Biometrics 69, 673–682
September 2013

DOI: 10.1111/biom.12065

Using Joint Utilities of the Times to Response and Toxicity to Adaptively Optimize Schedule–Dose Regimes

Peter F. Thall,[1,*] Hoang Q. Nguyen,[1] Thomas M. Braun,[2] Muzaffar H. Qazilbash[3]

[1]Department of Biostatistics, M.D. Anderson Cancer Center, Houston, Texas, U.S.A.
[2]Department of Biostatistics, School of Public Health, University of Michigan, Ann Arbor, Michigan, U.S.A.
[3]Department of Stem Cell Transplantation, M.D. Anderson Cancer Center, Houston, Texas, U.S.A.
email: rex@mdanderson.org

Summary. A Bayesian two-stage phase I–II design is proposed for optimizing administration schedule and dose of an experimental agent based on the times to response and toxicity in the case where schedules are non-nested and qualitatively different. Sequentially adaptive decisions are based on the joint utility of the two event times. A utility function is constructed by partitioning the two-dimensional positive real quadrant of possible event time pairs into rectangles, eliciting a numerical utility for each rectangle, and fitting a smooth parametric function to the elicited values. We assume that each event time follows a gamma distribution with shape and scale parameters both modeled as functions of schedule and dose. A copula is assumed to obtain a bivariate distribution. To ensure an ethical trial, adaptive safety and efficacy acceptability conditions are imposed on the (schedule, dose) regimes. In stage 1 of the design, patients are randomized fairly among schedules and, within each schedule, a dose is chosen using a hybrid algorithm that either maximizes posterior mean utility or randomizes among acceptable doses. In stage 2, fair randomization among schedules is replaced by the hybrid algorithm. A modified version of this algorithm is used for nested schedules. Extensions of the model and utility function to accommodate death or discontinuation of follow up are described. The method is illustrated by an autologous stem cell transplantation trial in multiple myeloma, including a simulation study.

Key words: Adaptive decision making; Bayesian design; Phase I/II clinical trial; Stem cell transplantation; Utility

1. Introduction

In clinical trials involving cytotoxic or other potentially harmful agents, adverse events (toxicities) generally occur at random times after the start of treatment. Most phase I clinical trial designs determine an optimal dose, or a maximum tolerable dose (MTD), using a binary indicator of toxicity occurring by a predetermined time from the start of therapy. These designs include the continual reassessment method (CRM, O'Quigley, Pepe, and Fisher, 1990) and many others. To use more available information, improve logistics, and protect patients from late onset toxicities, phase I designs based on Y_T =time to toxicity have been proposed, including the time-to-event (TiTE) CRM (Cheung and Chappell, 2000), and the designs of Braun et al. (2007) and Bekele et al. (2008). Many phase I/II designs based on binary or categorical response and toxicity have been proposed (Thall and Russell, 1998; Braun, 2002; Thall and Cook, 2004; Zhang, Sargent, and Mandrekar, 2006). Phase I/II designs also may be based on Y_T and Y_R =time to response. Denoting Y_m^o =time to the event or right-censoring and $\delta_m = I(Y_m = Y_m^o)$ for $m = R, T$, with $\boldsymbol{Y} = (Y_R, Y_T)$, $\boldsymbol{Y}^o = (Y_R^o, Y_T^o)$, and $\boldsymbol{\delta} = (\delta_R, \delta_T)$, dose-finding may be based on $(\boldsymbol{Y}^o, \boldsymbol{\delta})$ (cf. Yuan and Yin, 2009).

Most phase I and phase I/II designs focus on dose, but many agents have schedule-dependent effects. An example in oncology is a nucleoside analog for which the MTD of a 30-minute infusion is (i) 2100 mg/m^2 if given once, (ii) 1000 mg/m^2 if given weekly for three weeks with total dose 3000 mg/m^2 over 21 days, and (iii) 300 mg/m^2 if given twice in each of weeks 1, 3, and 5 for total dose 1800 mg/m^2 over 35 days. An example of an unexpected increase in toxicity after changing the schedule of a preparative agent in stem cell transplantation (SCT) from $(d/2, d/2)$ on Days $(-8, -3)$ to $(d/3, d/3, d/3)$ on Days $(-8, -6, -3)$ is described by Thall (2010, Section 1.1). Braun, Yuan, and Thall (2005) proposed a Bayesian design to optimize the schedule of administration times, $\boldsymbol{s} = (s_1, \ldots, s_k)$, based on (Y_T^o, δ_T), with fixed per-administration dose (PAD), assuming nested schedules with each $\boldsymbol{s}$ corresponding to a number of cycles of therapy. Braun et al. (2007) extended this to allow PAD to vary, and jointly optimized ($\boldsymbol{s}$, PAD) by minimizing the absolute difference between a fixed target probability and the posterior mean probability of toxicity by a specified time, t_T^{ref}, similar to the TiTE CRM. Li et al. (2008) proposed an approach to optimizing dose and schedule for the case of two nested schedules and bivariate binary outcomes. However, no designs currently exist that optimize either schedule or (schedule, dose) in the case of non-nested, qualitatively different schedules, or where the outcomes are bivariate event times.

We address the problem of sequential adaptive optimization of treatment regime $\tau = (\boldsymbol{s}, d)$ in a phase I/II clinical trial where schedules may differ qualitatively or quantitatively, and the outcomes are possibly right-censored event times $(\boldsymbol{Y}^o, \boldsymbol{\delta})$. The total dose is d, with fractions given at the successive administration times. No solution to this design problem currently exists. We propose an adaptive Bayesian method using a utility function, $U(\boldsymbol{y})$, defined on the positive

real quadrant $[0, \infty)^2$ of possible $\boldsymbol{Y}$ values. We construct $U(\boldsymbol{y})$ by partitioning a compact subset of $[0, \infty)^2$ where $\boldsymbol{Y}$ pairs are likely to occur into rectangles, eliciting a numerical utility on each rectangle from the physicians planning the trial, and fitting a parametric function to the elicited values. For each $Y_m,\ m = R, T$, we specify a gamma marginal with shape and scale parameters each modeled as functions of $(\boldsymbol{s}, d)$, and use a copula to obtain association.

The design has two stages, and only allows τ with both acceptable safety and efficacy to be assigned. In stage 1, patients are randomized fairly among schedules in blocks. Within each schedule, the acceptable dose with maximum posterior mean utility is chosen, unless the current sample size for the optimal dose is larger than all current sample sizes for the other doses. In that case, patients are randomized among the assigned schedule's acceptable doses with probabilities proportional to their posterior mean utilities. In stage 2, the block randomization among schedules is unbalanced using similar criteria, with each schedule's assignment probability proportional to the posterior mean utility of its optimal dose. We include randomization to reduce the chance of getting stuck at suboptimal τ, which may occur with a "greedy" algorithm that only maximizes posterior mean utility.

Our design differs from those of Braun et al. (2007) and Li et al. (2008) in that we (1) use both time-to-response and time-to-toxicity, (2) use utilities as decision criteria, (3) use unbalanced randomization to choose regimes, (4) assume a bivariate gamma regression model, and (5) allow non-nested schedules. We also consider the case where Y_R is evaluated at a sequence of times rather than continuously, hence is interval censored. We describe extensions of the model and utility to accommodate death or discontinuation of follow up at toxicity. Our methodology synthesizes the above and several other existing ideas, including use of a copula to obtain association, randomizing to avoid getting stuck at a suboptimal regime, and regression modeling of both scale and shape parameters.

To provide a concrete frame of reference, we describe the illustrative SCT trial in Section 2. Section 3 describes a method for constructing a utility function from elicited values. Gamma regression models for $[Y_R|\boldsymbol{s}, d]$ and $[Y_T|\boldsymbol{s}, d]$, and likelihoods for both continuous and interval censored Y_R are given in Section 4. The design is presented in Section 5. Section 6 illustrates the method by application to the SCT trial, which has two schedules and three doses (six regimes), including a simulation study with comparison to two alternative designs, and evaluation of robustness and of sensitivity to the prior, cohort size, and sample size. We close with a discussion in Section 7.

2. Motivating Application

Melphalan is an alkalating agent commonly used as part of the preparative regimen for autologous SCT to treat multiple myeloma (MM), but there is no consensus for what (schedule, dose) combination is best in older patients. To address this, we designed a phase I/II trial to evaluate total doses $d = 140$, 180, or 200 mg/m^2 of melphalan given either as a single 30-minute bolus infused on Day -2 before SCT, or with the dose split into two equal boluses infused on Days -3 and -2. Toxicity is defined as severe (grade 3 or 4) gastrointestinal toxicity or diarrhea, graft failure, or regimen-related death. Response is evaluated at 1, 3, 6, 9, and 12 months post transplant, so Y_R is interval censored while Y_T is observed continuously, which is common in early phase oncology trials. Response has three requirements, (i) normal bone marrow (<5% myeloma cells), (ii) no new lytic lesions on bone X-ray, and (iii) absence of β_2 microglobulin, a monoclonal protein characterizing MM in two consecutive tests.

Transforming pre-transplant administration Days $(-3, -2)$ to $(0, 1)$, so transplant is on Day 3 after the first administration, the six regimes in the MM trial are $\tau_1 = \{1, 140\}$, $\tau_2 = \{1, 180\}$, $\tau_3 = \{1, 200\}$, $\tau_4 = \{(0, 1), 140\}$, $\tau_5 = \{(0, 1), 180\}$, $\tau_6 = \{(0, 1), 200\}$. The subsets $\mathcal{T}_1 = \{\tau_1, \tau_2, \tau_3\}$ of 1-day schedules and $\mathcal{T}_2 = \{\tau_4, \tau_5, \tau_6\}$ of 2-day schedules have natural orderings, since the probabilities of toxicity and response each increase with d within each schedule. In contrast, for either response or toxicity, there is no obvious ordering among all six regimes in $\mathcal{T} = \mathcal{T}_1 \cup \mathcal{T}_2$. For example, although τ_1 and τ_4 deliver the same total dose, a 1- or 2-day schedule may either be more toxic or have higher response rate. The terms "escalate" or "de-escalate" thus are meaningful when assigning doses within $\mathcal{T}_1$ or $\mathcal{T}_2$, but not for assigning regimes within $\mathcal{T}$. The additive hazard model of Braun et al. (2007) does not accommodate qualitatively different schedules with all administrations given early, or bivariate event time outcomes. We address this by treating schedule as qualitative and dose as quantitative, and defining a joint utility function for the two event times.

3. Utility Functions

We obtain a utility function, $U(\boldsymbol{y})$, that represents the clinical desirability of each possible outcome pair. This utility may be used for decision-making with any sort of regime, and is not limited to dose–schedule optimization. Given follow up time $T_{\max}$, the physician is asked to partition the domains of Y_R and Y_T into subintervals that determine a grid of rectangular subsets partitioning $[0,\ T_{\max}]^2$. A numerical utility is elicited for each rectangular subset, and nonlinear regression is used to fit a smooth surface by treating the midpoints of the rectangles in the $\boldsymbol{Y}$ domain as predictors and the corresponding elicited utilities as outcomes. The partition should be sufficiently refined to provide a discretization of $\boldsymbol{Y}$ in terms of the anticipated joint probability distribution that is realistic based on clinical experience, but sufficiently coarse that the elicitation is feasible. To facilitate refinement of the elicited numerical utilities or the grid, it is useful to show the physician a plot of the fitted surface, and iterate this process until an acceptable utility surface is obtained.

Since smaller y_R and larger y_T are more desirable, $U(\boldsymbol{y})$ must decrease in y_R and increase in y_T, formally, $\partial U(y_R, y_T)/\partial y_R < 0 < \partial U(y_R, y_T)/\partial y_T$. We used the parametric function

$$U(y_R, y_T) = 100\ \frac{b_1\, e^{-c_1 y_R} + b_2\, e^{-c_2 y_T} + b_3\, e^{-c_1 y_R - c_2 y_T} - U_{\min}}{U_{\max} - U_{\min}} \quad \text{for } y_R, y_T > 0. \tag{1}$$

To obtain $0 \le U(y_R, y_T) \le 100$ with 0 corresponding to the worst and 100 to the best possible outcomes, we

Table 1
Utilities for rectangles of Y_R = time to response and Y_T = time to toxicity in the multiple myeloma autologous stem cell transplantation trial

Y_T = months to toxicity	Y_R = months to response				
	[0, 1)	[1, 3)	[3, 6)	[6, 9)	[9, 12)
[9, 12)	95, 93.9	88, 86.0	74, 74.5	64, 62.8	54, 53.1
[6, 9)	85, 85.3	76, 77.4	63, 65.8	53, 54.0	43, 44.3
[3, 6)	75, 73.5	64, 65.5	52, 53.8	42, 41.9	32, 32.1
[1, 3)	62, 60.2	52, 52.1	41, 40.3	31, 28.3	21, 18.4
[0, 1)	50, 50.3	40, 42.2	30, 30.2	20, 18.1	10, 8.1

For each (Y_R, Y_T) rectangle, the two tabled values are $U^{(e)}$ = the elicited utility and $\hat{U}$ =the fitted parametric function evaluated at the rectangle's midpoint.

used the norming constants $U_{\max} = U^o(y_{R,\min}, y_{T,\max})$ and $U_{\min} = U^o(y_{R,\max}, y_{T,\min})$, denoting $U^o(y_R, y_T) = b_1\, e^{-c_1 y_R} + b_2\, e^{-c_2 y_T} + b_3\, e^{-c_1 y_R - c_2 y_T}$. Any compact domain for U may be used, however. The inequalities $c_1, c_2 > 0$, $b_2 < 0 < b_1$, and $b_2 < -b_3 < b_1$ are sufficient to ensure monotonicity of $U(y_R, y_T)$ in each argument. Subject to these constraints, we solved for $(c_1, c_2, b_1, b_2, b_3)$ using nonlinear least squares with the midpoint of each rectangle as the X-variable and the elicited utilities $U^{(e)}$ on the rectangle as the Y-variable. For the autologous SCT trial design (Table 1) this gave estimates $(\hat{c}_1, \hat{c}_2, \hat{b}_1, \hat{b}_2, \hat{b}_3) = (0.0631, 0.1088, 9.3557, -7.8677, 0.5301)$. Table 1 also gives the fitted utilities $\hat{U}(\mathbf{y})$, and the surface obtained by plotting $\hat{U}(\mathbf{y})$ on $\mathbf{y}$ is illustrated by Figure 1, where $T_{\max} = 12$ months for the SCT trial. For example, the rectangle defined by $1 < y_R < 3$ and $3 < y_T < 6$ has midpoint $\boldsymbol{y}^{\text{mid}} = (2, 4.5)$ and elicited utility $U^{(e)} = 64$.

Our criterion for choosing each cohort's treatment regime is the posterior mean utility,

$$u(\tau, \text{data}) = E_{\boldsymbol{\theta}}[E_{\boldsymbol{Y}|\boldsymbol{\theta}}\{U(\boldsymbol{Y}) \mid \tau, \boldsymbol{\theta}\} \mid \text{data}] = E_{\boldsymbol{\theta}}[\bar{U}(\tau, \boldsymbol{\theta}) \mid \text{data}], \tag{2}$$

where we denote $\bar{U}(\tau, \boldsymbol{\theta}) = E_{\boldsymbol{Y}|\boldsymbol{\theta}}\{U(\boldsymbol{Y}) \mid \tau, \boldsymbol{\theta}\}$, the mean over $\boldsymbol{Y}$ of the utility $U(\boldsymbol{Y})$ of using regime τ for given $\boldsymbol{\theta}$. Another way to view $u(\tau, \text{data})$ is obtained by applying the Fubini-Tonelli Theorem to switch the order of expectations in (2). Denoting the joint pdf of $[\boldsymbol{Y} \mid \tau]$ by $f_{R,T}(\mathbf{y} \mid \tau, \boldsymbol{\theta})$, this gives

$$\begin{aligned} u(\tau, \text{data}) &= \int_{\mathbf{y}} U(\mathbf{y}) E_{\boldsymbol{\theta}|\boldsymbol{Y}=\mathbf{y}}\{f_{R,T}(\mathbf{y} \mid \tau, \boldsymbol{\theta}) \mid \text{data}\}\, \mathrm{d}\mathbf{y} \\ &= \int_{\mathbf{y}} U(\mathbf{y}) f_{R,T}(\mathbf{y} \mid \tau, \text{data})\, \mathrm{d}\mathbf{y}. \end{aligned}$$

The posterior expectation is the predictive distribution of $\boldsymbol{Y}$, given the current data, for a patient treated with regime τ. Thus, $u(\tau, \text{data})$ is the expected utility of τ for a newly enrolled patient. The design makes adaptive decisions based on the values of $\{u(\tau, \text{data}), \tau \in \mathcal{T}\}$, subject to safety and efficacy acceptability requirements.

4. Probability Model

4.1. *Marginal Model*

Our modeling strategy is to construct marginals for Y_R and Y_T that are functions of $\boldsymbol{s}$ and d, and use a bivariate copula (Nelsen, 2006) to obtain a joint distribution. For each outcome $m = R, T$ and regime $\tau = (\boldsymbol{s}, d)$, denote the pdf, cdf, and survivor function of Y_m at time $y > 0$ by $f_m(y \mid \tau, \boldsymbol{\theta}_m)$, $F_m(y \mid \tau, \boldsymbol{\theta}_m) = Pr(Y_m \le y \mid \tau, \boldsymbol{\theta}_m)$ and $\bar{F}_m(y \mid \tau, \boldsymbol{\theta}_m) = 1 - F_m(y \mid \tau, \boldsymbol{\theta}_m)$, where $\boldsymbol{\theta}_m$ is the marginal model parameter vector. The joint model parameter vector is $\boldsymbol{\theta} = (\boldsymbol{\theta}_R, \boldsymbol{\theta}_T, \zeta)$, where ζ is the copula's association parameter.

We assume that, given $\boldsymbol{s}$, larger d is associated with stochastically smaller Y_R and smaller Y_T. This says that, at any follow up time y, the probability of response, $F_R(y \mid (\boldsymbol{s}, d), \boldsymbol{\theta}_R)$, and the probability of toxicity, $F_T(y \mid (\boldsymbol{s}, d), \boldsymbol{\theta}_T)$, both increase in d for any $\boldsymbol{s}$. The marginals are formulated so that these probabilities may either vary qualitatively with schedule or have monotone schedule effects. The utility function addresses the conflict between the goals to choose τ to make $F_R(y_R|\tau, \boldsymbol{\theta}_R)$ large while not allowing $F_T(y_T \mid \tau, \boldsymbol{\theta}_T)$ to become unacceptably large by quantifying the desirability of each possible (y_R, y_T) pair.

Let $d_1 < d_2 < \cdots < d_J$ denote the doses being considered. A practical difficulty when using $u\{(\boldsymbol{s}, d), \text{data}\}$ for decision making based on bivariate outcomes is that simply assuming $F_R(t \mid (\boldsymbol{s}, d), \boldsymbol{\theta}_R)$ and $F_T(t \mid (\boldsymbol{s}, d), \boldsymbol{\theta}_T)$ both are monotone in d may not distinguish adequately between different values of $u\{(\boldsymbol{s}, d_j), \text{data}\}$ for doses d_j near the optimum, in the case $d_1 < d_j < d_J$. A given change in an intermediate d_j may produce changes of very different magnitudes in $F_R(t \mid (\boldsymbol{s}, d_j), \boldsymbol{\theta}_R)$ and $F_T(t \mid (\boldsymbol{s}, d_j), \boldsymbol{\theta}_T)$, which in turn may make it difficult to identify a middle dose for which $(\boldsymbol{s}, d_j)$ has true maximum utility. To address this problem, for each outcome we define *outcome-specific standardized doses*,

$$x_{m,j} = \frac{d_1}{\bar{d}} + \left(\frac{d_j - d_1}{d_J - d_1}\right)^{\lambda_m} \left(\frac{d_J - d_1}{\bar{d}}\right), m = R, T, \quad j = 1, \ldots, J,$$

denoting $\bar{d} = (d_1 + \cdots + d_J)/J$. The parameter λ_m controls the relative effects of doses that are not close to either d_1 or d_J. Note that $x_{R,1} = x_{T,1} = d_1/\bar{d}$ and $x_{R,J} = x_{T,J} = d_J/\bar{d}$, while all intermediate standardized doses for f_m are parameterized by λ_m.

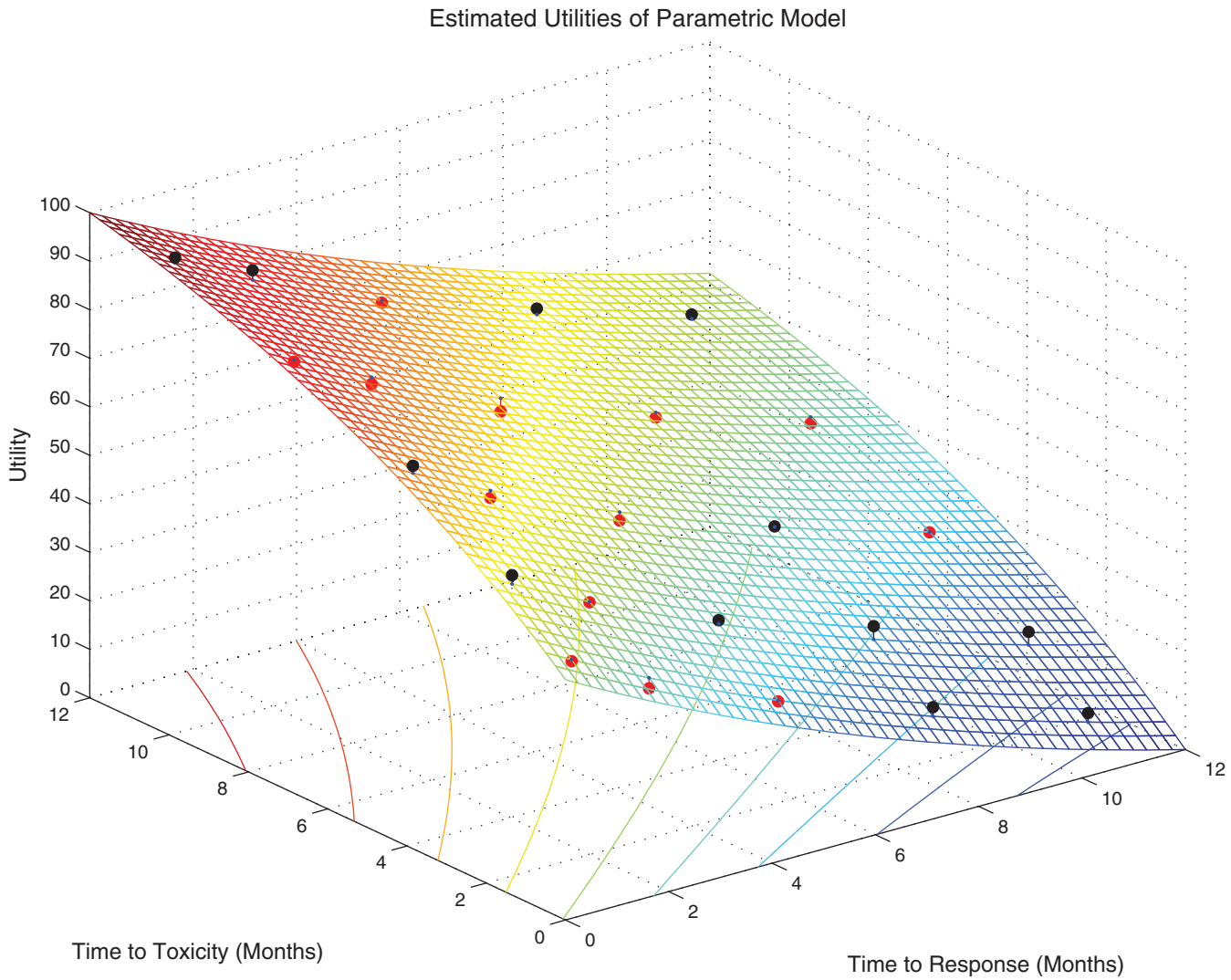

Figure 1. Fitted utility surface for the times to response and toxicity in the multiple myeloma stem cell transplantation trial. Black and red dots show elicited values above and below the fitted surface, respectively.

For brevity, hereafter we will index schedules by $k = 1, \ldots, K$ and denote $\tau = (k, j)$ for the kth schedule and dose d_j. To formulate flexible but reasonably parsimonious marginals for $[Y_m \mid \tau]$, $m = R, T$, in preliminary simulations we explored the lognormal, Weibull, and gamma distributions across a diverse set of regime-outcome scenarios and true event time distributions. We chose the gamma, since it had the best overall performance and robustness of the three distributions. We used gamma marginals having the parametric form

$$f_m(t \mid \tau, \boldsymbol{\theta}_m) = \frac{t^{\phi_{m,1}-1}\, e^{-t/\phi_{m,2}}}{\Gamma(\phi_{m,1})\, \phi_{m,2}^{\phi_{m,1}}},$$

where $\Gamma(\cdot)$ denotes the gamma function. The shape parameter $\phi_{m,1}$ and scale parameter $\phi_{m,2}$ both depend on dose and schedule as follows:

$$\phi_{m,1}\{(k, j), \boldsymbol{\theta}_m\} = \beta_{m,1}(\gamma_{m,k}\, x_{m,j})^{-\alpha_{m,1}} \tag{3}$$

and

$$\log\{\phi_{m,2}\{(k, j), \boldsymbol{\theta}_m\} + 1\} = \beta_{m,2}(\gamma_{m,k}\, x_{m,j})^{-\alpha_{m,2}}, \quad m = R, T. \tag{4}$$

We require $\alpha_{m,1}, \alpha_{m,2}, \beta_{m,1}, \beta_{m,2} > 0$, and assume the schedule effects, $\gamma_{m,1}, \ldots, \gamma_{m,K}$, have support $[0, 2]$. Different transformations are used for $\phi_{m,1}(k, j)$ and $\phi_{m,2}(k, j)$ because the shape and scale parameters play very different roles in determining the form of the gamma distribution, and we found that using a log transformation for $\phi_{m,2}\{(k, j), \boldsymbol{\theta}_m\}$ provided a well behaved dose-outcome model. For each outcome $m = R, T$ and gamma shape ($r = 1$) or scale ($r = 2$) parameter, if dose is fixed and only schedule is varied, the right-hand sides of (3) and (4) reduce to $\beta_{m,r}\gamma_{m,k}$, so there are $K + 1$ parameters for K effects. We thus define $\gamma_{m,1} = 2 - \{\prod_{k=2}^{K} \gamma_{m,k}\}^{1/(K-1)}$.

The utility $U(\boldsymbol{Y})$ reduces the two-dimensional outcome $\boldsymbol{Y}$ to a one-dimensional value, which in turn yields the posterior

mean utility, $u\{(k, j), \text{data}\}$, that is used for decision making. In the models (3) for shape and (4) for scale, the relative magnitudes of the parametric contributions of k and $x_{m,j}$ must reflect their actual effects on $u\{(k, j), \text{data}\}$. In these models, $\beta_{m,r}$ may be thought of as the gamma's usual shape ($r = 1$) or scale ($r = 2$) parameter, modified by the effects of dose and schedule. For each $m = R, T$, the same λ_m is used to define each standardized dose $x_{m,j}$ and, for each schedule k, the same parameter $\gamma_{m,k}$ is used as a multiplicative effect on $x_{m,j}$, for both $\phi_{m,1}$ and $\phi_{m,2}$.

4.2. *Likelihood for Continuously Observed Response Times*

Let t^* denote study time, defined as the time from the start of the trial to the current time when a new patient is enrolled and an interim decision must be made. Let n^* denote the number of patients accrued by t^*. For the ith patient, $i = 1, \ldots, n^*$, denote the treatment regime by τ_i and the outcome vectors evaluated at t^* by $\boldsymbol{Y}^o_{i,t^*} = (Y^o_{i,R,t^*}, Y^o_{i,T,t^*})$ and $\boldsymbol{\delta}_{i,t^*} = (\delta_{i,R,t^*}, \delta_{i,T,t^*})$. For a patient with entry time $e_i < t^*$, the patient time at trial time t^* is $t_i = t^* - e_i$. Each patient's outcome data change over time, starting at $\boldsymbol{Y}_{i,t^*} = \boldsymbol{Y}^o_{i,t^*} = (0, 0)$ and $\boldsymbol{\delta}_{i,t^*} = (0, 0)$ at accrual when $t_i = 0$. Thereafter, each $Y^o_{i,m,t^*} = t_i$ as long as $\delta_{i,m,t^*} = 0$, and Y^o_{i,m,t^*} achieves the final value $Y_{i,m}$ if and when the patient experiences event m, when δ_{i,m,t^*} jumps from 0 to 1. That is, each $(\boldsymbol{Y}^o_{i,t^*}, \boldsymbol{\delta}_{i,t^*})$ is a bivariate sample path of two step functions, jumping from 0 to 1 at their respective event times, with administrative right-censoring, from the time of that patient's accrual to the most recent follow up time. Consequently, before computing posterior quantities used for making outcome-adaptive interim decisions at any study time t^*, it is essential to update the trial data. We denote the interim data at trial time t^* by $\text{data}^* = \{(e_i, \tau_i, \boldsymbol{Y}^o_{i,t^*}, \boldsymbol{\delta}_{i,t^*}) : \ i = 1, \ldots, n^*\}$.

Denote the joint cdf and survivor function of $[\boldsymbol{Y} \mid \tau]$ by $F_{R,T}(\boldsymbol{y} \mid \tau, \boldsymbol{\theta})$ and $\bar{F}_{R,T}(\boldsymbol{y} \mid \tau_i, \boldsymbol{\theta}) = \Pr(Y_R > y_R, Y_T > y_T \mid \tau_i, \boldsymbol{\theta})$. When both Y_R and Y_T are observed continuously, the likelihood for patient i at study time t^* is

$$\begin{aligned}
&\mathcal{L}(\boldsymbol{Y}^o_{i,t^*}, \boldsymbol{\delta}_{i,t^*} \mid \tau_i, \boldsymbol{\theta}) \\
&\quad = \Big\{ f_{R,T}(Y^o_{i,R,t^*}, Y^o_{i,T,t^*} \mid \tau_i, \boldsymbol{\theta}) \Big\}^{\delta_{i,R,t^*}\delta_{i,T,t^*}} \\
&\qquad \times \Big\{ \int_{v=Y^o_{i,T,t^*}}^{\infty} f_{R,T}(Y^o_{i,R,t^*}, v \mid \tau_i, \boldsymbol{\theta})\, \mathrm{d}v \Big\}^{\delta_{i,R,t^*}(1-\delta_{i,T,t^*})} \\
&\qquad \times \Big\{ \int_{u=Y^o_{i,R,t^*}}^{\infty} f_{R,T}(u, Y^o_{i,T,t^*} \mid \tau_i, \boldsymbol{\theta})\, \mathrm{d}v \Big\}^{(1-\delta_{i,R,t^*})\delta_{i,T,t^*}} \\
&\qquad \times \Big\{ \bar{F}_{R,T}(Y^o_{i,R,t^*}, Y^o_{i,T,t^*} \mid \tau_i, \boldsymbol{\theta}) \Big\}^{(1-\delta_{i,R,t^*})(1-\delta_{i,T,t^*})}. \qquad (5)
\end{aligned}$$

Once the marginals have been specified, a joint distribution of Y_R and Y_T may be defined in numerous ways. To obtain a parsimonious and tractable model, we use the bivariate Farlie–Gumbel–Morgenstern (FGM) copula (Nelsen, 2006). Hereafter, we will suppress t^*, i, τ_i, and $\boldsymbol{\theta}$ for brevity when no meaning is lost. The FGM copula is given in terms of the marginals and one association parameter $\zeta \in [-1, \ 1]$ by

$$F_{R,T}(y_R, y_T \mid \zeta) = F_R(y_R)F_T(y_T)\{1 + \zeta\ \bar{F}_R(y_R)\bar{F}_T(y_T)\}. \qquad (6)$$

To obtain the terms in (5) under the FGM copula, for $(\delta_R, \delta_T) = (1, 1)$ the joint pdf is

$$f_{R,T}(y_R, y_T \mid \zeta) = f_R(y_R) f_T(y_T)[1 + \zeta\ \{1-2F_R(y_R)\}\{1-2F_T(y_T)\}],$$

and $\bar{F}_{R,T}(y_R, y_T) = F_{R,T}(y_R, y_T) + \bar{F}_R(y_R) + \bar{F}_T(y_T) - 1$. For $(\delta_R, \delta_T) = (0, 1)$ and $a > 0$,

$$\int_a^{\infty} f_{R,T}(y, y_T)\, \mathrm{d}y \ = \ \bar{F}_R(a) f_T(y_T) \Big[1 - \zeta F_R(a)\{1 - 2F_T(y_T)\Big\}],$$

and the term for $(\delta_R, \delta_T) = (1, 0)$ is obtained by symmetry. All likelihood contributions thus are determined by ζ and the marginal pdfs, with F_R and F_T and terms corresponding to administratively censored event times computed by numerical integration.

4.3. *Likelihood for Interval Censored Response Times*

To account for interval censoring when response is evaluated at successive times $0 = a_0 < a_1 < \cdots < a_{L-1} < a_L = \infty$, rather than continuously, let $A_l = (a_{l-1}, a_l]$ denote the lth subinterval. If a response did not occur by a_{l-1} but did occur by a_l, then $Y_R \in A_l$. Let $\delta_{1,l}$ denote this event. Given the partition $\{A_1, \ldots, A_L\}$ of $[0, \ \infty)$, the pair $(Y^o_{i,R}, \delta_{i,R})$ for continuously observed $Y_{i,R}$ are replaced by the vector of indicators $\boldsymbol{\delta}_{i,R} = (\delta_{i,R,1}, \ldots, \delta_{i,R,L})$, having one entry 1 and all other entries 0. At study time t^*, the observed data of the ith patient are $\{\boldsymbol{\delta}_{i,R}(t^*), Y^o_{i,T,t^*}, \delta_{i,T,t^*}\}$. When $Y^o_{i,T} = Y_{i,T}$ has been observed by study time t^*, so that $\delta_{i,T,t^*} = 1$, the ith patient's likelihood contribution is

$$\begin{aligned}
&\mathcal{L}(\boldsymbol{\delta}_{i,R}(t^*), Y^o_{i,T,t^*}, 1 \mid \tau_i, \boldsymbol{\theta}) \\
&\quad = \prod_{l=1}^{L} \Big\{ \int_{a_{l-1}}^{a_l} f_{R,T}(y, Y^o_{i,T,t^*} \mid \tau_i, \boldsymbol{\theta})\, \mathrm{d}y \Big\}^{\delta_{i,R,l}(t^*)} \\
&\quad = \prod_{l=1}^{L} \Big\{ P^{(1)}_{R,T}(A_l, Y^o_{i,T,t^*} \mid \tau_i, \boldsymbol{\theta}) \Big\}^{\delta_{i,R,l}(t^*)}, \qquad (7)
\end{aligned}$$

denoting $P^{(1)}_{R,T}(A_l, Y^o_T) = \int_{A_l} f_{R,T}(y, Y^o_T)\, \mathrm{d}y$. Under the copula (6), this takes the form

$$\begin{aligned}
P^{(1)}_{R,T}(A_l, Y^o_T) = f_T(Y^o_T)\{F_R(a_l) - F_R(a_{l-1})\}[1 + \zeta\{2F_T(Y^o_T) - 1\} \\
\times \{F_R(a_l) + F_R(a_{l-1}) - 1\}].
\end{aligned}$$

Similarly, when patient i has not yet experienced toxicity, so $\delta_{i,T,t^*} = 0$ and $Y_{i,T}$ is censored at study time t^*, the likelihood

contribution is

$$\begin{aligned}&\mathcal{L}(\boldsymbol{\delta}_{i,R}(t^*), Y^o_{i,T,t^*}, 0 \mid \tau_i, \boldsymbol{\theta})\\ &= \prod_{l=1}^{L}\left\{\int_{Y^o_{i,T,t^*}}^{\infty}\int_{a_{l-1}}^{a_l} f_{R,T}(y, w \mid \tau_i, \boldsymbol{\theta})\,\mathrm{d}y\,\mathrm{d}w\right\}^{\delta_{i,l}(t^*)}\\ &= \prod_{l=1}^{L}\left\{P^{(2)}_{R,T}(A_l, Y^0_{i,T}(t^*) \mid \tau_i, \boldsymbol{\theta})\right\}^{\delta_{i,l}(t^*)}, \qquad (8)\end{aligned}$$

denoting $P^{(2)}_{R,T}(A_l, Y^o_T) = \int_{Y^o_T}^{\infty}\int_{A_l} f_{R,T}(y, w)\,\mathrm{d}y\,\mathrm{d}w$. Under the copula (6), this takes the form

$$\begin{aligned}P^{(2)}_{R,T}(A_l, Y^o_T) = \bar{F}_T(Y^o_T)\{&F_R(a_l)\\ &- F_R(a_{l-1})\}[1 + \zeta F_T(Y^o_T)\{F_R(a_l) + F_R(a_{l-1}) - 1\}].\end{aligned}$$

Combining terms (7) and (8), if Y_R is interval censored the likelihood at trial time t^* is

$$\begin{aligned}&\mathcal{L}(Y^o_{i,T,t^*}, \boldsymbol{\delta}_i(t^*) \mid \tau_i, \boldsymbol{\theta})\\ &= \prod_{l=1}^{L}\left[P^{(1)}_{R,T}(A_l, Y^0_{i,T}(t^*)|\tau_i, \boldsymbol{\theta})^{\delta_{i,T,t^*}} P^{(2)}_{R,T}(A_l, Y^0_{i,T}(t^*)|\tau_i, \boldsymbol{\theta})^{1-\delta_{i,T,t^*}}\right]^{\delta_{i,l}(t^*)}.\end{aligned}$$

5. Trial Design

5.1. *Treatment Regime Acceptability*

While using utilities is a sensible way to combine efficacy and toxicity for optimizing treatment regimes, in practice some regimes may be excessively toxic or inefficacious. Such regimes should not be used to treat patients, and in the extreme case where all regimes are found to be either too toxic or inefficacious the trial should be terminated. We thus employ the following acceptability criteria, similar to those used by Thall and Cook (2004) and others for phase I/II trials. For $m = R, T$, let t^{ref}_m be a reference time from the start of therapy used to specify a limit on $F_m(t^{\mathrm{ref}}_m|\tau, \boldsymbol{\theta})$. Let $\bar{\pi}_T$ be a fixed upper limit on $F_T(t^{\mathrm{ref}}_T|\tau, \boldsymbol{\theta})$ and $\underline{\pi}_R$ be a fixed lower limit on $F_R(t^{\mathrm{ref}}_R|\tau, \boldsymbol{\theta})$, both specified by the physician. Given upper probability cut-offs p_T and p_R, a regime τ is *unacceptable* if

$$\begin{aligned}&Pr\{F_T(t^{\mathrm{ref}}_T|\tau, \boldsymbol{\theta}) > \bar{\pi}_T \mid \mathrm{data}^*\} > p_T \quad \text{or}\\ &Pr\{F_R(t^{\mathrm{ref}}_R|\tau, \boldsymbol{\theta}) < \underline{\pi}_R \mid \mathrm{data}^*\} > p_R \qquad (9)\end{aligned}$$

and we denote the set of acceptable regimes by $\mathcal{A}$.

5.2. *A Design for Non-Nested Schedules*

The problem that a "greedy" sequential search algorithm, that always chooses the optimal action, may get stuck at a suboptimal action is well-known in optimization, but only recently has been addressed in dose-finding trials (Azriel, Mandel, and Rinott, 2011; Thall and Nguyen, 2012). Our proposed design is a hybrid of a greedy design that always chooses $\tau = (k, j)$ to maximize posterior mean utility, and a nonadaptive, hence unethical design that simply randomizes patients fairly among regimes. The idea is to avoid getting stuck at a suboptimal regime, but still conduct the trial ethically by using adaptive rules.

For each successive cohort of c patients, τ is chosen adaptively, as follows. Denote the regime maximizing $u(\tau, \mathrm{data}^*)$ among all $\tau \in \mathcal{A}$ by τ^{opt}. Denote the index of the optimal dose among acceptable regimes having schedule k by

$$j^{\mathrm{opt}}(k) \;=\; \underset{1\le j\le J,\ (k,j)\in\mathcal{A}}{\mathrm{argmax}} \; u\{(k, j), \mathrm{data}^*\}.$$

Because the posterior mean utility $u(\tau, \mathrm{data}^*)$ is highly variable throughout much of the trial, randomizing among regimes with $u\{(k, j), \mathrm{data}^*\}$ close to τ^{opt} is ethical, and reduces the risk of getting stuck at a suboptimal regime. The proposed hybrid design, Design 1, has two stages. Let $n^*(k, j)$ denote the number of patients up to trial time t^* treated with $\tau = (k, j)$. Since only $\tau \in \mathcal{A}$ may be chosen, if $\mathcal{A}$ is empty then the trial is stopped and no τ is selected. If $\mathcal{A}$ is not empty, then for qualitatively different, non-nested schedules Design 1 proceeds as follows. Let N be the maximum overall sample size, and N_1 the maximum stage 1 sample size, with N_1 chosen to be a multiple of Kc reasonably close to $N/2$. The following design first distributes patients evenly among schedules and optimizes dose within each schedule in Stage 1, then optimizes (schedule, dose) globally in Stage 2.

Stage 1. Randomize K cohorts of size c fairly among the schedules, restricted so that each schedule is assigned to exactly c patients. Repeat this until N_1 patients have been treated. Within schedule k, starting at the lowest dose and not skipping an untried dose when escalating, treat the next patient at dose $j^{\mathrm{opt}}(k)$, unless

$$n\{k, j^{\mathrm{opt}}(k)\} \ge \max_j\{n(k, j) :\ j \ne j^{\mathrm{opt}}(k),\ (k, j) \in \mathcal{A}\} + \Delta_1\, c, \qquad (10)$$

where Δ_1 is a small positive integer. If (10) holds, then within schedule k choose an acceptable dose randomly with probability proportional to $u\{(k, j), \mathrm{data}^*\}$.

Stage 2. For $N - N_1$ more patients, choose (k, j) to maximize $u\{k, j, \mathrm{data}^*\}$, unless

$$\begin{aligned}&n\{k^{\mathrm{opt}}, j^{\mathrm{opt}}(k^{\mathrm{opt}})\}\\ &\quad \ge \max_k[n\{k, j^{\mathrm{opt}}(k)\} :\ k \ne k^{\mathrm{opt}}, \{k, j^{\mathrm{opt}}(k)\} \in \mathcal{A}] + \Delta_2\, c, \qquad (11)\end{aligned}$$

where Δ_2 is a small positive integer. If (11) holds, choose a schedule with probability proportional to $u\{(k, j^{\mathrm{opt}}(k)), \mathrm{data}^*\}$.

The inequality (10) in Stage 1 says that the current sample size at the best acceptable dose within schedule k is at least $\Delta_1\, c$ larger than the current sample size at any other acceptable dose with that schedule. One may use $\Delta_1 = \Delta_2 = 1$, or slightly larger values, depending on c and possibly N, to control the amount of sample size imbalance between regimes. The randomization probabilities among doses within schedule k in Stage 1 at t^* are

$$r_1(k, j) \;=\; \frac{u\{(k, j), \mathrm{data}^*\}\, I\{(k, j) \in \mathcal{A}\}}{\sum_{j'=1}^{J} u\{(k, j'), \mathrm{data}^*\}\, I\{(j', k) \in \mathcal{A}\}}, \quad j = 1, \ldots, J.$$

Similarly, the inequality (11) says that the current sample size at the best acceptable regime is at least $\Delta_2\, c$ larger than the current sample size at any other acceptable regime. The randomization probabilities among schedules in Stage 2 at t^* are

$$r_2(k) = \frac{u\{(k,\, j^{\rm opt}(k)),\, {\rm data}^*\}\, I\{(k,\, j^{\rm opt}(k)) \in \mathcal{A}\}}{\sum_{k'=1}^{K} u\{(k',\, j^{\rm opt}(k')),\, {\rm data}^*\}\, I\{(k',\, j^{\rm opt}(k')) \in \mathcal{A}\}},$$
$$k = 1, \ldots, K.$$

Design 2, the "greedy" design, is a much simpler version of Design 1 that chooses $\tau \in \mathcal{A}$ by simply maximizing $u\{(k,\, j),\, {\rm data}^*\}$, subject to the constraint that an untried dose may not be skipped when escalating within any schedule. With Design 2, schedules are chosen by fair randomization without replacement, as in the hybrid Design 1, but this is done throughout the trial, and within schedule k the current dose $j^{\rm opt}(k)$ is chosen.

If schedules are nested, then $\gamma_{m,1} < \gamma_{m,2} < \cdots < \gamma_{m,K}$ for $m = R, T$, and consequently Y_R and Y_T are stochastically increasing in k as well as j, so the regime-finding algorithm must reflect this. Since in this case the word "escalate" pertains to both schedule and dose, i.e. to both k and j, the trial could be conducted by choosing $(k,\, j)$ to maximize $u\{(k,\, j),\, {\rm data}^*\}$ subject to a two-dimensional "do-not-skip" rule similar to that used by Braun et al. (2007), with escalation from $(k,\, j)$ to optimize $u\{\tau,\, {\rm data}^*\}$ restricted to the three adjacent untried combinations $(k+1,\, j)$, $(k,\, j+1)$, or $(k+1,\, j+1)$. This could be elaborated, as in Design 1, to include randomization among regimes based on $u\{(k,\, j),\, {\rm data}^*\}$.

5.3. *Accommodating Death During Follow Up*

The model and utility may be modified to account for death during follow up, or discontinuation of follow up due to toxicity, possibly because the regime was changed at Y_T. This may be done parsimoniously using a semi-competing risks model, wherein we call either death or discontinuation of follow up at Y_T "fatal" toxicity, indicated by δ_{TD}, with δ_{TA} indicating "non-fatal" toxicity that allows follow up to continue for Y_R. Thus, $\delta_{TD} + \delta_{TA} = \delta_T$, and $(\delta_{TD},\, \delta_{TA})$ has possible values (1, 0) or (0, 1) if $\delta_T = 1$ and (0, 0) if $\delta_T = 0$. If $\delta_{TD} = 1$ and $Y_T < Y_R$ then response will not occur. In this case, we define $Y_R = +\infty$ and $\delta_R = 1$, and extend the domain of $(Y_R,\, Y_T)$ from $E_2 = [0,\, \infty)^2$ to $E_2^+ = E_2 \cup [\{+\infty\} \times [0,\, \infty)]$. We do not assume that Y_R censors Y_T, however. Suppressing τ and $\boldsymbol{\theta}$, we define an extended distribution $f^+_{R,T,D}(Y_R, Y_T, \delta_{TD})$ in terms of $\pi_{TD} = Pr(\delta_{TD} = 1)$ and the conditional probabilities $f^+_{R,T|D}(y_R, y_T \mid \delta_{TD} = 0) = f_{R,T}(y_R, y_T)$ and $f^+_{R,T|D}(y_R, y_T \mid \delta_{TD} = 1) = f_{R,T}(y_R, y_T) I(y_R < y_T) + f_T(y_T)\pi_{NR} I(y_R > y_T)$, where $\pi_{NR} = Pr(Y_R > Y_T)$ is the probability of death before response if $\delta_{TD} = 1$. It follows that $f^+_{R,T,D}$ is a probability distribution on E_2^+, since $\int_{E_2^+} \sum_{a=0}^{1} f^+_{R,T|D}(y_R, y_T \mid \delta_{TD} = a) Pr(\delta_{TD} = a)\, {\rm d}y_R\, {\rm d}y_T = 1$.

To extend the likelihood (5) to this case, we first note that lines 2 and 4 of (5) are unchanged since in these cases Y_T is right-censored. The first line of (5) becomes

$$\left[\{f_{R,T}(Y_R^o, Y_T^o)\}^{I(Y_R^o < +\infty)}\, \{f_T(Y_T^o)\pi_{NR}\}^{I(Y_R^o = +\infty)}\, \pi_{TD}^{\delta_{TD}} (1-\pi_{TD})^{\delta_{TA}}\right]^{\delta_R \delta_T}.$$

For line 3 of (5), if Y_R is censored at the time of fatal toxicity, then $\delta_{TD} = 1$ and $Y_R = +\infty$, a case already accounted for by line 1. If Y_R is censored at the time of non-fatal toxicity, this is accounted for by simply replacing δ_T with δ_{TA} in line 3.

The utility may be modified to accommodate death by considering the scaled original utility, $U(y_R, y_T)/100$, as a multiplicative discount factor for survival time on the follow-up interval $[0,\ T_{\max}]$. A utility function that does this is $U^+(y_R, y_{TA}) = T_{\max} U(y_R, y_{TA})/100$ if $\delta_{TA} = 1$ and $U^+(y_R, y_{TD}) = y_{TD}\, U(y_R, y_{TD})/100$ if $\delta_{TD} = 1$. This definition ensures that $U^+(y_R, y_{TD}) < U^+(y_R, y_{TA})$ if $y_{TD} = y_{TA}$. The trial is conducted as described above. The model may be extended similarly if follow up is stopped at Y_R, although this is not commonly done if toxicity occurring during follow up period $[0,\ T_{\max}]$ is considered important. One also might model π_{TD} as a function of $(\mathbf{s}, d)$, if the death rate is sufficiently high to estimate the additional parameters, although this may be unlikely in practice.

6. Application to the SCT Trial

6.1. *Prior and Design Parameters*

We assumed that the positive real valued parameters $\alpha_{R,1}, \alpha_{R,2},\ \beta_{R,1}, \beta_{R,2},\ \lambda_R, \alpha_{T,1},\ \alpha_{T,2},\ \beta_{T,1},\ \beta_{T,2},\ \lambda_T$ followed log-normal priors. The means were determined from the elicited values in Table 2 using a pseudo-sample based method similar to that described in Section 4.2 of Thall et al. (2011). Prior variances were calibrated to obtain a design with good performance across a broad range of scenarios. We assumed that ζ followed a beta distribution with parameters (1.1, 1.1), rescaled to have support on $(-1, +1)$. Numerical values of the prior hyperparameters are given in Supplementary Table S1.

Since each schedule effect acts multiplicatively on the outcome-specific standardized doses, we found that $\gamma_{m,j} \leq 0.80$ or ≥ 1.20 may have a large effect on $x_{m,j}$. For example, $\gamma_{m,j} = 0.80$ would reduce some $x_{m,j}$ values by more than one full dose level. A more disperse prior on the $\gamma_{m,j}$'s also may cause the method to misinterpret a dose effect for a schedule effect in certain cases, especially those where a middle dose has highest utility. Consequently, we specified the priors of the $\gamma_{m,j}$'s to be highly concentrated beta distributions with domain $[0,\ 2]$ and parameters (47.3, 47.3), which gives $\Pr[0.80 < \gamma_{m,j} < 1.20] = 0.95$. Although these priors may appear to be overly informative, in fact small changes within the subdomain [0.80, 1.20] of [0, 2] allow the posterior mean utility to change substantively, so that one may detect true differences between schedules. In this case, the observed data easily have the necessary effect on the posterior distributions of the $\gamma_{m,j}$'s.

6.2. *Simulation Study*

We simulated the SCT trial with $N_1 = 36$, $N = 72$, and $c = 3$, with Y_T monitored continuously and Y_R interval censored per the actual evaluation schedule at 1, 3, 6, 9, 12 months. We studied three competing designs: the hybrid design (Design 1), the greedy design (Design 2), and a randomized design with no interim decisions, restricting the randomization to treat exactly 12 patients at each of the six τ pairs, with the regime maximizing $u(\tau, {\rm data})$ selected at the end. We considered eight simulation scenarios (Supplementary Table S2). In

Table 2
Elicited prior means of $F_R(t \mid \tau)$ and $F_T(t \mid \tau)$ for the autologous stem cell transplantation trial to optimize (schedule, dose) of melphalan

Days of follow-up	Prior means of $F_R(t \mid \tau)$			Prior means of $F_T(t \mid \tau)$		
	Total dose of Melphalan (mg/m^2)			Total dose of Melphalan (mg/m^2)		
	$d = 140$	$d = 180$	$d = 200$	$d = 140$	$d = 180$	$d = 200$
$t = 1$	—	—	—	0.01	0.02	0.03
$t = 3$	—	—	—	0.05	0.07	0.09
$t = 6$	—	—	—	0.15	0.18	0.20
$t = 10$	—	—	—	0.25	0.30	0.32
$t = 14$	—	—	—	0.30	0.35	0.40
$t = 28$	—	—	—	0.33	0.38	0.43
$t = 30$	0.05	0.08	0.10	—	—	—
$t = 90$	0.09	0.11	0.15	—	—	—
$t = 180$	0.13	0.16	0.19	0.34	0.39	0.44
$t = 270$	0.16	0.19	0.24	—	—	—
$t = 360$	0.20	0.25	0.30	0.35	0.40	0.45

For each total dose d, the prior means for the regimes $\tau = (-2, d)$ and $\tau = ((-3, -2), d)$ were identical.

Scenario 1, there is no schedule effect, toxicity is acceptable, and efficacy increases with dose. Scenario 2 also has no schedule effect, but toxicity is much higher, so the lowest dose has the highest utility. In Scenario 3, the 2-day schedule is superior due to higher efficacy. In Scenario 4, the 1-day schedule is superior. Scenario 5 has no schedule effect, but the middle dose is best. In Scenario 6, for both schedules, the utility is "V" shaped, lowest for the middle dose with the highest dose optimal. All regimes are unacceptably toxic in Scenario 7, and unacceptably inefficacious in Scenario 8. Each case was simulated 3000 times.

We use the statistic $R_{\text{select}} = \{u^{\text{true}}(\tau_{\text{select}}) - u_{\min}\}/(u_{\max} - u_{\min})$, (cf. Thall and Nguyen, 2012) to quantify reliability of regime selection. This is the proportion of the difference between the utilities of the best and worst possible regimes achieved by τ_{select}. A statistic quantifying the ethics of how well the method assigns regimes to patients in the trial is $R_{\text{treat}} = \{N^{-1} \sum_{i=1}^{N} u^{\text{true}}(\tau_{[i]}) - u_{\min}\}/(u_{\max} - u_{\min})$, where $u^{\text{true}}(\tau_{[i]})$ is the true utility of the regime given to patient i, and N is the final sample size.

The main simulation results are summarized in Table 3. In each of Scenarios 1–6, the hybrid design does a good job of selecting regimes with high true utilities, and is very likely to correctly stop early in both Scenarios 7 and 8. Table 4 compares the hybrid, greedy, and balanced designs in terms of R_{treat} and R_{select}. More detailed summaries of the simulations of the greedy and balanced non-adaptive designs are given in Supplementary Tables S3a and S3b, respectively. The main messages from Scenarios 1–6 in Table 3 are that (i) compared to the greedy design, the hybrid design has the same or higher R_{select} while neither design is uniformly superior in terms of R_{treat}; (ii) compared to the balanced design, the hybrid design has nearly identical R_{select} but much higher values of R_{treat}, so is much more ethical; and (iii) in Scenarios 7 and 8, both the hybrid and greedy designs correctly stop early with high probability, and both have much higher R_{select} and R_{treat} than the balanced design. In summary, the hybrid design has the best overall performance of the three designs and, as may be expected, the balanced design is ethically unacceptable.

In Supplementary Table S4, we evaluate robustness of the hybrid design to the true event time distribution (lognormal, gamma, Weibull, or uniform). It shows that that (i) R_{treat} is insensitive to the distributions studied, (ii) R_{select} is insensitive to whether the true distribution is lognormal, gamma, or Weibull, but (iii) for a uniform distribution R_{select} may be lower (Scenarios 1 and 3) or higher (Scenarios 2 and 4) than for the other distributions. Supplementary Table S5a shows that the hybrid design is insensitive to changes in prior hyperparameter $\tilde{\sigma} = 8$ to 14, the assumed common prior sd of $\log(\alpha_{m,l})$, $\log(\beta_{m,l})$, and $\log(\lambda_m)$ for all $m = R, T$ and $l = 1, 2$. Supplementary Table S5b shows that the hybrid design is insensitive to changes in the 95% prior interval for the $\gamma_{m,j}$'s varying from (0.9, 1.1) to (0.6, 1.4), although R_{select} and R_{treat} both decrease slightly with the width of this interval in Scenario 5. This motivated our use of the 95% prior interval (0.8, 1.2). Supplementary Table S6 shows that R_{select} is insensitive to $c = 1$, 2, or 3, and that R_{treat} may increase or decrease slightly with c depending on the scenario. Supplementary Table S7 shows that, as N is increased from 48 to 360, both R_{select} and R_{treat} increase substantially.

We also evaluated the hybrid design for an extended version of the SCT trial, with 4 doses and 3 schedules (12 regimes), obtained by interpolating the elicited priors and scenarios of the original 6-regime design. The three doses of the original trial are mapped into the first, third and fourth doses of the extended trial, with a new, second lowest dose corresponding to $d = 160$ mg/m^2 added. Elicited prior and scenario probabilities of the original first two doses were interpolated to obtain values for the new dose. A new third schedule was obtained by averaging the prior and scenario probabilities of the two original schedules. Results for this 12-regime setting are given in Supplementary Tables S8–S14.

Supplementary Table S8a gives a hypothetical utility that places greater weight on quick responses. For either 6 regimes

Table 3
The main simulation results using the hybrid algorithm with sample size 72 and cohort 3

Scenario		1-day schedule			2-day schedule			None	R_{select} / R_{treat}
		Dose 1	Dose 2	Dose 3	Dose 1	Dose 2	Dose 3		
1	$\bar{u}^{\text{true}}(s,d)$	52.2	57.5	62.9	52.2	57.5	62.9		
	% Sel	5	8	38	5	8	35	1	0.82
	# Pats	11.6	9.8	14.5	11.6	9.5	14.4		0.54
2	$\bar{u}^{\text{true}}(s,d)$	59.0	53.7	48.1	59.0	53.7	48.1		
	% Sel	39	7	4	39	6	5	1	0.85
	# Pats	22.6	8.2	5.0	22.7	8.2	4.9		0.75
3	$\bar{u}^{\text{true}}(s,d)$	53.1	58.4	63.8	56.8	62.1	67.6		
	% Sel	3	5	16	6	12	58	1	0.81
	# Pats	11.1	9.1	12.9	11.6	10.5	16.5		0.54
4	$\bar{u}^{\text{true}}(s,d)$	58.6	54.6	49.7	55.4	51.4	46.5		
	% Sel	54	12	5	18	5	5	1	0.80
	# Pats	23.1	9.0	5.1	21.0	8.3	5.1		0.69
5	$\bar{u}^{\text{true}}(s,d)$	52.9	63.6	50.2	52.9	63.6	50.2		
	% Sel	8	34	6	9	36	6	1	0.74
	# Pats	13.0	16.7	6.2	12.8	16.7	6.2		0.54
6	$\bar{u}^{\text{true}}(s,d)$	53.5	48.1	56.5	53.5	48.1	56.5		
	% Sel	21	4	23	21	4	25	2	0.76
	# Pats	17.2	7.3	11.0	17.2	7.2	11.2		0.62
7	$\bar{u}^{\text{true}}(s,d)$	35.3	34.2	33.0	35.3	34.2	33.0		
	% Sel	0	0	0	0	0	0	100	0.87
	# Pats	5.0	1.6	0.5	5.0	1.6	0.5		0.81
8	$\bar{u}^{\text{true}}(s,d)$	39.9	37.8	35.6	39.9	37.8	35.6		
	% Sel	1	0	1	1	0	1	96	0.54
	# Pats	5.8	4.0	4.5	5.8	4.0	4.5		0.54

The true event time distribution is assumed to be lognormal.

(Supplementary Table S8b) or 12 regimes (Supplementary Table S15), the hybrid design's behavior for this different utility, compared to the actual utility, has an equally high probability of correctly stopping the trial early in Scenarios 7 and 8, and in Scenarios 1–6 is better in three cases and worse in three cases. This is desirable, since otherwise there would be little point in using a utility as an objective function.

7. Discussion

We have proposed an adaptive Bayesian method for jointly optimizing schedule of administration and dose in phase I-II trials based on event times for efficacy and toxicity. We modeled schedules qualitatively because either of the two outcome events may occur long after administration. This is very different from the additive hazard model, with a component

Table 4
Summary statistics for the hybrid design, greedy design, and non-adaptive balanced allocation, for the (3-dose, 2-schedule) trial

Scenario	Hybrid		Greedy		Balanced	
	R_{select}	R_{treat}	R_{select}	R_{treat}	R_{select}	R_{treat}
1	0.82 (1)	0.54	0.78 (1)	0.45	0.85 (0)	0.50
2	0.85 (1)	0.75	0.85 (1)	0.84	0.85 (0)	0.51
3	0.81 (1)	0.54	0.78 (1)	0.48	0.83 (0)	0.50
4	0.80 (1)	0.69	0.80 (0)	0.77	0.80 (0)	0.51
5	0.74 (1)	0.54	0.65 (1)	0.48	0.77 (0)	0.40
6	0.76 (2)	0.62	0.75 (3)	0.62	0.77 (0)	0.55
7	0.87 (100)	0.81	1.00 (100)	0.83	0.94 (98)	0.50
8	0.54 (96)	0.54	0.54 (96)	0.59	0.35 (72)	0.50

The "balanced" method assigns 12 patients to each (schedule, dose) pair and does only one posterior computation, at the end of the trial. The number in parentheses after each R_{select} is the percentage of times the trial is stopped with no (schedule, dose) selected. Because scenarios 7 (too toxic) and 8 (too inefficacious) have no acceptable treatments, the R_{select} values are less relevant and thus are shown with a gray background.

for each administration, used by Braun et al. (2007), who dealt with time to toxicity occurring over a much shorter time frame. For regimes administered over a period longer than a few days, our methodology could be extended to allow each patient's initial dose to be changed adaptively based on interim events or new data from other patients.

Our design uses a regime assignment algorithm that is a hybrid of a greedy algorithm and adaptive randomization. Extensive simulations show that, for a maximum sample size of 72, the proposed model and method provide a design that is reliable, safe, and robust, and that it works well in the cases of either six or 12 regimes.

8. Supplementary Materials

Web Appendix 1 referenced in Section 6 is available with this paper at the Biometrics website on Wiley Online Library.

Acknowledgements

The research of Peter Thall and Hoang Nguyen was supported by NCI grant R01 CA 83932. Thomas Braun's research was supported by NIH grant 1 R01 CA148713-01. The authors thank the co-editor, an associate editor, and a referee for helpful comments and suggestions.

References

Azriel, D., Mandel, M., and Rinott, Y. (2011). The treatment versus experimentation dilemma in dose-finding studies. *Journal of Statistical Planning and Inference* **141**, 2759–2758.

Bekele, B. N., Ji, Y., Shen, Y., and Thall, P. F. (2008). Monitoring late onset toxicities in phase I trials using predicted risks. *Biostatistics* **9** 442–457, 2008

Braun, T. M. (2002). The bivariate continual reassessment method: extending the CRM to phase I trials of two competing outcomes. *Contemporary Clinical Trials* **23**, 240–256.

Braun, T. M. (2010). Generalizing the TITE-CRM to adapt for early- and late-onset toxicities. *Statistics in Medicine* **25**, 2071–2083.

Braun, T.M., Yuan, Z., and Thall, P.F. (2005). Determining a maximum tolerated schedule of a cytotoxic agent. *Biometrics* **61**, 335–343.

Braun, T. M., Thall, P. F., Nguyen, H., and de Lima, M. (2007). Simultaneously optimizing dose and schedule of a new cytotoxic agent. *Clinical Trials* **4**, 113–124.

Cheung, Y. and Chappell, R. (2000). Sequential designs for Phase I clinical trials with late-onset toxicities. *Biometrics* **56**, 1177–1182.

Li, Y., Bekele, B. N., Ji, Y., and Cook, J. D. (2008). Dose–schedule finding in phase I/II clinical trials using a Bayesian isotonic transformation. *Statistics In Medicine* **27**, 4895–4913.

Nelsen, R. B. (2006). *An Introduction to Copulas*, 2nd edition. New York: Springer-Verlag.

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.

Thall, P. F. (2010). Bayesian models and decision algorithms for complex early phase clinical trials. *Statistical Science* **25**, 227–244.

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.

Thall, P. F. and Russell, K. T. (1998). A strategy for dose finding and safety monitoring based on efficacy and adverse outcomes in phase I/II clinical trials. *Biometrics* **54**, 251–264.

Thall, P. F., Szabo, A., Nguyen, H. Q., Amlie-Lefond, C. M., and Zaidat, O. O. (2011), Optimizing the concentration and bolus of a drug delivered by continuous infusion. *Biometrics* **67**, 1638–1646.

Yuan, Y. and Yin, G. (2009), Bayesian dose-finding by jointly modeling toxicity and efficacy as time-to-event outcomes. *Journal of the Royal Statistical Society, Series C* **58**, 954–968.

Zhang, W., Sargent, D.J. and Mandrekar, S. (2006), An adaptive dose-finding design incorporating both efficacy and toxicity. *Statistics In Medicine* **25**, 2365–2383.

Received October 2012. Revised February 2013. Accepted April 2013.