.10,0.20)		(0.20,0.35)		**(0.30,0.55)**		(0.50,0.55)		
HRDK	3.9	(10.0)	5.4	(10.0)	6.4	(10.0)	5.2	(8.9)	**33.8**	(8.5)	45.2	(8.3)	0.0
TC	0.0	(4.8)	0.5	(3.7)	11.5	(8.3)	37.0	(18.8)	**41.5**	(17.6)	5.5	(5.2)	4.0
MTA-RA	0.0	(4.7)	1.0	(5.8)	4.7	(8.6)	18.5	(13.0)	**55.2**	(17.8)	12.7	(7.2)	8.0
MTA-PM	3.2	(8.7)	0.6	(5.7)	8.6	(9.3)	19.9	(11.9)	**39.5**	(15.5)	2.6	(2.6)	25.7
Scenario 7													
(ψ_k, ϕ_k)	(0.02,0.30)		(0.07,0.50)		**(0.13,0.70)**		(0.17,0.73)		(0.25,0.76)		(0.30,0.77)		
HRDK	1.0	(10.0)	4.5	(10.0)	**20.1**	(10.0)	22.6	(9.8)	24.6	(8.8)	27.2	(6.1)	0.0
TC	57.5	(37.2)	2.0	(2.4)	**12.0**	(7.0)	23.5	(10.9)	4.0	(2.0)	1.0	(0.5)	0.0
MTA-RA	1.4	(6.1)	8.6	(9.0)	**38.7**	(15.1)	22.9	(13.8)	16.6	(11.1)	11.8	(4.9)	0.0
MTA-PM	10.1	(8.8)	22.9	(15.1)	**48.9**	(24.6)	16.2	(9.3)	1.5	(1.7)	0.0	(0.4)	0.5
Scenario 8													
(ψ_k, ϕ_k)	(0.03,0.30)		(0.06,0.50)		(0.10,0.52)		(0.20,0.54)		(0.40,0.55)		(0.50,0.55)		
HRDK	4.2	(10.0)	**23.5**	(10.0)	24.3	(10.0)	21.8	(8.9)	14.7	(5.8)	11.5	(3.5)	0.0
TC	55.5	(35.7)	**2.5**	(2.6)	14.0	(7.2)	25.5	(12.4)	2.0	(1.8)	0.0	(0.4)	0.0
MTA-RA	13.5	(10.4)	**43.7**	(15.0)	20.0	(12.8)	12.5	(12.0)	8.3	(8.1)	2.0	(1.8)	0.1
MTA-PM	25.5	(16.0)	**43.9**	(22.4)	24.6	(15.1)	5.1	(4.8)	0.4	(1.3)	0.0	(0.2)	0.6
Scenario 9													
(ψ_k, ϕ_k)	(0.05,0.01)		(0.10,0.02)		(0.25,0.05)		(0.55,0.35)		(0.70,0.55)		(0.90,0.70)		
HRDK	4.4	(10.0)	1.1	(10.0)	2.6	(10.0)	2.3	(9.2)	15.8	(9.1)	73.6	(9.0)	0.0
TC	0.0	(3.4)	0.0	(3.2)	8.5	(9.5)	9.5	(11.5)	0.5	(1.6)	0.0	(0.3)	**81.5**
MTA-RA	0.0	(5.8)	0.0	(5.9)	3.0	(7.5)	5.5	(10.9)	0.0	(2.3)	0.0	(0.3)	**91.6**
MTA-PM	0.1	(6.1)	0.0	(6.0)	2.7	(6.9)	6.6	(10.2)	0.2	(3.2)	0.0	(0.3)	**90.4**
Scenario 10													
(ψ_k, ϕ_k)	(0.50,0.40)		(0.60,0.55)		(0.69,0.65)		(0.76,0.65)		(0.82,0.65)		(0.89,0.65)		
HRDK	4.6	(10.0)	14.4	(10.0)	28.7	(10.0)	23.8	(9.1)	16.7	(6.6)	11.8	(3.6)	0.0
TC	10.5	(18.7)	2.0	(0.8)	0.0	(0.3)	0.0	(0.1)	0.0	(0.0)	0.0	(0.0)	**87.5**
MTA-RA	10.9	(16.1)	0.2	(4.0)	0.0	(1.0)	0.0	(0.2)	0.0	(0.0)	0.0	(0.0)	**89.0**
MTA-PM	10.9	(16.3)	0.1	(4.7)	0.0	(0.7)	0.0	(0.0)	0.0	(0.0)	0.0	(0.0)	**89.0**

3.4 Results

As shown in Table 1 and Table 2, the proposed MTA-RA and MTA-PM designs generally outperformed HRDK design. TC performed better than MTA approaches in scenarios 2 and 4, while MTA-RA performed as well as or better in terms of both the selection of the optimal and correct dose levels and the number of patients allocated to the optimal doses in scenarios 1, 3, 6, 7 and 8. In scenario 1, all doses were safe with toxicity probabilities lower than the upper limit, and thus the dose selection was largely guided by efficacy. The percentage of correct selection of the optimal dose (PCS-OD) was 38.2%, 45.0% , 58.5% and 68.7% under HRDK, TC, MTA-RA and MTA-PM, respectively. The MTA-PM design performed best with the highest PCS-OD. In scenarios 2 the plateau was reached at the lowest dose level, in this case TC performs better. We can observe for this scenario that the PCS-OD of TC and the proposed designs more than doubled that of the HRDK design. In scenario 3, the third dose is the optimal dose. The PCS-OD of the proposed designs were about 50%, while that of the HRDK design was 28.9% and that of the TC design fall down to 4.5%. The TC design tended to select a higher dose (i.e., the fourth dose) as the optimal dose. Similar results were observed in scenarios 4 to 6. In scenarios 7 and 8, where the efficacy does not exactly plateau but increases by small amount until a difference of 10% from the efficacy plateau, the proposed designs were still able to select the target dose with the highest percentage and substantially outperformed the HRDK and TC designs. In scenarios 9 and 10, none of the doses were admissible and the trial should be terminated. Specifically, in

Table 2: Selection percentages of the optimal dose and correct doses.

	Scenarios							
Design	1	2	3	4	5	6	7	8
	Optimal dose selection percentage							
HRDK	38.2	25.3	28.9	27.0	27.5	33.8	20.1	23.5
TC	45.0	79.5	4.5	62.0	30.5	41.5	12.0	2.5
MTA-RA	58.5	63.5	48.3	53.9	55.4	55.2	38.7	43.7
MTA-PM	68.7	70.2	49.8	47.0	53.4	39.5	48.9	43.9
	Correct dose selection percentage							
HRDK	96.6	87.8	93.6	58.6	54.1	33.8	94.5	69.6
TC	95.5	100.0	63.0	97.5	92.0	41.5	40.5	42.0
MTA-RA	94.2	96.5	83.6	81.4	81.8	55.2	90.0	76.2
MTA-PM	78.8	99.0	67.2	71.8	80.9	39.5	66.6	73.6

scenario 9, the first three doses was safe but their efficacy was unacceptably low, and the remaining doses were overly toxic; in scenario 10, all doses were overly toxic. In these two scenarios, the two proposed designs early terminated the trial more than 90% of the time, as well as TC in 81.5%. The HRDK design did not consider toxicity, thus failed to stop the trial in both scenarios.

Table 2 summarizes the selection percentage of the optimal dose and correct doses under 8 scenarios. We can see that the selection percentage of the "correct" doses under the proposed designs were mostly over 70%. Between the two proposed design, MTA-RA appeared to have slightly better and more robust performance than the MTA-PM, thus we recommend it for general use in practice.

3.5 Sensitivity analysis

3.5.1 Non-monotonic dose-efficacy relationships

Our designs assume that the dose-efficacy curve increases initially and then plateaus. We conducted a sensitivity analysis to evaluate the performance of the proposed design when the dose-efficacy curve was umbrella shaped (i.e., efficacy first increases and then decreases, see scenarios 1 and 2 in Table 3), monotonically increasing (scenario 3 in Table 3), or monotonically decreasing (scenario 4 in Table 3). The sample size was 36 patients with a cohort size of 3, and we took $\theta = 0.35$ and $\xi = 0.20$. The simulation results (see Table 3) show that the proposed design performed well under these different shapes of dose-efficacy curves. The selection percentages of the target dose were comparable to these reported in Table 1, suggesting that the proposed designs were not sensitive to the violation of the increasing-then-plateau assumption.

3.5.2 Prior distributions

We also evaluate the robustness of the proposed designs in terms of prior distribution. We changed the prior distributions of slope parameters β_1 (for toxicity) and γ_1 (for efficacy) from $Exp(1)$ to gamma distribution $Gamma(0.5, 0.5)$. Results are provided in Table 4. We observe that both proposed methods are generally robust to the choice of prior distributions. The results are similar to these reported in Table 1.

Table 3: Sensitivity analysis of the proposed MTA-RA and MTA-PM designs. In bold are given the optimal doses.

Design	Dose levels								None
	1		2		3		4		
	Scenario 1								
	(0.01, 0.10)		(0.05, 0.35)		**(0.15, 0.60)**		(0.25, 0.30)		
HRDK	0.1	(9.0)	7.6	(9.0)	**81.5**	(9.0)	10.8	(8.9)	0.0
TC	2.0	(4.8)	6.0	(5.0)	**46.0**	(12.6)	45.0	(13.4)	1.0
MTA-RA	1.2	(4.8)	19.6	(8.6)	**73.9**	(17.1)	3.5	(5.3)	2.0
MTA-PM	13.0	(9.6)	26.5	(10.1)	**49.1**	(12.9)	0.8	(2.6)	10.7
	Scenario 2								
	(0.10, 0.50)		**(0.20, 0.70)**		(0.30, 0.60)		(0.50, 0.40)		
HRDK	7.4	(9.0)	**54.6**	(9.0)	34.6	(9.0)	3.3	(5.2)	0.0
TC	67.0	(25.0)	**25.0**	(7.5)	7.0	(2.7)	0.0	(0.6)	1.0
MTA-RA	30.4	(11.6)	**59.6**	(14.9)	9.4	(7.3)	0.5	(2.1)	0.2
MTA-PM	38.2	(15.2)	**58.4**	(17.9)	3.2	(2.4)	0.0	(0.4)	0.3
	Scenario 3								
	(0.05, 0.02)		(0.10, 0.28)		(0.16, 0.50)		**(0.22, 0.80)**		
HRDK	0.0	(9.0)	0.7	(9.0)	5.1	(9.0)	**94.2**	(8.9)	0.0
TC	0.0	(3.2)	2.0	(4.4)	19.0	(9.0)	**78.0**	(19.0)	1.0
MTA-RA	0.0	(3.8)	4.4	(5.6)	10.7	(8.0)	**82.7**	(18.2)	2.3
MTA-PM	1.8	(5.9)	13.4	(7.7)	27.8	(11.1)	**49.1**	(10.4)	8.0
	Scenario 4								
	(0.05,0.80)		(0.10,0.50)		(0.16,0.28)		(0.22,0.02)		
HRDK	**86.5**	(9.0)	12.5	(9.0)	1.0	(9.0)	0.0	(0.0)	0.0
TC	**99.0**	(35.0)	1.0	(0.7)	0.0	(0.2)	0.0	(0.1)	0.0
MTA-RA	**99.4**	(26.3)	0.5	(5.2)	0.0	(3.1)	0.0	(1.4)	0.1
MTA-PM	**98.5**	(29.5)	1.5	(3.3)	0.0	(2.0)	0.0	(1.2)	0.1

Table 4: Sensitivity analysis with different prior distributions.

	Different prior distributions: β_1, $\gamma_1 \sim \Gamma(0.5, 0.5)$							
	sc1	sc2	sc3	sc4	sc5	sc6	sc7	sc8
	Optimal dose selection percentage							
MTA-RA	61.1	61.9	45.7	56.0	54.8	50.2	38.2	43.5
MTA-PM	68.5	85.0	45.9	49.3	50.5	40.5	39.7	34.3
	Correct dose selection percentage							
MTA-RA	90.9	96.4	80.8	81.4	78.6	50.2	86.4	74.9
MTA-PM	78.4	99.2	62.8	71.8	76.6	40.5	56.2	57.3

4 Conclusion

We have developed a phase I/II Bayesian dose-finding design for a molecularly targeted agent alone or in combination with a fixed dose of a cytotoxic agent. Our design takes into account that the efficacy curve may plateau as the dose level increases and focuses on selecting the optimal dose, i.e. the dose with the lowest toxicity among those with highest efficacy.

We have proposed two different allocation methods. One based on adaptive randomization with posterior probabilities for the plateau parameter, and the other one based on the difference between the posterior mean of efficacy probabilities according to the plateau parameter. On the contrary to our previous work [13] that used a greedy algorithm for dose estimation-selection based on highest posterior probability, our proposed approaches have found to be robust, and enable to avoid getting stuck in local maxima when considering many doses. Both allocations give good and similar perfor-

mance in terms of PCS-OD, but MTA-RA seems more robust across scenarios (always above 50%). Moreover, it also gives a higher percentage of selection of a correct dose level (Table 2), that is dose levels with the highest efficacy but not necessarily lowest toxicity under toxicity restrictions. Our approach outperforms the HRDK design [4], and performs as well as or better than TC [38] in terms of both the selection of the optimal and correct dose levels and the number of patients allocated to the optimal doses in most scenarios (1, 3, 6, 7 and 8, while TC performed better in scenarios 2 and 4). It must be noted that TC contains many settings that need to be calibrated to which the design are very sensitive. In the absence of clinical trials data, we used the default values of the provided software.

Furthermore, we have also considered the possibility to apply our design to non-monotone relationships. In these cases, our design also gives good performance in general, but MTA-RA performed better in the common case where no plateau is observed across all dose levels. For all these reasons, when a statistician is involved in a clinical trial where a plateau efficacy or a unimodal relationship is expected, we recommend to use the MTA-RA design. We developed an R-package *dfmta* which is available on CRAN.

We have assumed that toxicity is binary and quickly ascertainable while efficacy may be delayed or late-onset. In some cases, toxicity may also require a long period of time to be evaluated. We can accommodate such delayed toxicity using a similar weighting approach. Jin et al [12] considered a more principled missing data approach to

handle both delayed toxicity and efficacy, which can also be adopted here. We also plan to develop a new methodology to deal with more complex trials. We are developing a method for ordinal toxicity and efficacy outcomes using a multinomial logistic regression model, and selecting the dose based on a utility function. To avoid the exploitation versus exploration dilemma, dynamic and non-myopic rules could be explored [26, 39, 28], which enable to appropriately balance the dilemma between learning about doses' efficacy and allocating more patients to the best dose level. We are pursuing this work with new collaborations; some authors of this paper are exploring the possibility to apply Bandit algorithms [40, 41] to standard CRM and then to extend it to unimodal relationship.

Supplementary material

Results of sensitivity analysis on prior distributions can be found in supplementary material.

Acknowledgments

We would like to thank the two reviewers for their advice and constructive comments that have served to substantially improve the manuscript. This work was partially funded by grants from the ANRT (Association Nationale de la Recherche et de la Technologie), "Laboratoires Servier" (CIFRE number 2011/0900), Denis Diderot University-

Paris 7 ("Aide à la mobilité internationale 2012"), INSERM (French National Institute of Health and Medical Research), and by Award Number CA154591, 5P50CA098258 and CA016672 from the National Cancer Institute.

References

[1] Department of Health US, Human Services F, Drug Administration CfBE, Research. Clinical Considerations for Therapeutic Cancer Vaccines. Guidance for Industry. 2011;.

[2] Ellis LM. Antiangiogenic therapy: more promise and, yet again, more questions. J Clin Oncol. 2003;21(21):3897–3899.

[3] Morgan B, Thomas AL, Drevs J, Hennig J, Buchert M, Jivan A, et al. Dynamic contrast-enhanced magnetic resonance imaging as a biomarker for the pharmacological response of PTK787/ZK 222584, an inhibitor of the vascular endothelial growth factor receptor tyrosine kinases, in patients with advanced colorectal cancer and liver metastases: results from two phase I studies. J Clin Oncol. 2003;21(21):3955–3964.

[4] Hunsberger S, Rubinstein LV, Dancey J, Korn EL. Dose escalation trial designs based on a molecularly targeted endpoint. Stat Med. 2005;24(14):2171–2181.

[5] Hirakawa A. An adaptive dose-finding approach for correlated bivariate binary and continuous outcomes in phase I oncology trials. Stat Med. 2012;31(6):516–532.

[6] Cai C, Yuan Y, Ji Y. A Bayesian Phase I/II Design for Oncology Clinical Trials of Combining Biological Agents. Journal of the Royal Statistical Society: Series C. 2013;Early view article:DOI: 10.1111/rssc.12039.

[7] Ivanova A, Xiao C. Dose finding when the target dose is on a plateau of a dose-response curve: comparison of fully sequential designs. Pharm Stat. 2013;12(5):309–314.

[8] Zhang W, Sargent DJ, Mandrekar S. An adaptive dose-finding design incorporating both toxicity and efficacy. Stat Med. 2006;25:2365–2383.

[9] Stiles JM, Amaya C, Rains S, Diaz D, Pham R, Battiste J, et al. Targeting of beta adrenergic receptors results in therapeutic efficacy against models of hemangioendothelioma and angiosarcoma. PLoS ONE. 2013;8(3):e60021.

[10] Liu S, Yin G, Yuan Y. Bayesian data augmentation dose finding with continual reassessment method and delayed toxicity. Ann Appl Stat. 2013;7(4):2138–2156.

[11] Yuan Y, Yin G. Bayesian Phase I/II Adaptively Randomized Oncology Trials With Combined Drugs. Ann Appl Stat. 2011;5(2A):924–942.

[12] Jin IH, Liu S, Thall PF, Yuan Y. Using Data Augmentation to Facilitate Conduct of Phase III Clinical Trials With Delayed Outcomes. Journal of the American Statistical Association. 2014;109(506):525–536.

[13] Riviere MK, Yuan Y, Dubois F, Zohar S. A Bayesian dose finding design for clinical trials combining a cytotoxic agent with a molecularly targeted agent. Journal of the Royal Statistical Society: Series C (Applied Statistics). 2014;64(1):215–229.

[14] 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.

[15] Chevret S. The continual reassessment method in cancer phase I clinical trials: a simulation study. Stat Med. 1993;12(12):1093–1108.

[16] Liu S, Ning j. Bayesian dose-finding design for drug combination trials with delayed toxicities. Bayesian Analysis. 2013;8:703–722.

[17] Yuan Y, Yin G. Robust EM Continual Reassessment Method in Oncology Dose Finding. Journal of the American Statistical Association. 2011;106:818–831.

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

[19] Cheung YK, Thall PF. Monitoring the rates of composite events with censored data in phase II clinical trials. Biometrics. 2002;58(1):89–97.

[20] Thompson WR. On the Likelihood that one Unknown Probability Exceeds Another in View of the Evidence of Two Samples. Biometrika. 1933;25:285–294.

[21] Robbins H. Some aspects of the sequential design of experiments. Bulletin of the American Mathematical Society. 1952;58:527–535.

[22] Gittins JC. Bandit processes and dynamic allocation indices. JRSS, Series B. 1979;41:148–177.

[23] Azriel D, Mandel M, Rinott Y. The treatment versus experimentation dilemma in dose-finding studies. Journal of Statistical Planning and Inference. 2011;141:2759–2768.

[24] Thall PF, Nguyen HQ. Adaptive randomization to improve utility-based dose-finding with bivariate ordinal outcomes. J Biopharm Stat. 2012;22(4):785–801.

[25] Oron AP, Hoff PD. Small-sample behavior of novel phase I cancer trial designs. Clin Trials. 2013;10(1):63–80.

[26] Berry DA, Eick SG. Adaptive assignment versus balanced randomization in clinical trials: a decision analysis. Stat Med. 1995 Feb;14(3):231–246.

[27] Sutton RS, Barto AG. Reinforcement Learning: An Introduction. Cambridge, MA: MIT Press. 1998;.

[28] Villar SS, Bowden J, Wason J. Multi-armed Bandit Models for the Optimal Design of Clinical Trials: Benefits and Challenges. Statist Sci. 2015;30(2):199–215.

[29] Thall PF, Simon R. Practical Bayesian guidelines for phase IIB clinical trials. Biometrics. 1994;50(2):337–349.

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

[31] Yuan Y, Yin G. Bayesian Dose-finding by Jointly Modeling Toxicity and Efficacy as Time-to-Event Outcomes. JRSS. 2009;58:719–736.

[32] Yin G, Yuan Y. Bayesian dose finding in oncology for drug combinations by copula regression. JRSS. 2009;58:211–224.

[33] Yin G, Yuan Y. A latent contingency table approach to dose finding for combinations of two agents. Biometrics. 2009;65:866–875.

[34] Storer BE. Design and analysis of phase I clinical trials. Biometrics. 1989;45:925–37.

[35] Korn EL, Midthune D, Chen TT, Rubinstein LV, Christian MC, Simon RM. A comparison of two phase I trial designs. Stat Med. 1994;13:1799–806.

[36] Ivanova A, Montazer-Haghighi A, Mohanty SG, Durham SD. Improved up-and-down designs for phase I trials. Stat Med. 2003;22(1):69–82.

[37] Ivanova A, Wang K. A non-parametric approach to the design and analysis of two-dimensional dose-finding trials. Stat Med. 2004;23:1861–1870.

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

[39] Jiang F, Jack Lee J, Muller P. A Bayesian decision-theoretic sequential response-adaptive randomization design. Stat Med. 2013 May;32(12):1975–1994.

[40] Audibert JY, Bubeck S, Munos R. Best Arm Identification in Multi-Armed Bandits. In: Proceedings of the 23th annual conference on Computational Learning Theory. Hafa (Isral); 2010. .

[41] Kalyanakrishnan S, Tewari A, Auer P, Stone P. PAC Subset Selection in Stochastic Multi-armed Bandits. In: Langford J, Pineau J, editors. In proceedings of the 29th International Conference on Machine Learning (ICML). New York, NY, USA: Omnipress; 2012. p. 655–662.