



An adaptive trial design to optimize dose-schedule regimes with delayed outcomes

Ruitao Lin | Peter F. Thall | Ying Yuan

Department of Biostatistics, The University of Texas MD Anderson Cancer Center, Houston, Texas

Correspondence

Ruitao Lin, Department of Biostatistics, The University of Texas MD Anderson Cancer Center, Houston, TX 77030.
Email: RLin@mdanderson.org

Funding information

National Institutes of Health, Grant/Award Numbers: CA016672, 1P50CA217685, 5P50CA098258

Abstract

This paper proposes a two-stage phase I-II clinical trial design to optimize dose-schedule regimes of an experimental agent within ordered disease subgroups in terms of the toxicity-efficacy trade-off. The design is motivated by settings where prior biological information indicates it is certain that efficacy will improve with ordinal subgroup level. We formulate a flexible Bayesian hierarchical model to account for associations among subgroups and regimes, and to characterize ordered subgroup effects. Sequentially adaptive decision-making is complicated by the problem, arising from the motivating application, that efficacy is scored on day 90 and toxicity is evaluated within 30 days from the start of therapy, while the patient accrual rate is fast relative to these outcome evaluation intervals. To deal with this in a practical manner, we take a likelihood-based approach that treats unobserved toxicity and efficacy outcomes as missing values, and use elicited utilities that quantify the efficacy-toxicity trade-off as a decision criterion. Adaptive randomization is used to assign patients to regimes while accounting for subgroups, with randomization probabilities depending on the posterior predictive distributions of utilities. A simulation study is presented to evaluate the design's performance under a variety of scenarios, and to assess its sensitivity to the amount of missing data, the prior, and model misspecification.

KEYWORDS

adaptive randomization, Bayesian design, missing data, optimal treatment regime, ordered subgroups, phase I-II clinical trial

1 | INTRODUCTION

The primary objective of a phase I clinical trial is to estimate a maximum tolerable dose (MTD) based on a toxicity variable defined in terms of one or more adverse events. Numerous phase I designs have been proposed, such as the algorithm-based 3 + 3 design (Storer, 1989), many model-based methods including the continual reassessment method (CRM) (O'Quigley *et al.*, 1990), escalation with overdose control (EWOC) (Babb *et al.*, 1998), Bayesian model averaging CRM (Yin and Yuan,

2009), and model-assisted methods (Liu and Yuan, 2015; Zhou *et al.*, 2018). For a comprehensive review on existing phase I designs, see Zhou *et al.* (2018). Some Bayesian model-based methods have been extended to deal with late-onset toxicity (Cheung and Chappell, 2000; Yuan and Yin, 2011a).

For molecularly targeted agents and immunotherapies, a toxicity-based MTD is not necessarily the optimal dose. Many phase I-II trial designs have been proposed to use both efficacy and toxicity for decision-making. Thall and Cook (2004) proposed a Bayesian phase I-II design

based on toxicity-efficacy probability trade-offs. Nebiyou Bekele and Shen (2005) introduced a Bayesian approach to jointly model toxicity and biomarker expression. Zhang *et al.* (2006) utilized a continuation-ratio model to adaptively estimate a biologically optimal dose. Houede *et al.* (2010) optimized the dose pair of a two-agent combination using ordinal toxicity and efficacy, by maximizing posterior mean utility. This approach was extended by using adaptive randomization (AR) to reduce the chance of getting stuck at a suboptimal dose pair (Thall and Nguyen, 2012). Yuan and Yin (2011b) considered phase I-II drug-combination trials with late-onset efficacy. Guo and Yuan (2017) proposed a Bayesian phase I-II design for precision medicine that incorporates biomarker subgroups. Liu *et al.* (2018) developed a Bayesian utility-based phase I-II design for immunotherapy trials. Reviews of phase I-II designs are given by Yuan *et al.* (2016).

In many settings, multiple administration schedules are considered, along with different doses. This motivates more complex phase I or I-II designs to optimize dose-schedule treatment regimes (Braun *et al.*, 2005; 2007; Zhang and Braun, 2013; Lee *et al.*, 2015; Guo *et al.*, 2016). This paper is motivated by a planned phase I-II trial for optimizing the dose-schedule of propylene glycol-free (PGF) Melphalan as a single agent preparative regimen for autologous stem cell transplantation in patients with multiple myeloma (MM). This disease is heterogeneous, dichotomized in terms of pathogenesis pathways determined by genetic and cytogenetic abnormalities as hyperdiploid or not. Hyperdiploid patients are believed to have a better response rate than nonhyperdiploid (Chng *et al.*, 2006). A review is given by Fonseca *et al.* (2009). The trial considers three doses, 200, 225, and 250 mg m⁻², and three infusion schedules, 30 minutes, 12 hours, and 24 hours, yielding nine treatment regimes. Toxicity is defined as the binary indicator of grade 3 mucositis lasting >3 days or any grade 4 or 5 nonhematologic or noninfectious toxicity within 30 days from start of infusion. Efficacy is defined as the binary indicator of complete remission evaluated at day 90. Thus, while toxicity is observed soon enough to apply a usual sequentially adaptive toxicity-based decision rule feasibly, the efficacy outcome is evaluated much later. This greatly complicates making outcome-adaptive decisions to optimize dose or dose-schedule based on both toxicity and efficacy.

Our proposed design optimizes the dose-schedule regime in terms of a toxicity-efficacy risk-benefit trade-off quantified by a utility function, allowing the possibility that the optimal regime may differ between disease subgroups. We formulate a Bayesian hierarchical model to characterize associations among dose, schedule, subgroup, and the bivariate toxicity-efficacy outcome.

The design includes a two-stage AR scheme that randomizes each newly enrolled patient to a treatment regime using the posterior predictive probability of the regime being the best with respect to the utility.

The rest of the paper is organized as follows. Section 2 presents the probability model, likelihood, and priors. Section 3 gives the trial design, including the utility function, prior elicitation, and rules for trial conduct. In Section 4, we apply the proposed design to the motivating example and conduct simulation studies to examine the design's performance. We close with a brief discussion in Section 5.

2 | PROBABILITY MODEL

2.1 | Bayesian hierarchical model

We consider a phase I-II trial with C ordered subgroups and a total of DS treatment regimes obtained from D doses and S treatment schedules. Let n denote the number of patients accrued at an interim point in the trial, and $c_i \in \{1, \dots, C\}$ be the subgroup of the i th patient, $i = 1, \dots, n$. Denote the dose-schedule treatment regime assigned to patient i by $r_i = (d_i, s_i)$, for $d_i \in \{1, \dots, D\}$, $s_i \in \{1, \dots, S\}$, and the joint toxicity and efficacy outcome $\mathbf{Y}_i = (Y_i^T, Y_i^E)$. As \mathbf{Y}_i may depend on both c_i and r_i , the objective of the trial is to find optimal subgroup-specific dose-schedule regimes that can maximize a given utility function.

Similarly to Albert and Chib (1993) and Lee *et al.* (2015), to facilitate posterior computation, we assume a latent normal vector to characterize the joint distribution of the observed discrete outcome vector. Let $\xi_i = (\xi_i^T, \xi_i^E)$ be real-valued bivariate normal latent variables with means that vary with c_i and r_i . We define \mathbf{Y}_i by assuming that $Y_i^j = I(\xi_i^j > 0)$, $j = T, E$, where $I(\cdot)$ is the indicator function, so the joint distribution of the latent vector $[\xi_i | c_i, r_i]$ induces that of the observed vector $[\mathbf{Y}_i | c_i, r_i]$. Each Y_i^j is assumed to be a binary outcome. Extension to ordinal outcomes is straightforward, but introduces additional complexity in the likelihood and utility. We assume the following Bayesian hierarchical model for $[\xi_i | r_i]$:

- (a) *Level 1 prior on ξ_i .* Using patient-specific random effects $\epsilon_i = (\epsilon_i^T, \epsilon_i^E)$, we assume the following conditional distribution for the latent variables:

$$\xi_i^j | c_i, r_i, \epsilon_i, \tilde{\xi}_{c_i, r_i}^j, \sigma_\xi^2 \sim N\left(\tilde{\xi}_{c_i, r_i}^j + \epsilon_i^j, \sigma_\xi^2\right),$$

$$j = E, T, \quad i = 1, \dots, n \quad (1)$$

with the variance σ_ξ^2 a hyperparameter and $\tilde{\xi}_i = (\tilde{\xi}_{c_i, r_i}^T, \tilde{\xi}_{c_i, r_i}^E)$ the mean effects of regime $r_i = (d_i, s_i)$ in subgroup c_i . The following second-level priors on ϵ_i and $\tilde{\xi}_{c_i, r_i}$ induce association between ξ_i^T and ξ_i^E , which, in turn, induces association between Y_i^T and Y_i^E .

(b) *Level 2 prior on ϵ_i .* Assume

$$(\epsilon_i^T, \epsilon_i^E) \stackrel{\text{i.i.d.}}{\sim} \text{BN}(\mathbf{0}_2, \Sigma_\epsilon), \quad i = 1, \dots, n, \quad (2)$$

where ‘‘i.i.d.’’ represents independent and identically distributed, BN denotes a bivariate normal distribution, $\mathbf{0}_2 = (0, 0)$ and Σ_ϵ is the 2×2 matrix with both diagonal elements ζ^2 and both off-diagonal elements $\rho\zeta^2$. The fixed hyperparameters $\rho \in (-1, 1)$ and ζ^2 quantify the association between Y_i^T and Y_i^E via the latent variable model.

(c) *Level 2 prior on $\tilde{\xi}_{c,r}$.* To facilitate information sharing across subgroups, for each regime $r = (d, s)$ we assume that $\tilde{\xi}_{c,r} = \tilde{\xi}_r + \sum_{c'=2}^C \nu_{c,r} I(c' = c)$, $c = 1, \dots, C$, where $\tilde{\xi}_r = (\tilde{\xi}_r^T, \tilde{\xi}_r^E)$ can be treated as the baseline effects for regime r , and the baseline subgroup is $c = 1$ with $\nu_{1,r} = \mathbf{0}_2$ for all r . The ordering constraint is imposed by choosing the support of $\nu_{c,r} = (\nu_{c,r}^T, \nu_{c,r}^E)$ to satisfy the corresponding order constraint. For example, if prior information indicates that $\text{Pr}(\text{efficacy})$ in subgroup 1 is greater than $\text{Pr}(\text{efficacy})$ in subgroup 2, then we require $\nu_{1,r}^E > \nu_{2,r}^E$. Thus, this model ensures that the efficacy probabilities are heterogeneous across subgroups. When $\nu_{c,r} \rightarrow \mathbf{0}_2$ for all (c, r) , the model shrinks to the homogeneous case where regime effects in different subgroups are the same.

To define the priors of $\tilde{\xi}_r$, for schedule s and outcome $j = T, E$, we denote by $\tilde{\xi}_{-d,s}^j$ the subvector of $\tilde{\xi}_s^j = (\tilde{\xi}_{1,s}^j, \dots, \tilde{\xi}_{D,s}^j)$ with $\tilde{\xi}_{d,s}^j$ deleted, for $d = 1, \dots, D$. Our model includes the common assumption that the risk of toxicity increases monotonically with dose, which we formalize as

$$\tilde{\xi}_{d,s}^T | \tilde{\xi}_{-d,s}^T \sim N(\xi_0^T, \sigma_0^{T,2}) I(\tilde{\xi}_{d-1,s}^T < \tilde{\xi}_{d,s}^T < \tilde{\xi}_{d+1,s}^T), \quad (3)$$

where $N(\xi_0^T, \sigma_0^{T,2})$ denotes the hyperprior normal distribution with mean ξ_0^T and variance $\sigma_0^{T,2}$, with the truncated support of $\tilde{\xi}_{d,s}^T$ given by the indicator function. That is, the conditional distribution of the mean $\tilde{\xi}_{d,s}^T$ is restricted to the subset of the reals determined by the values of the other means, $\tilde{\xi}_{-d,s}^T$,

through the order constraint. These order constraints induce association among different dose levels, ensuring that the latent variable for toxicity increases stochastically in dose d for each schedule s , hence the probability of toxicity increases with d for each s . Such an order constraint can be achieved at each Markov chain Monte Carlo (MCMC) step by generating the proposals of $(\tilde{\xi}_{1,s}^T, \tilde{\xi}_{D,s}^T)$ from a multivariate normal distribution, subject to $\tilde{\xi}_{1,s}^T < \dots < \tilde{\xi}_{D,s}^T$. In contrast, we do not impose any monotonicity restriction on efficacy in d , and simply assume that

$$\tilde{\xi}_{d,s}^E | \tilde{\xi}_{-d,s}^E \sim N(\xi_0^E, \sigma_0^{E,2}), \quad (4)$$

where $N(\xi_0^E, \sigma_0^{E,2})$ denotes the unconstrained hyperprior normal distribution. Thus, for each s and c , the dose-efficacy probability relationship can take a wide variety of possible forms.

The priors on $\nu_{c,r}$ should be elicited while accounting for prior order. Based on the MM trial with $C = 2$ ordered subgroups, the toxicity distribution is homogeneous across subgroups, so we assume $\nu_{c,r}^T = 0$ for each (c, r) combination. As efficacy in the second subgroup ($c = 2$) is greater than in the first group ($c = 1$), we estimate $\nu_{2,r}^E$ by borrowing information across dose levels, as follows: $\nu_{2,r}^E = \nu_{2,d,s}^E \stackrel{\text{i.i.d.}}{\sim} N_+(\nu_{1,r}^E, \tau_0^2)$, $d = 1, \dots, D$, $s = 1, \dots, S$, where $N_+(\nu_{1,r}^E, \tau_0^2)$ is a truncated normal distribution with support $(\nu_{1,r}^E, \infty)$, and the variance τ_0^2 is prespecified. This ensures that, given r , the efficacy probability in subgroup 2 is strictly greater than that in subgroup 1. For more general trials, if there is an ordering relationship among subgroups in terms of toxicity, say, subgroup 2 has a higher toxicity probability than subgroup 1, the design may account for this by assuming the prior, $\nu_{2,r}^T = \nu_{2,d,s}^T \stackrel{\text{i.i.d.}}{\sim} N_+(\nu_{1,r}^T, \tau_0^2)$, $d = 1, \dots, D$, $s = 1, \dots, S$, with $\nu_{1,r}^T = 0$. However, as the MM trial considers a homogeneous toxicity distribution across subgroups, we simply take $\nu_{2,r}^T = \nu_{1,r}^T = 0$ in this paper.

To specify the likelihood and posterior, we denote $\tilde{\xi} = \{\tilde{\xi}_{c,r}, c = 1, \dots, C, r = 1, \dots, DS\}$. Combining Equations (1) and (2), the joint distribution of (ξ_i^T, ξ_i^E) can be derived by integrating out ϵ_i , yielding

$$(\xi_i^T, \xi_i^E) | c_i, r_i, \tilde{\xi} \stackrel{\text{indep}}{\sim} \text{BN}(\mu_{c_i, r_i}, \Sigma_\xi), \quad (5)$$

where the mean vector μ_{c_i, r_i} depends on the i th patient’s subgroup c_i and treatment regime $r_i = (d_i, s_i)$. More precisely, $\mu_{c_i, r_i} = (\tilde{\xi}_{c_i, r_i}^T, \tilde{\xi}_{c_i, r_i}^E)$, and Σ_ξ is the covariance matrix with the diagonal elements being $\sigma_\xi^2 + \zeta^2$ and the off-diagonal elements being $\rho\zeta^2$. As a result, the

individual likelihood for the observations of a patient with outcome (y_i^T, y_i^E) ($y_i^T, y_i^E = 0, 1$) can be parameterized as

$$\begin{aligned} f(y_i^T, y_i^E | c_i, r_i, \tilde{\xi}) &= \Pr(\gamma_{y_i^T} \leq \xi_i^T < \gamma_{y_i^T+1}, \gamma_{y_i^E} \leq \xi_i^E \\ &< \gamma_{y_i^E+1} | c_i, r_i, \tilde{\xi}) \\ &= \int_{\gamma_{y_i^T}}^{\gamma_{y_i^T+1}} \int_{\gamma_{y_i^E}}^{\gamma_{y_i^E+1}} f(\xi_i^T, \xi_i^E | c_i, r_i, \tilde{\xi}) d\xi_i^E d\xi_i^T, \end{aligned}$$

where $f(\xi^T, \xi^E | c, r, \tilde{\xi})$ is given by (5), and the cutoff vector $(\gamma_0, \gamma_1, \gamma_2) = (-\infty, 0, \infty)$.

2.2 | Delayed outcomes

In the MM study, the toxicity outcome is evaluated within $V^T = 30$ days, while efficacy is defined as complete remission based on disease evaluation on day $V^E = 90$. Thus, toxicity can occur at any time during the 30 day assessment window, but efficacy is not known until a patient has reached day 90 of follow-up. Consequently, when a regime must be assigned for a newly enrolled patient, the outcomes of some previously treated patients might not have been fully assessed. Formally, at the time of interim decision-making, both Y^E and Y^T of previously treated patients are subject to missingness. In the MM study, the amount of missing Y^T data would be much less than the amount of missing Y^E data because the 30-day assessment window for Y^T is much shorter than the 90 days required to evaluate Y^E .

At an interim decision-making time, suppose that patient i has been followed for t_i days. We introduce an indicator vector $\delta_i = (\delta_i^T, \delta_i^E)$ to denote the respective missingness of toxicity and efficacy for the i th patient, where $\delta_i^j = 1$ if Y_i^j has been evaluated and $\delta_i^j = 0$ if not, for $j = T, E$. In the MM trial setting, Y^T can be defined as a time-to-event outcome. Suppose X_i denotes the i th patient's time to toxicity, then $Y_i^T = 1$ if $X_i \leq V^T$ and $Y_i^T = 0$ if $X_i > V^T$. We have $\delta_i^T = I(t_i \geq \min(X_i, V^T))$ and $\delta_i^E = I(t_i \geq V^E)$. Because $V^T < V^E$, we have $\delta_i^T \geq \delta_i^E$, so $(\delta_i^T, \delta_i^E) = (0, 1)$ is impossible.

When both $Y_i^T = y_i^T$ and $Y_i^E = y_i^E$ are observed for patient i , that is, $t_i > V^E$, the individual likelihood is $f(y_i^T, y_i^E | \delta_i^T = 1, \delta_i^E = 1, c_i, r_i, \tilde{\xi}) = f(y_i^T, y_i^E | c_i, r_i, \tilde{\xi})$. Under the mechanism of missing at random, the likelihood of a patient with observed $Y_i^T = y_i^T$ and missing Y_i^E , that is, $\min(X_i, V^T) \leq t_i < V^E$ and $(\delta_i^T, \delta_i^E) = (1, 0)$, is $f(y_i^T | \delta_i^T = 1, \delta_i^E = 0, c_i, r_i, \tilde{\xi}) = \Pr(\gamma_{y_i^T} \leq \xi_i^T < \gamma_{y_i^T+1} | c_i, r_i, \tilde{\xi})$, where $\Pr(\gamma_{y_i^T} \leq \xi_i^T < \gamma_{y_i^T+1} | c_i, r_i, \tilde{\xi})$ is the marginal likelihood of Y_i^T evaluated at y_i^T . When both y_i^T and y_i^E are missing, that is, $t_i < \min(X_i, V^T)$ and $(\delta_i^T, \delta_i^E) = (0, 0)$, it only is known that the i th patient's

time to toxicity is greater than t_i . In this case, for $0 < t_i < V^T$, the likelihood is given by

$$\begin{aligned} &\Pr(X_i > t_i | \delta_i^T = 0, \delta_i^E = 0, c_i, r_i, \tilde{\xi}) \\ &= \Pr(Y_i^T = 0 | c_i, r_i, \tilde{\xi}) \Pr(X_i > t_i | Y_i^T = 0, c_i, r_i, \tilde{\xi}) \quad (6) \\ &+ \Pr(Y_i^T = 1 | c_i, r_i, \tilde{\xi}) \Pr(X_i > t_i | Y_i^T = 1, c_i, r_i, \tilde{\xi}) \\ &= 1 - w_i \Pr(Y_i^T = 1 | c_i, r_i, \tilde{\xi}), \quad (7) \end{aligned}$$

where we denote $w_i = \Pr(X_i \leq t_i | Y_i^T = 1, c_i, r_i, \tilde{\xi})$. The first equality above is due to the fact that, given $\delta_i^T = 0$ and $\delta_i^E = 0$, all values of (Y_i^T, Y_i^E) are possible. Suppressing $c_i, r_i, \tilde{\xi}$ for brevity, as $0 < t_i < V^T$, $\Pr(X_i > t_i | Y_i^T = 0) = \Pr(V^T \geq X_i > t_i | Y_i^T = 0) + \Pr(X_i > V^T | Y_i^T = 0) = 0 + 1 = 1$, hence, the first summand (6) equals $\Pr(Y_i^T = 0 | c_i, r_i, \tilde{\xi}) = 1 - \Pr(Y_i^T = 1 | c_i, r_i, \tilde{\xi})$.

We assume that, conditional on Y_i^T , the time-to-toxicity distribution is independent of $(c_i, r_i, \tilde{\xi})$, that is, X_i does not depend on subgroup or treatment regime given the indicator of toxicity on $[0, V^T]$. We thus need to estimate w_i in order to obtain a working likelihood for $\Pr(X_i > t_i | \delta_i^T = 0, \delta_i^E = 0, c_i, r_i, \tilde{\xi})$. Noting that the support of $[X_i | Y_i^T = 1]$ is $(0, V^T)$, we model the conditional samples of $[X_i | Y_i^T = 1]$ based on a scaled Beta distribution, given by

$$\begin{aligned} &\text{Sampling model: } X_i | Y_i^T = 1 \sim V^T \times \text{Beta}(\alpha, \beta), \quad (8) \\ &\text{Priors: } \alpha, \beta \sim \text{Gamma}(\lambda_0, \eta_0), \end{aligned}$$

where λ_0 and η_0 are the hyperparameters for the Gamma prior distribution. As a result, w_i can be obtained based on the posterior distribution of $[X_i | Y_i^T = 1]$.

Let $\mathcal{D}_n = \{(c_i, r_i, y_i^T, y_i^E, x_i, t_i, \delta_i^T, \delta_i^E)\}_{i=1}^n$ be the observed data at the arrival time of the $(n + 1)$ th patient. The joint likelihood for the first n patients can be written as

$$\begin{aligned} L(\mathcal{D}_n | \tilde{\xi}, \alpha, \beta) &= \prod_{i=1}^n \left\{ g(x_i | \alpha, \beta)^{y_i^T} \right. \\ &\times f(y_i^T, y_i^E | \delta_i^T = 1, \delta_i^E = 1, c_i, r_i, \tilde{\xi})^{\delta_i^T \delta_i^E} \\ &\times f(y_i^T | \delta_i^T = 1, \delta_i^E = 0, c_i, r_i, \tilde{\xi})^{\delta_i^T (1 - \delta_i^E)} \\ &\left. \times \Pr(X_i \geq t_i | \delta_i^T = 0, \delta_i^E = 0, c_i, r_i, \tilde{\xi})^{(1 - \delta_i^T)(1 - \delta_i^E)} \right\} \end{aligned}$$

where $g(\cdot | \alpha, \beta)$ is the density function of the scaled Beta distribution.

Denote the vector of all hyperparameters by $\theta_0 = (\rho, \zeta^2, \sigma_\xi^2, \lambda_0, \eta_0, \xi_0^T, \xi_0^E, \sigma_0^{T,2}, \sigma_0^{E,2}, \tau_0^2)$, and let $\pi(\tilde{\xi}, \alpha, \beta | \theta_0)$ be the joint prior distribution of $(\tilde{\xi}, \alpha, \beta)$ induced by the hierarchical model (1) to (4) and model (8) for time to toxicity. The joint posterior distribution of $(\tilde{\xi}, \alpha, \beta)$ is then given by

$\pi(\xi, \alpha, \beta | \theta_0, \mathcal{D}_n) \propto \pi(\xi, \alpha, \beta | \theta_0) L(\mathcal{D}_n | \xi, \alpha, \beta)$, where the posterior samples of $[\xi, \alpha, \beta | \theta_0, \mathcal{D}_n]$ can be obtained via standard MCMC sampling methods. The sampling procedure is carried out in two steps. Since the posterior sampling of model (8) only depends on the time to toxicity data $\{(X_i, \delta_i^T), i = 1, \dots, n\}$, in the first step, we simulate the posterior samples of (α, β) , as well as those of w_i , because w_i depends solely on the model assumption (8). In the second step, we plug the posterior samples of w_i into Equation (7) to obtain samples of the remaining parameters. The R code for implementing the proposed design is available in the Supporting Information.

3 | TRIAL DESIGN

We define the admissibility criteria to screen out any regimes with excessively high toxicity or unacceptably low efficacy adaptively based on the interim data. Let $\pi_{c,r}^j(\xi) = \Pr(Y^j = 1 | c, r, \xi)$ be the marginal probability of outcome $j = T, E$. Recall that θ_0 denotes the vector of all fixed hyperparameters. Given a fixed upper limit $\bar{\pi}^T$ on $\pi_{c,r}^T(\xi)$, a fixed lower limit $\underline{\pi}^E$ on $\pi_{c,r}^E(\xi)$, and fixed cutoff probabilities η^T and η^E , for each subgroup c , we define the set \mathcal{A}_c of admissible regimes to be all $r = (d, s)$ satisfying the two criteria

$$\Pr\{\pi_{c,r}^T(\xi) > \bar{\pi}^T | \theta_0, \mathcal{D}_n\} < \eta^T \quad \text{and} \quad \Pr\{\pi_{c,r}^E(\xi) < \underline{\pi}^E | \theta_0, \mathcal{D}_n\} < \eta^E, \tag{9}$$

similarly to Thall and Cook (2004).

To choose regimes from \mathcal{A}_c for each subgroup c , we utilize a utility-based criterion to quantify efficacy-toxicity risk-benefit trade-offs. To do this, a numerical utility $U(y^T, y^E)$ is elicited for each of the four elementary outcome pairs $(y^T, y^E) = (0, 0), (1, 0), (0, 1)$, and $(1, 1)$. For illustrations of the choice of U in a variety of settings, see Houede *et al.* (2010), Thall and Nguyen (2012), Yuan *et al.* (2016), or Liu *et al.* (2018). As (Y^T, Y^E) are random variables that depend on the patient’s regime r and subgroup c , $U(Y^T, Y^E)$ also is a random variable. Denote $\mathbf{y}_u = \{(y_T, y_E) : U(y_T, y_E) = u\}$. The posterior predictive distribution (PPD) of $U(Y^T, Y^E)$ for future values (Y^T, Y^E) is derived as follows:

$$\begin{aligned} & \Pr\{U(Y^T, Y^E) = u | c, r, \theta_0, \mathcal{D}_n\} \\ &= \sum_{\mathbf{y}_u} \Pr\{(Y^T, Y^E) = \mathbf{y}_u | c, r, \theta_0, \mathcal{D}_n\} \\ &= \sum_{\mathbf{y}_u} \int_{\xi} \Pr\{(Y^T, Y^E) = \mathbf{y}_u | c, r, \xi\} \pi(\xi | \theta_0, \mathcal{D}_n) d\xi, \end{aligned}$$

where $\pi(\xi | \theta_0, \mathcal{D}_n)$ is the marginal posterior of ξ obtained by integrating $\pi(\xi, \alpha, \beta | \theta_0, \mathcal{D}_n)$ over (α, β) under the gamma hyperprior. We denote the random variable $[U(Y^T, Y^E) | c, r, \theta_0, \mathcal{D}_n]$ by $U_{c,r} = U_{c,d,s}$. Let $u_c^{\max} = \max_{\{c,r\}} \{u_{c,r}\}$ be the maximum utility among all considered treatment regimes for subgroup $c = 1, \dots, C$, where $u_{c,r}$ denotes the true mean utility for combination (c, r) . The optimal treatment regime (OTR) for each subgroup $c = 1, \dots, C$, is defined as the regime with $u_{c,r} > u_c^{\max} - \underline{u}$, where \underline{u} is an indifference margin. In the MM study, we consider $\underline{u} = 5$.

The AR procedure of the proposed trial design, which will be given in detail below, depends on the PPD of $U_{c,r}$. We define AR with the probability of assignment to regime (d, s) within each subgroup c proportional to

$$\omega_c(d, s) = \Pr\{U_{c,d,s} = \max_{d' \in \{1, \dots, D\}} U_{c,d',s}\} I\{(d, s) \in \mathcal{A}_c\}. \tag{10}$$

In Equation (10), as $U_{c,d,s}$ is a random variable, the quantity $\max_{d' \in \{1, \dots, D\}} U_{c,d',s}$ is the maximum among D random variables, and may not equal the maximum utility $U(0, 1)$. This equation implies that the AR probability is proportional to the posterior predictive probability of attaining the maximum predicted utility among all admissible dose levels within treatment schedule s , for a future patient. Thus, $\omega_c(d, s)$ accounts for both the mean and variation of the utilities from different regimes within the schedule. This PPD-based approach is fundamentally different from the procedures used by Thall and Nguyen (2012), Lee *et al.* (2015), and others, where AR probabilities are defined in terms of posterior mean utilities. In the present context, these would be

$$\begin{aligned} u_c(d, s, \theta_0, \mathcal{D}_n) &= E[E\{U(Y^T, Y^E) | c, d, s, \xi\} | \theta_0, \mathcal{D}_n] \\ &= \sum_{(y^T, y^E)} U(y^T, y^E) \int_{\xi} f(Y^T = y^T, Y^E = y^E | c, d, s, \xi) \pi(\xi | \theta_0, \mathcal{D}_n) d\xi. \end{aligned}$$

AR probabilities $\omega'_c(d, s)$ among (d, s) pairs for subgroup c then are defined to be proportional to $u_c(d, s, \theta_0, \mathcal{D}_n) I\{(d, s) \in \mathcal{A}_c\}$.

Compared to the approach of defining AR probabilities $\omega'_c(d, s)$ based on posterior mean utilities (Thall and Nguyen, 2012; Lee *et al.*, 2015), the proposed approach of using the PPD of $U_{c,r}$ to define the AR probabilities $\omega_c(d, s)$ leads to a more extensive exploration of the regime space, and thus it addresses the “exploitation versus exploration” problem, which is well

known in phase I-II trials (Yuan *et al.*, 2016, chapter 2.6) and more generally in sequential analysis (Sutton and Barto, 1998). This is because the PPD of $U_{c,r}$ accounts for distributions of future observations. The AR procedure based on the PPD of $U_{c,r}$ thus tends to have a smaller chance of getting stuck at suboptimal regimes. As AR treats patients with suboptimal regimes, care must be taken to ensure that its use does not expose patients to unacceptably high risks of high toxicity or low efficacy.

To estimate the optimal subgroup-specific treatment regime that maximizes $U(y^T, y^E)$ across different combinations of (c, r) , we divide the trial into two stages. In stage 1, for each subgroup $c = 1, \dots, C$, N_{1c} patients are randomized fairly among the schedules. This is different from standard phase I methods, such as the CRM or EWOC, which use deterministic allocation to treat patients. In the motivating MM study, the toxicity assessment window is 30 days, while efficacy is evaluated much later, at day 90. Thus, toxicity outcomes are observed much sooner than efficacy outcomes, and in stage 1, the data available for making the adaptive decisions are largely toxicity data, with efficacy data collected to facilitate decision-making in stage 2. At the end of stage 1, as some previously missing Y^E outcomes may be observed for patients followed to V^E , preliminary estimates of the subgroup-specific optimal regimes can be obtained. For each subgroup c , stage 2 enrolls the remaining N_{2c} patients with the goal of estimating the globally optimal regime. To achieve this, we propose a procedure that does the optimization within schedule, combined with AR across schedules. An optimal dose first is selected within each treatment schedule, and then patients are adaptively randomized among the optimal dose set across schedules. This hybrid approach, of selecting optimal doses and randomizing, balances exploitation versus exploration by allowing sufficient dose exploration (through AR) to reduce the risk of being stuck at suboptimal doses, but also avoids allocating too many patients to suboptimal doses.

Let N^{\max} be the maximum total sample size, and p_c the prevalence of subgroup $c = 1, \dots, C$, so $\sum_{c=1}^C p_c = 1$. We bind the maximum sample size N_c^{\max} for subgroup c by $p_c N^{\max}$. Assume patients are recruited sequentially to each schedule within each subgroup. Let κ be the proportion of patients assigned to each schedule in stage 1, that is, we randomize κN_c^{\max} patients to each schedule in stage 1 for each subgroup. Thus, $N_{1c} = \kappa S N_c^{\max} = \kappa S p_c N^{\max}$ and $N_{2c} = N_c^{\max} - N_{1c}$. The two-stage trial proceeds as follows:

Stage 1. If the next patient enrolled is in subgroup c ,

1.1. Randomly choose a schedule, s , with probability $1/S$ each.

1.2. If s has never been tested before, then start the subtrial in this schedule at the lowest dose. Otherwise, based on (9), determine the admissible set \mathcal{A}_c in subgroup c based on the most recent data \mathcal{D}_n . Subject to the constraint that no untried dose may be skipped when escalating, randomly choose an acceptable dose for the next patient with AR probability proportional to $\omega_c(d, s)$, $d = 1, \dots, D$. Thus, the AR probability in subgroup c is proportional to the probability that regime (d, s) induces the maximum utility within schedule s , with all (d, s) that are unacceptable in subgroup c given AR probability 0.

1.3. The subtrial for subgroup c is either stopped when the maximum sample size κN_c^{\max} is reached, or terminated early if no dose within this schedule is admissible for subgroup c .

Stage 2. For each newly enrolled patient in subgroup c , first determine the optimal dose $d_c^*(s)$ that has the largest probability of having the maximum utility within each s ; that is, $d_c^*(s) = \arg \max_{d \in \{1, \dots, D\}} \{\omega_c(d, s)\}$, $s = 1, \dots, S$, where $\omega_c(d, s)$ is given by Equation (10). Then choose the schedule s across schedules with the AR probability proportional to

$$\omega_c(d_c^*(s), s) = \Pr \left\{ U_{c,d_c^*(s),s} = \max_{s' \in \{1, \dots, S\}} U_{c,d_c^*(s'),s'} \right\}$$

$$I\{d_c^*(s), s \in \mathcal{A}_c\}, \quad s = 1, \dots, S,$$

and assign the new patient dose $d_c^*(s)$. In other words, in stage 2, the dose first is optimized within each schedule without use of AR, and then AR is applied to randomize patients among doses across all schedules. Repeat this until N_{2c} patients have been treated in the second stage, and then stop the trial for subgroup c . If no regime is admissible for subgroup c as given by (9), then stop the trial in that subgroup.

At the end of the study, based on the complete data $\mathcal{D}_{N^{\max}}$, for each subgroup $c = 1, \dots, C$, the OTR is defined as that with the largest probability of having the maximum utility among all regimes, formally $r_c^* = (d_c^*, s_c^*) = \arg \max_{(d,s) \in \mathcal{A}_c} \Pr\{U_{c,d,s} = \max_{(d',s')} U_{c,d',s'}\}$.

To implement the design in practice, one must prespecify values of the hyperparameters θ_0 , the utility function $U(y^T, y^E)$, and the design parameters $(N^{\max}, \kappa, \bar{\pi}^T, \underline{\mu}^E, \eta^T, \eta^E)$. A detailed description of the calibration procedure for the proposed design is provided in the Supporting Information.

4 | SIMULATION STUDY

In this section, we summarize results of a simulation study to investigate the proposed design’s operating characteristics (OCs), using the PGF Melphalan trial as a basis for the simulation study design. We consider $C = 2$ subgroups with equal prevalences $p_1 = p_2 = 1/2$, and assume the toxicity probabilities $\pi_{c,r}^T$ are homogeneous across subgroups, but that the efficacy probabilities satisfy $\pi_{2,r}^E > \pi_{1,r}^E$ for all $r = (d, s)$. We will evaluate the sensitivity of the design’s performance to different prevalences. We study $D = 3$ doses (200, 225, and 250 mg m⁻²) combined with $S = 3$ infusion schedules, for a total of nine treatment regimes, and 18 different subgroup-specific dose-schedule regime combinations. We assume $N^{\max} = 120$ patients are accrued, so on average 6.6 patients are allocated to each subgroup-

specific dose-schedule regime. This sample size is reasonable, as the maximum sample size using a 3 + 3 design to find an MTD for each of six (c, s) pairs would be similar. Based on preliminary simulations, we take $\kappa = 0.2$, leading to $N_{11} + N_{12} = 72$ patients treated in stage 1. When $p_1 = p_2 = 1/2$, the stage 1 sample size per subgroup is $N_{11} = N_{12} = 36$. Toxicity is monitored during the first $V^T = 30$ days, and efficacy is evaluated on day $V^E = 90$. We assume patients are accrued at a rate of 6 per month, arriving according to a Poisson process, so the average interarrival time is 5 days. Thus, the expected time to accrue 120 patients is 20 months.

We evaluated the proposed design’s OCs under eight different scenarios, characterized in terms of fixed marginal probabilities of toxicity and efficacy ($\pi_{c,r}^T, \pi_{c,r}^E$) given in Table 1. The trial data were simulated using (1) and (2), where we set $\sigma_{\xi}^{2,\text{true}} = 0.5^2, \rho^{\text{true}} = -0.2$, and $\zeta^{2,\text{true}} = 0.3^2$.

TABLE 1 True toxicity and efficacy probabilities and utilities ($\pi_{c,r}^T, \pi_{c,r}^E, u_{c,r}$) under eight simulation scenarios, for each dose, schedule, and subgroup

Scenario	s	Subgroup 1			Subgroup 2		
		d = 1	d = 2	d = 3	d = 1	d = 2	d = 3
1	1	(0.03,0.10,62.2)	(0.05,0.20,65.0)	(0.15,0.60,75.0)	(0.03,0.13,63.4)	(0.05,0.23,66.2)	(0.15,0.63,76.2)
	2	(0.05,0.50,77.0)	(0.15,0.40,67.0)	(0.30,0.40,58.0)	(0.05,0.56,79.4)	(0.15,0.46,69.4)	(0.30,0.46,60.4)
	3	(0.13,0.35,66.2)	(0.45,0.40,49.0)	(0.60,0.45,42.0)	(0.13,0.47,71.0)	(0.45,0.52,53.8)	(0.60,0.57,46.8)
2	1	(0.10,0.30,66.0)	(0.27,0.40,59.8)	(0.55,0.50,47.0)	(0.10,0.40,70.0)	(0.27,0.50,63.8)	(0.55,0.60,51.0)
	2	(0.25,0.25,55.0)	(0.30,0.30,54.0)	(0.40,0.40,52.0)	(0.25,0.35,59.0)	(0.30,0.40,58.0)	(0.40,0.50,56.0)
	3	(0.08,0.15,61.2)	(0.12,0.35,66.8)	(0.25,0.35,59.0)	(0.08,0.25,65.2)	(0.12,0.45,70.8)	(0.25,0.45,63.0)
3	1	(0.05,0.10,61.0)	(0.15,0.40,67.0)	(0.40,0.10,40.0)	(0.05,0.30,69.0)	(0.15,0.45,69.0)	(0.40,0.30,48.0)
	2	(0.05,0.40,73.0)	(0.18,0.20,57.2)	(0.40,0.10,40.0)	(0.05,0.45,75.0)	(0.18,0.30,61.2)	(0.40,0.20,44.0)
	3	(0.03,0.15,64.2)	(0.08,0.23,64.4)	(0.15,0.45,69.0)	(0.03,0.30,70.2)	(0.08,0.30,67.2)	(0.15,0.50,71.0)
4	1	(0.03,0.10,62.2)	(0.05,0.30,69.0)	(0.10,0.60,78.0)	(0.03,0.20,66.2)	(0.05,0.35,71.0)	(0.10,0.63,79.2)
	2	(0.07,0.30,67.8)	(0.15,0.40,67.0)	(0.30,0.50,62.0)	(0.07,0.33,69.0)	(0.15,0.60,75.0)	(0.30,0.65,68.0)
	3	(0.05,0.25,67.0)	(0.10,0.34,67.6)	(0.15,0.25,61.0)	(0.05,0.28,68.2)	(0.10,0.38,69.2)	(0.15,0.29,62.6)
5	1	(0.05,0.10,61.0)	(0.12,0.25,62.8)	(0.20,0.33,61.2)	(0.05,0.15,63.0)	(0.12,0.30,64.8)	(0.20,0.38,63.2)
	2	(0.07,0.05,57.8)	(0.13,0.45,70.2)	(0.25,0.30,57.0)	(0.07,0.20,63.8)	(0.13,0.50,72.2)	(0.25,0.40,61.0)
	3	(0.02,0.23,68.0)	(0.05,0.15,63.0)	(0.08,0.10,59.2)	(0.02,0.28,70.0)	(0.05,0.45,75.0)	(0.08,0.28,66.4)
6	1	(0.05,0.05,59.0)	(0.07,0.07,58.6)	(0.09,0.09,58.2)	(0.05,0.15,63.0)	(0.07,0.47,74.6)	(0.09,0.49,74.2)
	2	(0.08,0.10,59.0)	(0.13,0.35,66.2)	(0.30,0.40,58.0)	(0.08,0.15,61.2)	(0.13,0.40,68.2)	(0.30,0.45,60.0)
	3	(0.11,0.30,65.4)	(0.13,0.20,60.2)	(0.20,0.10,52.0)	(0.11,0.33,66.6)	(0.13,0.23,61.4)	(0.20,0.13,53.2)
7	1	(0.05,0.10,61.0)	(0.15,0.45,69.0)	(0.30,0.45,60.0)	(0.05,0.13,62.2)	(0.15,0.48,70.2)	(0.30,0.48,61.2)
	2	(0.12,0.20,60.8)	(0.23,0.55,68.2)	(0.55,0.60,51.0)	(0.12,0.50,72.8)	(0.23,0.58,69.4)	(0.55,0.63,52.2)
	3	(0.45,0.30,45.0)	(0.50,0.35,44.0)	(0.60,0.30,36.0)	(0.45,0.33,46.2)	(0.50,0.38,45.2)	(0.60,0.31,36.4)
8	1	(0.05,0.10,61.0)	(0.10,0.20,62.0)	(0.15,0.50,71.0)	(0.05,0.15,63.0)	(0.10,0.25,64.0)	(0.15,0.53,72.2)
	2	(0.10,0.15,60.0)	(0.25,0.15,51.0)	(0.40,0.15,42.0)	(0.10,0.20,62.0)	(0.25,0.20,53.0)	(0.40,0.20,44.0)
	3	(0.07,0.25,65.8)	(0.08,0.25,65.2)	(0.18,0.25,59.2)	(0.07,0.25,65.8)	(0.08,0.30,67.2)	(0.18,0.60,73.2)

Note: These values for optimal treatment regimes with $u_{c,r} > u_c^{\max} - 5$ are bold, where $u_c^{\max} = \max_r\{u_{c,r}\}$ is the maximum utility for subgroup $c, c = 1, 2$. The regimes with a toxicity rate above 15% and an efficacy rate below 20% are considered inadmissible in scenarios 1 to 6. The regimes with a toxicity rate above 30% and an efficacy rate below 30% are considered inadmissible in scenarios 7 and 8.

The true values of $\xi_{c,r}^{\text{true}}$ were determined by matching $\pi_{c,r}^j = 1 - \Phi(0 | \xi_{c,r}^{j,\text{true}}, \sigma_{\xi}^{2,\text{true}} + \zeta^{2,\text{true}}), j = T, E$. In scenarios 1 to 6, a regime with an efficacy probability $< \bar{\pi}^E = 0.20$ is considered clinically unimportant, and a toxicity probability $> \bar{\pi}^T = 0.15$ is considered unsafe. To assess the applicability of the design to more general scenarios, regimes with a toxicity rate above 30% and an efficacy rate below 30% are considered inadmissible in scenarios 7 and 8, which are different from the MM trial. The utility function is $U(0, 1) = 100, U(0, 0) = 60, U(1, 1) = 40, U(1, 0) = 0$, reflecting the belief that avoiding toxicity is more important than achieving efficacy. The expected utility of each regime under the eight scenarios is displayed in Table 1, and the true OTRs are bold.

We denote the proposed two-stage design by TD. The design configuration of TD is given in the Supporting Information. To show the advantage of borrowing information between subgroups, we compare the TD with a design that conducts a separate trial independently for each subgroup in parallel, with the sample size for subgroup c bounded by $p_c N^{\text{max}}, c = 1, \dots, C$. We denote this design by ITD. As a benchmark for comparison, we also implement the complete-data version of the proposed two-stage design, denoted by TD_C , which waits until all toxicity and efficacy outcomes of previously treated patients are completely observed before choosing a regime for the next patient. The TD_C design thus requires repeatedly suspending accrual of new patients prior to each new regime assignment. Therefore, TD_C has a very lengthy trial duration, which is not feasible in practice with late-onset outcomes. To examine the benefit

of using the proposed AR probabilities $\omega_c(d, s)$, we also include the design using AR probabilities $\omega'_c(d, s)$ that depends on posterior mean utilities. We denote this design by TD_U . Each design was simulated 1000 times under each scenario.

Table 2 shows the percentages of selecting OTRs and the average trial durations of the four designs. In general, the average OTR selection percentages of TD are 71.8 and 74.9 for subgroups 1 and 2, respectively. As scenarios 7 and 8 have different definitions of admissible regimes than scenarios 1 to 6, the desirable performance of TD in scenarios 7 and 8 also indicates that the proposed TD design is flexible and can be applied to different trial settings. The selection percentage of each regime using TD are given in Table 3. We find that TD is efficient in identifying inadmissible regimes. For example, TD has very small probabilities of selecting toxic treatment regimes with schedule $s = 2$ in scenario 2. Similarly, in scenario 8, where the efficacy probabilities of treatment regimes $r = (d, 2), d = 1, 2, 3$, all are below the lower limit $\bar{\pi}^E = 0.30$, TD is unlikely to select these inefficacious regimes as the OTRs.

Table 2 shows that the OTR selection percentages of TD are very close to those of TD_C , indicating that TD recovers from efficiency loss due to missing efficacy data early in the trial. Once the missing outcomes are observed, TD efficiently incorporates the new data for subsequent decision-making. Since TD_C repeatedly suspends accrual of new patients to wait for full assessments of previously treated patients, on average it would require 360 months to complete a trial with 120 patients. In

TABLE 2 Selection percentage for the optimal treatment regime (OTR) within each subgroup, and mean trial durations, for the four designs under the eight scenarios in Table 1

Scenario	Selection percentage of OTR								Trial duration (in months)			
	Subgroup 1				Subgroup 2				TD	ITD	TD_C	TD_U
	TD	ITD	TD_C	TD_U	TD	ITD	TD_C	TD_U				
1	88.1	75.4	89.4	88.6	85.0	69.4	89.5	88.8	23.0	26.0	360.0	23.0
2	69.2	58.7	72.4	72.1	67.1	56.2	69.0	70.9	23.1	26.0	360.0	23.0
3	73.7	64.0	74.1	74.7	73.5	62.2	72.6	75.9	23.0	25.9	360.0	23.0
4	60.9	46.1	63.4	47.9	78.4	66.0	79.7	78.0	23.1	26.0	360.0	23.0
5	72.9	70.0	78.9	75.8	85.9	79.0	87.5	86.9	23.0	26.0	360.0	23.0
6	79.8	74.4	78.0	72.9	48.0	59.8	50.8	45.0	23.0	26.0	360.0	23.0
7	82.5	80.9	83.9	72.4	95.6	93.1	96.7	95.3	23.0	26.0	360.0	23.0
8	47.2	35.2	50.9	38.1	65.6	58.6	68.0	54.0	23.0	26.0	360.0	23.0
Average	71.8	63.1	73.9	67.8	74.9	68.0	75.6	74.4	23.0	26.0	360.0	23.0

Note: "TD" is the proposed two-stage trial design; " TD_C " denotes the two-stage trial design based on complete (Y^T, Y^E) data; " TD_U " denotes the two-stage trial design based on AR probabilities $\omega'_c(d, s)$; "ITD" denotes the independent two-stage design that conducts a separate regime-finding trial for each subgroup. The accrual rate is six patients per month

TABLE 3 Regime selection percentage based on the proposed two-stage design under the eight scenarios in Table 1

Scenario	Schedule/dose	Subgroup 1			Subgroup 2		
		1	2	3	1	2	3
1	1	0.0	0.3	38.7	0.1	0.7	33.6
	2	49.4	6.7	0.8	51.4	6.1	0.9
	3	4.0	0.1	0.0	7.1	0.1	0.0
2	1	31.7	11.8	0.3	30.0	13.1	0.2
	2	3.8	0.8	0.6	4.8	1.4	0.2
	3	9.1	37.5	4.4	8.3	37.1	4.9
3	1	1.0	17.3	0.0	3.5	16.6	0.0
	2	56.3	0.7	0.0	49.4	1.4	0.0
	3	3.1	4.2	17.4	5.2	5.0	18.9
4	1	0.1	6.3	60.9	0.9	5.8	56.7
	2	5.8	13.1	3.2	3.9	21.7	3.9
	3	2.7	7.4	0.5	1.8	5.1	0.2
5	1	1.1	8.7	3.1	1.2	5.4	3.2
	2	0.6	57.9	1.6	1.1	50.3	1.8
	3	15.0	11.1	0.9	11.0	24.6	1.4
6	1	0.9	5.2	3.2	0.8	30.6	17.4
	2	2.4	37.1	3.5	1.1	22.6	2.7
	3	42.7	5.0	0.0	22.7	2.1	0.0
7	1	1.1	45.4	4.2	1.0	35.4	2.7
	2	11.2	37.1	0.6	28.9	31.3	0.3
	3	0.4	0.0	0.0	0.4	0.0	0.0
8	1	2.4	7.2	47.2	1.9	5.3	42.3
	2	2.7	0.2	0.0	2.6	0.3	0.0
	3	21.5	11.7	7.1	13.9	10.4	23.3

Note: The accrual rate is six patients per month. These values for optimal treatment regimes with $u_{c,r} > u_c^{\max} - 5$ are bold, where $u_c^{\max} = \max_r\{u_{c,r}\}$ is the maximum utility for subgroup c , $c = 1, 2$.

contrast, TD facilitates real-time decision-making with no suspension of accrual, requiring approximately 23 months for the trial, with a negligible drop in OTR selection percentage. Comparing the OTR selection percentages of TD and TD_U in Table 2 shows that, on average, TD performs better than TD_U . When there is only one OTR, as in scenarios 4 and 8 for subgroup 1, TD yields approximately 10% higher OTR selection percentages than TD_U , indicating that the use of AR probabilities by TD leads to a more thorough exploration of the treatment regime space. We summarize the total number of patients treated with a toxic regime having true $\pi_{c,r}^T > \bar{\pi}^T$ in Figure 1, and the total number of patients treated with an inefficacious regime having $\pi_{c,r}^E < \bar{\pi}^E$ in Figure 2. The results show that, compared to TD_U , TD only exposes two and three more patients to overly toxic treatment regimes. For a maximum sample size of 120

patients, such a risk is generally acceptable. However, because it explores more doses and schedules, TD tends to treat fewer patients at inefficacious regimes than TD_U (see Figure 2).

Table 2 shows the advantage of borrowing information across subgroups, in terms of within-subgroup OTR selection percentage. For nearly all scenarios and subgroups, TD has a larger OTR selection percentage than ITD, with the relative performance between TD and ITD depending on the degree of homogeneity of treatment effects across subgroups. In scenarios 1 and 2, where the locations of the OTRs are the same for the two subgroups, TD greatly outperforms ITD in terms of selection percentages of OTRs, with at least a 10% advantage over ITD for all scenario-subgroup combinations. This is because TD borrows information between subgroups. When the subgroups are relatively homogeneous in terms of treatment effects, TD may be expected to perform better than ITD. In scenarios 3 and 5, subgroup 2 has one more OTR than subgroup 1, with the remaining OTRs of subgroup 2 the same as those of subgroup 1. In these scenarios, TD again has substantially larger OTR selection percentages than ITD. However, in extremely heterogenous cases, borrowing information may harm TD's performance. This is shown by scenario 6, where the OTRs are very different for the two subgroups, and the OTR selection percentage in subgroup 2 for TD is less than that of ITD. An advantage of information sharing by the TD method is that, as the toxicity outcomes are assumed to be homogenous, borrowing toxicity information across subgroups helps screen out overly toxic regimes. As ITD does not borrow information between subgroups, it is more likely to treat patients with overly toxic regimes, illustrated by Figure 1, which gives the total numbers of patients overdosed. Thus, in terms of OTR selection, trial duration, and safety, TD is superior to ITD.

The Supporting Information reports extensive sensitivity analyses to examine the OCs of the proposed TD for different maximum sample sizes, N^{\max} . These show that the probability that TD correctly identifies the OTR increases substantially with N^{\max} . For $|\mathcal{D}| \leq \lfloor |V(G')|/2 \rfloor + 1 \leq \lfloor n/2 \rfloor$, the average selection percentage of OTR across the eight considered scenarios is as high as 90%. This indicates that the proposed design can recover from situations where it may get stuck early on at suboptimal regimes. We also show that TD is very robust to various subgroup prevalence ratios, patient accrual rates, true underlying models (data-generating processes), and prior distributions (with reasonably noninformative priors).

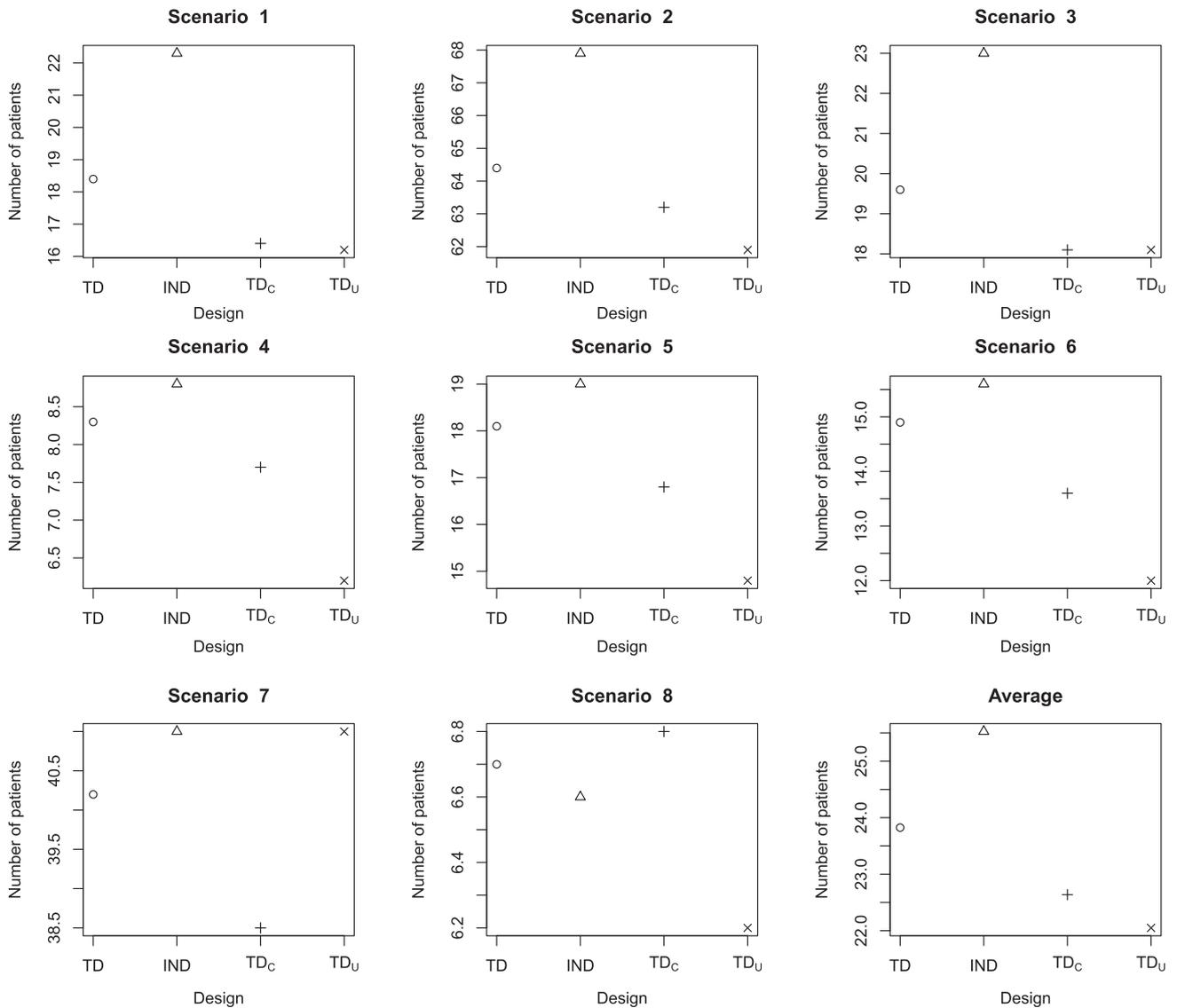


FIGURE 1 Average total number of patients overdosed, that is, treated with a regime having true $\pi_{c,r}^T > \bar{\pi}^T = .15$ in scenarios 1 to 6 and $\pi_{c,r}^T > \bar{\pi}^T = .30$ in scenarios 7 and 8, for the TD (circle ○), ITD (triangle △), TD_c (plus +), and TD_u (cross ×) under the simulation scenarios in Table 1. “TD” is the proposed two-stage trial design; “TD_c” denotes the two-stage trial design based on complete (Y^T, Y^E) data; “TD_u” denotes the two-stage trial design based on AR probabilities $\omega'_c(d, s)$; “ITD” denotes the independent two-stage design that conducts a separate regime-finding trial for each subgroup

5 | CONCLUDING REMARKS

We have proposed a two-stage phase I-II clinical trial design that does subgroup-specific dose-schedule finding based on a Bayesian hierarchical model, with specific attention to settings where the efficacy outcome is evaluated long after the start of treatment. To accommodate subgroups, the model exploits prior ordering information that the drug should be more effective in one subgroup than the other. The PPD of the utility of each dose-schedule regime is used as a

basis for regime selection and AR, which is employed to improve reliability. Within-subgroup regime acceptability rules are included for both toxicity and efficacy.

Late-onset outcomes complicate outcome-adaptive trial conduct. We have addressed this problem by using a hybrid two-stage design with AR. In stage 1, little efficacy data are available, and mainly toxicity data are utilized for early decision-making, primarily to screen out unsafe treatment regimes. When efficacy outcomes of more patients have been

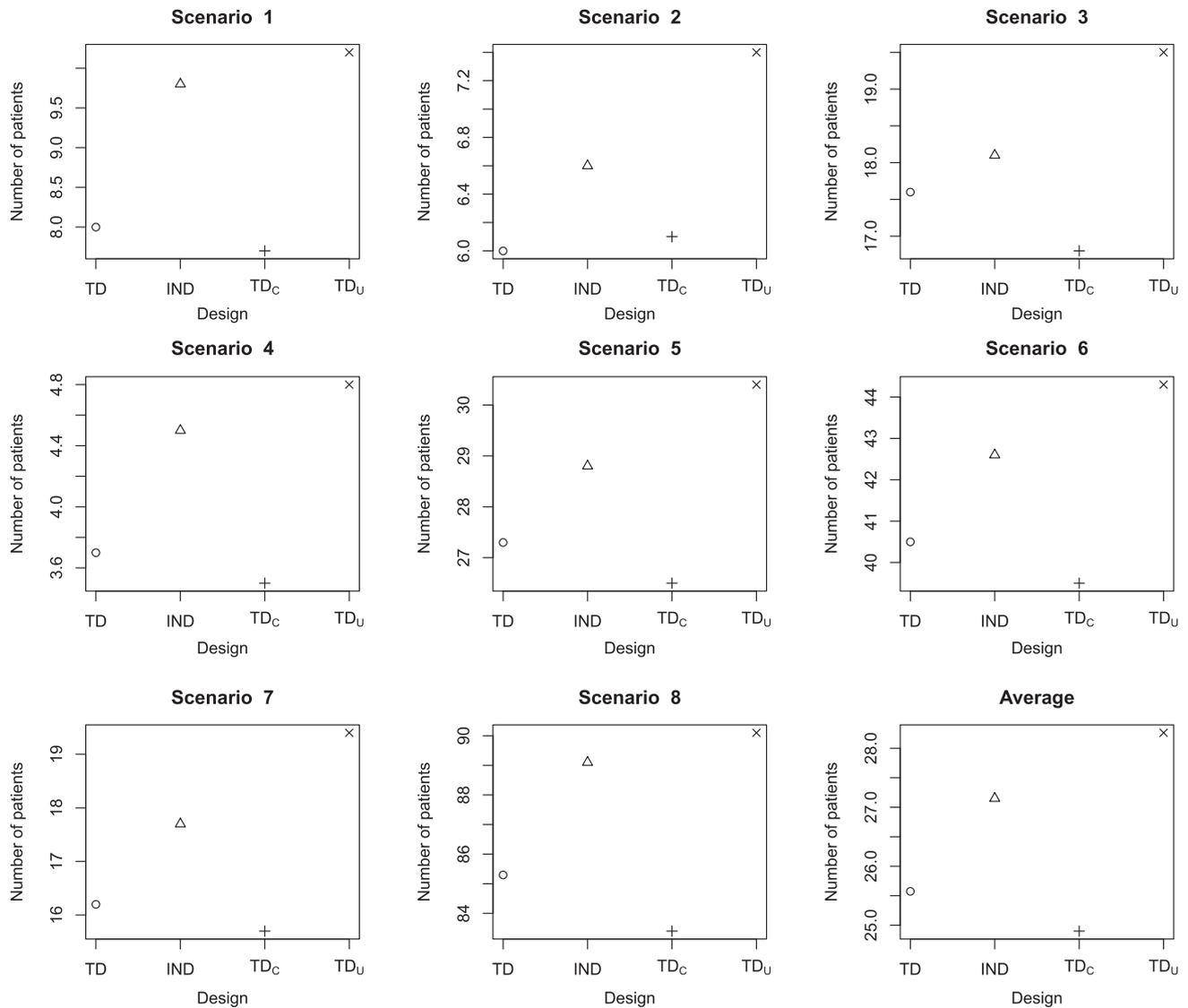


FIGURE 2 Average total number of patients treated with an inefficient regime having true $\pi_{c,r}^E < \underline{\pi}^E = .20$ in scenarios 1 to 6 and $\pi_{c,r}^E > \underline{\pi}^E = .30$ in scenarios 7 and 8, for the TD (circle ○), ITD (triangle △), TD_c (plus +), and TD_u (cross ×) under the simulation scenarios in Table 1. “TD” is the proposed two-stage trial design; “TD_c” denotes the two-stage trial design based on complete (Y^T , Y^E) data; “TD_u” denotes the two-stage trial design based on AR probabilities $\omega'_c(d, s)$; “ITD” denotes the independent two-stage design that conducts a separate regime-finding trial for each subgroup

assessed in stage 2, efficacy plays a more prominent role in choosing regimes for the remaining patients, and for making the final within-subgroup optimal regime selections. Simulations show that the operating characteristics of the proposed design are very similar to those of the benchmark complete-data design, which would require an unrealistically long time to complete the trial. Thus, the proposed design has a minimal loss in efficiency due to accommodating late-onset toxicity/efficacy, while providing a realistic trial duration.

ACKNOWLEDGMENTS

We thank the associate editor, two referees, and the editor for many constructive and insightful comments that led to significant improvements in the article. YY was partially supported by NIH grants 5P50CA098258, 1P50CA217685, and CA016672.

ORCID

Ruitao Lin  <http://orcid.org/0000-0003-2244-131X>
Ying Yuan  <http://orcid.org/0000-0003-3163-480X>

REFERENCES

- Albert, J.H. and Chib, S. (1993) Bayesian analysis of binary and polychotomous response data. *Journal of the American Statistical Association*, 88, 669–679.
- Babb, J., Rogatko, A. and Zacks, S. (1998) Cancer phase I clinical trials: efficient dose escalation with overdose control. *Statistics in Medicine*, 17, 1103–1120.
- Braun, T.M., Thall, P.F., Nguyen, H. and DeLima, M. (2007) Simultaneously optimizing dose and schedule of a new cytotoxic agent. *Clinical Trials*, 4, 113–124.
- Braun, T.M., Yuan, Z. and Thall, P.F. (2005) Determining a maximum tolerated schedule of a cytotoxic agent. *Biometrics*, 61, 335–343.
- Cheung, Y.K. and Chappell, R. (2000) Sequential designs for phase I clinical trials with late-onset toxicities. *Biometrics*, 56, 1177–1182.
- Chng, W.J., Ketterling, R.P. and Fonseca, R. (2006) Analysis of genetic abnormalities provides insights into genetic evolution of hyperdiploid myeloma. *Genes, Chromosomes and Cancer*, 45, 1111–1120.
- Fonseca, R., Bergsagel, P.L., Drach, J., Shaughnessy, J., Gutierrez, N., Stewart, A.K. *et al.* (2009) International Myeloma Working Group molecular classification of multiple myeloma: spotlight review. *Leukemia*, 23, 2210–2221.
- Guo, B., Li, Y. and Yuan, Y. (2016) A dose-schedule finding design for phase I–II clinical trials. *Journal of the Royal Statistical Society, Series C*, 65, 259–272.
- Guo, B. and Yuan, Y. (2017) Bayesian phase I/II biomarker-based dose finding for precision medicine with molecularly targeted agents. *Journal of the American Statistical Association*, 112, 508–520.
- Houede, N., Thall, P.F., Nguyen, H., Paoletti, X. and Kramar, A. (2010) Utility-based optimization of combination therapy using ordinal toxicity and efficacy in phase I–II trials. *Biometrics*, 66, 532–540.
- Liu, S., Guo, B. and Yuan, Y. (2018) A Bayesian phase I/II design for immunotherapy trials. *Journal of the American Statistical Association*, 113, 1016–1027.
- Liu, S. and Yuan, Y. (2015) Bayesian optimal interval designs for phase I clinical trials. *Journal of the Royal Statistical Society, Series C*, 64, 507–523.
- Lee, J., Thall, P.F., Ji, Y. and Müller, P. (2015) Bayesian dose-finding in two treatment cycles based on the joint utility of efficacy and toxicity. *Journal of the American Statistical Association*, 110, 711–722.
- Nebiyu Bekele, B. and Shen, Y. (2005) A Bayesian approach to jointly modeling toxicity and biomarker expression in a phase I–II dose-finding trial. *Biometrics*, 61, 343–354.
- 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.
- Storer, B.E. (1989) Design and analysis of phase I clinical trials. *Biometrics*, 45, 925–937.
- Sutton, R.S. and Barto, A.G. (1998) *Reinforcement Learning: An Introduction*. Cambridge, MA: MIT Press.
- Thall, P.F. and Cook, J.D. (2004) Dose-finding based on efficacy–toxicity trade-offs. *Biometrics*, 60, 684–693.
- Thall, P.F. and Nguyen, H.Q. (2012) Adaptive randomization to improve utility-based dose-finding with bivariate ordinal outcomes. *Journal of Biopharmaceutical Statistics*, 22, 785–801.
- Yin, G. and Yuan, Y. (2009) Bayesian model averaging continual reassessment method in phase I clinical trials. *Journal of the American Statistical Association*, 104, 954–968.
- Yuan, Y., Nguyen, H.Q. and Thall, P.F. (2016) *Bayesian Designs for Phase I–II Clinical Trials*. New York: Chapman & Hall/CRC.
- Yuan, Y. and Yin, G. (2011a) Robust EM continual reassessment method in oncology dose finding. *Journal of the American Statistical Association*, 106, 818–831.
- Yuan, Y. and Yin, G. (2011b) Bayesian phase I–II adaptively randomized oncology trials with combined drugs. *Annals of Applied Statistics*, 5, 924–942.
- Zhang, J. and Braun, T.M. (2013) A phase I Bayesian adaptive design to simultaneously optimize dose and schedule assignments both between and within patients. *Journal of the American Statistical Association*, 108, 892–901.
- Zhang, W., Sargent, D.J. and Mandrekar, S. (2006) An adaptive dose-finding design incorporating both toxicity and efficacy. *Statistics in Medicine*, 25, 2365–2383.
- Zhou, H., Yuan, Y. and Nie, L. (2018) Accuracy, safety, and reliability of novel phase I trial designs. *Clinical Cancer Research*, 24, 4357–4364.

SUPPORTING INFORMATION

The prior calibration procedure referenced in Section 3, the design configuration and sensitivity analysis referenced in Section 4, and the R code for implementing the proposed design are available with this paper at the Biometrics website on Wiley Online Library.

How to cite this article: Lin R, Thall PF, Yuan Y. An adaptive trial design to optimize dose–schedule regimes with delayed outcomes. *Biometrics*. 2020;76:304–315. <https://doi.org/10.1111/biom.13116>