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: rexnmdanderson.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 $Y_m^T =$ time to the event or right-censoring and $\delta_m = I(Y_m = Y_m^T)$ for $m = R, T$, with $Y = (Y_R, Y_T)$, $Y'_m = (Y'_R, Y'_T)$, and $\delta = (\delta_R, \delta_T)$, dose-finding may be based on $(Y', \delta)$ (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, $s = (s_1, \ldots, s_k)$, based on $(Y'^T, \delta_T)$, with fixed peradministration dose (PAD), assuming nested schedules with each $s$ corresponding to a number of cycles of therapy. Braun et al. (2007) extended this to allow PAD to vary, and jointly optimized $(s, PAD)$ by minimizing the absolute difference between a fixed target probability and the posterior mean probability of toxicity by a specified time, $t^{ref}_T$, 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 $\tau = (s, d)$ in a phase I/II clinical trial where schedules may differ qualitatively or quantitatively, and the outcomes are possibly right-censored event times $(Y'_T, \delta)$. 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, $U(y)$, defined on the positive
real quadrant $[0, \infty)^2$ of possible $Y$ values. We construct $U(y)$ by partitioning a compact subset of $[0, \infty)^2$ where $Y$ 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 $(s, d)$, and use a copula to obtain association.

The design has two stages, and only allows $\tau$ 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 $\tau$, 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.

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 $Y_R[s, d]$ and $Y_T[s, d]$, 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, limited to dose–schedule optimization. Given follow up time $T_{\text{max}}$, the physician is asked to partition the domains of $Y_R$ and $Y_T$ into subintervals that determine a grid of rectangular subsets partitioning $[0, T_{\text{max}}]$. 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 $Y$ domain as predictors and the corresponding elicited utilities as outcomes. The partition should be sufficiently refined to provide a discretization of $Y$ 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, $U(y)$ 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

$$U(y_R, y_T) = 100 \frac{b_1 e^{-y_T} + b_2 e^{-2y_T} + b_3 e^{-(y_R - 2y_T)}}{U_{\text{max}} - U_{\text{min}}} - U_{\text{min}},$$

for $y_R, y_T > 0$. (1)

To obtain $0 \leq U(y_R, y_T) \leq 100$ with 0 corresponding to the worst and 100 to the best possible outcomes, we
For each \((Y_R, Y_T)\) rectangle, the two tabled values are \(U^{(c)}\) = the elicited utility and \(\bar{U}\) = the fitted parametric function evaluated at the rectangle’s midpoint.

used the norming constants \(U_{\text{max}} = U^*(y_{R_{\text{min}}, y_{T_{\text{max}}}})\) and \(U_{\text{min}} = U^*(y_{R_{\text{max}}, y_{T_{\text{min}}}})\), denoting \(U^*(y_R, y_T) = b_1 e^{-c_1 y_R} + b_2 e^{-c_2 y_T} + b_3 e^{-c_3 y_R - c_4 y_T}\). Any compact domain for \(U\) may be used, however. The inequalities \(c_1, c_2 > 0, b_2 < 0 < b_1\), and \(b_2 < -b_1 < b_1\) 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 \(U^{(c)}\) on the rectangle as the \(Y\)-variable. For the autologous SCT trial design (Table 1) this gave estimates \((\hat{c}_1, \hat{c}_2, b_1, b_2, b_3) = (0.0631, 0.1088, 9.3557, -7.8677, 0.5301)\).

Table 1 also gives the fitted utilities \(\bar{U}(y)\), and the surface obtained by plotting \(\bar{U}(y)\) on \(y\) is illustrated by Figure 1, where \(T_{\text{max}} = 12\) months for the SCT trial. For example, the rectangle defined by \(1 < y_R < 3\) and \(3 < y_T < 6\) has midpoint \(y^{\text{mid}} = (2, 4.5)\) and elicited utility \(U^{(c)} = 64\).

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

\[
u(\tau, \text{data}) = E_{\theta}[E_{\theta_\theta}[U(Y) \mid \tau, \theta] \mid \text{data}] = E_{\theta}[\bar{U}(\tau, \theta) \mid \text{data}],
\]

where we denote \(\bar{U}(\tau, \theta) = E_{\theta}[\bar{U}(Y) \mid \tau, \theta]\), the mean over \(Y\) of the utility \(U(Y)\) of using regime \(\tau\) for given \(\theta\). 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 \([Y \mid \tau]\) by \(f_{R,T}(y \mid \tau, \theta)\), this gives

\[
u(\tau, \text{data}) = \int_U U(y)E_{\theta}[f_{R,T}(y \mid \tau, \theta) \mid \text{data}] dy = \int_{Y} U(y)f_{R,T}(y \mid \tau, \text{data}) dy.
\]

The posterior expectation is the predictive distribution of \(Y\), given the current data, for a patient treated with regime \(\tau\). Thus, \(u(\tau, \text{data})\) is the expected utility of \(\tau\) 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 \(s\) and \(d\), and use a bivariate copula (Nelsen, 2006) to obtain a joint distribution. For each outcome \(m = R, T\) and regime \(s = (s, d)\), denote the pdf, cdf, and survivor function of \(Y_m\) at time \(y > 0\) by \(f_{m}(y \mid \tau, \theta_m)\), \(F_{m}(y \mid \tau, \theta_m)\), and \(F_{m}(y \mid \tau, \theta_m) = 1 - F_{m}(y \mid \tau, \theta_m)\), where \(\theta_m\) is the marginal model parameter vector. The joint model parameter vector is \(\theta = (\theta_R, \theta_T, \zeta)\), where \(\zeta\) is the copula’s association parameter.

We assume that, given \(s\), 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, \(F_R(y \mid (s, d), \theta_R)\), and the probability of toxicity, \(F_T(y \mid (s, d), \theta_T)\), both increase in \(d\) for any \(s\). 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 \(\tau\) to make \(F_R(y_R, \tau, \theta_R)\) large while not allowing \(F_T(y_T \mid \tau, \theta_T)\) to become unacceptably large by quantifying the desirability of each possible \((y_R, y_T)\) pair.

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

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

denoting \(\bar{d} = (d_1 + \cdots + d_J)/J\). The parameter \(\lambda_m\) 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 \(f_m(\text{data})\) are parameterized by \(\lambda_m\).
For brevity, hereafter we will index schedules by $k = 1, \ldots, K$ and denote $\tau = (k, j)$ for the $k$th schedule and dose $d_j$. To formulate flexible but reasonably parsimonious marginals for $[Y_m | \tau]$, $m = R, T$, 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

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

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:

$$\phi_{m,1}(k, j, \theta_m) = \beta_{m,1}(\gamma_{m,k} x_{m,j})^{-\alpha_{m,1}},$$

and

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

(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 $\phi_{m,2}(k, j, \theta_m)$ 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 $U(Y)$ reduces the two-dimensional outcome $Y$ to a one-dimensional value, which in turn yields the posterior

**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.
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 \( x_m \) 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 \( \lambda_m \) is used to define each standardized dose \( x_m,j \) and, for each schedule \( k \), the same parameter \( y_{m,k} \) 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 \( r^* \) 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 \( r^* \). For the \( i \)th patient, \( i = 1, \ldots, n^* \), denote the treatment regime by \( \tau_i \) and the outcome vectors evaluated at \( r^* \) by \( Y_{i,r^*} = (Y_{i,R}, Y_{i,T}) \) and \( \delta_{i,r^*} = (\delta_{i,R}, \delta_{i,T}) \). For a patient with entry time \( e_i < r^* \), the patient time at trial time \( r^* \) is \( t_i = r^* - e_i \). Each patient’s outcome data change over time, starting at \( Y_{i,0} = Y_{i,m} = (0, 0) \) and \( \delta_{i,0} = (0, 0) \) at accrual when \( t_i = 0 \). Thereafter, each \( Y_{i,m,r} = t_i \) as long as \( \delta_{i,m,r} = 0 \), and \( y_{i,m,r} \) achieves the final value \( Y_{i,m} \) if and when the patient experiences event \( m \), when \( \delta_{i,m,r} \) jumps to \( 0 \) to \( 1 \). That is, each \( (Y_{i,m,r}^* \), \( \delta_{i,m,r}^* ) \) is a bivariate sample path of two step functions, jumping from \( 0 \) to \( 1 \) at their respective event times, with administratively 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 \( r^* \), it is essential to update the trial data. We denote the interim data at trial time \( r^* \) by data* = \( \{(e_i, \tau_i, Y_{i,m}, \delta_{i,m}): i = 1, \ldots, n^* \} \).

Denote the joint cdf and survivor function of \([Y \mid \tau] \) by \( F_{R,T}(Y \mid \tau, \theta) \) and \( S_{R,T}(Y \mid \tau, \theta) \) for \( Y = (Y_R, Y_T) > \tau = \tau_i, \theta \). When both \( Y_R \) and \( Y_T \) are observed continuously, the likelihood for patient \( i \) at study time \( r^* \) is

\[
\mathcal{L}(Y_{i,r^*}, \delta_{i,r^*} \mid \tau, \theta) = \left\{ f_{R,T}(Y_{i,R,r^*}, Y_{i,T,r^*} \mid \tau_i, \theta) \right\}^{\delta_{i,R,r^*}} \times \int_{v=r_i}^{\infty} f_{R,T}(v \mid \tau_i, \theta) \, dv \times \left\{ f_{R,T}(u \mid \tau_i, \theta) \right\}^{(1-\delta_{i,R,r^*})} \times F_{R,T}(Y_{i,R,r^*}, Y_{i,T,r^*} \mid \tau_i, \theta) \left\{ (1-\delta_{i,R,r^*})^{(1-\delta_{i,R,r^*})} \right\}.
\]

(5)

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 \( r^* \), \( i \), \( \tau_i \), and \( \theta \) for brevity when no meaning is lost. The FGM copula is given in terms of the marginals and one association parameter \( \xi \in [-1, 1] \) by

\[
F_{R,T}(y_R, y_T \mid \xi) = F_R(y_R)F_T(y_T)[1 + \xi \tilde{F}_R(y_R)\tilde{F}_T(y_T)].
\]

(6)

To obtain the terms in (5) under the FGM copula, for \( (\delta_{i,R}, \delta_{i,T}) = (1,1) \) the joint pdf is

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

and \( \tilde{F}_R(y_R) = F_R(y_R), \tilde{F}_T(y_T) = F_T(y_T) - 1 \). For \( (\delta_{i,R}, \delta_{i,T}) = (0, 1) \) and \( a > 0 \),

\[
\int_{a}^{\infty} f_{R,T}(y, y_T \mid \xi) \, dy = F_R(a)f_T(y_T)[1 - \xi F_R(a)[1 - 2F_T(y_T)]]
\]

and the term for \( (\delta_{i,R}, \delta_{i,T}) = (1, 0) \) is obtained by symmetry. All likelihood contributions thus are determined by \( \xi \) 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 \( l \)th subinterval. If a response did not occur by \( a_{l-1} \) but did occur by \( a_l \), then \( Y_R \in A_l \). Let \( \delta_{i,l} \) denote this event. Given the partition \( \{A_1, \ldots, A_L\} \) of \([0, \infty)\), the pair \((Y_{i,R}^o, \delta_{i,R})\) for continuously observed \( Y_{i,R} \) are replaced by the vector of indicators \( \delta_{i,R} = (\delta_{i,R,1}, \ldots, \delta_{i,R,L}) \), having one entry \( 1 \) and all other entries \( 0 \). At study time \( r^* \), the observed data of the \( i \)th patient are \((\delta_{i,R}(r^*), Y_{i,R}^o, \delta_{i,T}(r^*)) \). When \( Y_{i,T} = Y_{i,T} \) has been observed by study time \( r^* \), so that \( \delta_{i,T,r^*} = 1 \), the \( i \)th patient’s likelihood contribution is

\[
\mathcal{L}(Y_{i,R}^o, Y_{i,T}^o, 1 \mid \tau_i, \theta) = \prod_{l=1}^{L} \left\{ \int_{a_{l-1}}^{a_l} f_{R,T}(y, Y_{i,T}^o \mid \tau_i, \theta) \, dy \right\}^{\delta_{i,R}(r^*)} \times \prod_{l=1}^{L} \left\{ P_{R,T}^{(1)}(A_l, Y_{i,T}^o \mid \tau_i, \theta) \right\}^{\delta_{i,R}(r^*)},
\]

(7)

denoting \( P_{R,T}^{(1)}(A_l, Y_{i,T}^o) = \int_{a_{l-1}}^{a_l} f_{R,T}(y, Y_{i,T}^o) \, dy \). Under the copula (6), this takes the form

\[
P_{R,T}^{(1)}(A_l, Y_{i,T}^o) = f_T(Y_{i,T}^o)(F_R(a_l) - F_R(a_{l-1}))[1 + \xi - 2F_T(Y_{i,T}^o) - 1] \times (F_R(a_l) + F_R(a_{l-1}) - 1).
\]

Similarly, when patient \( i \) has not yet experienced toxicity, so \( \delta_{i,T,r^*} = 0 \) and \( Y_{i,T} \) is censored at study time \( r^* \), the likelihood
A suboptimal regime, but still conduct the trial ethically by using adaptive rules.

For each successive cohort of a patients, is chosen adaptively, as follows. Denote the regime maximizing \( u(\tau, \text{data}^*) \) among all \( \tau \in \mathcal{A} \) by \( \tau^\text{opt} \). Denote the index of the optimal dose among acceptable regimes having schedule \( k \) by \( f^\text{opt}(k) = \arg \max_{1 \leq i \leq c \mid \mathcal{A}(k, j) \in \mathcal{A}} \ u((k, j), \text{data}^*) \).

Because the posterior mean utility \( u(\tau, \text{data}^*) \) is highly variable throughout much of the trial, randomizing among regimes with \( u((k, j), \text{data}^*) \) close to \( \tau^\text{opt} \) 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 \( \tau \) 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.

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 \( t^\text{ref}_m \) be a reference time from the start of therapy used to specify a limit on \( F_m(t^\text{ref}_m | \tau, \theta) \). Let \( \bar{t}_\tau \) be a fixed upper limit on \( F_T(\bar{t}_\tau | \tau, \theta) \) and \( \underline{t}_\tau \) be a fixed lower limit on \( F_R(\underline{t}_\tau | \tau, \theta) \), both specified by the physician. Given upper probability cut-offs \( p_T \) and \( p_R \), a regime \( \tau \) is unacceptable if

\[
Pr[F_T(\bar{t}_\tau \mid \tau, \theta) > \bar{t}_\tau \mid \text{data}^*] > p_T \quad \text{or} \quad Pr[F_R(\underline{t}_\tau \mid \tau, \theta) < \underline{t}_\tau \mid \text{data}^*] > p_R
\]

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
Similarly, the inequality (11) says that the current sample size at the best acceptable regime is at least \( \Delta_2 \) larger than the current sample size at any other acceptable regime. The randomization probabilities among schedules in Stage 2 at \( t^* \) are
\[
r_2(k) = \frac{u((k, j^{opt}(k)), data^*) I((k, j^{opt}(k)) \in A)}{\sum_{k=1}^{K} u((k, j^{opt}(k)), data^*) I((k, j^{opt}(k)) \in A)},
\]
where \( k = 1, \ldots, K \).

Design 2, the “greedy” design, is a much simpler version of Design 1 that chooses \( r \in A \) by simply maximizing \( u((k, j), data^*) \), 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 \( j^{opt}(k) \) is chosen.

If schedules are nested, then \( \gamma_n, 1 < \gamma_{m, 2} \leq \cdots \leq \gamma_{n, K} \) 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 \( u((k, j), data^*) \) subject to a two-dimensional “do-not-skip” rule similar to that elaborated, as in Design 1, to include randomization among 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 \( u((k, j), data^*) \).

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 \( \delta_{TD} \), with \( \delta_{TD} \) indicating “non-fatal” toxicity that allows follow up to continue for \( Y_T \). 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 \( Y_T < Y_K \) 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 \([E_2, E_2] \cup [0, \infty] \times [0, \infty] \). We do not assume that \( Y_K \) censors \( Y_T \), however. Suppressing \( r \) and \( \theta \), 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 \( f_{R,T,D}(Y_R, Y_T | \delta_{TD} = 0) = f_{R,T}(Y_R, Y_T) \) and \( f_{R,T,D}(Y_R, Y_T | \delta_{TD} = 1) = f_{R,T}(Y_R, Y_T)I(y_R < y_T) + f_T(y_T)\pi_{NK}I(y_R > y_T) \), where \( \pi_{NK} = Pr(Y_K > Y_T) \) is the probability of death before response if \( \delta_{TD} = 1 \). It follows that \( f_{R,T,D} \) is a probability distribution on \( E_2^2 \), since
\[
\int_{E_2^2} \sum_{\delta_{TD}=0}^{1} f_{R,T,D}(Y_R, Y_T | \delta_{TD} = a)Pr(\delta_{TD} = a)dy_R dy_T = 1.
\]

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
\[
\left[ f_{R,T}(Y_R, Y_T)I(Y_T < \infty) \{ f_T(y_T)\pi_{NK} \} I(Y_T = \infty) \pi_{TD}^{\delta_{TD}} (1 - \pi_{TD})^{1 - \delta_{TD}} \right]^{\delta_{TD}}.
\]
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

<table>
<thead>
<tr>
<th>Days of follow-up</th>
<th>Prior means of $F_R(t \mid \tau)$</th>
<th>Prior means of $F_T(t \mid \tau)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total dose of Melphalan (mg/m$^2$)</td>
<td>$d = 140$</td>
<td>$d = 180$</td>
</tr>
<tr>
<td>$t = 1$</td>
<td>0.05</td>
<td>0.08</td>
</tr>
<tr>
<td>$t = 3$</td>
<td>0.09</td>
<td>0.11</td>
</tr>
<tr>
<td>$t = 6$</td>
<td>0.13</td>
<td>0.16</td>
</tr>
<tr>
<td>$t = 10$</td>
<td>0.16</td>
<td>0.19</td>
</tr>
<tr>
<td>$t = 14$</td>
<td>0.20</td>
<td>0.25</td>
</tr>
<tr>
<td>$t = 28$</td>
<td>0.33</td>
<td>0.38</td>
</tr>
<tr>
<td>$t = 30$</td>
<td>0.34</td>
<td>0.39</td>
</tr>
<tr>
<td>$t = 90$</td>
<td>0.35</td>
<td>0.40</td>
</tr>
</tbody>
</table>

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

We use the statistic $R_{select} = \{u^{true}(\tau_{select}) - u_{min}\}/(u_{max} - u_{min})$, (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 $\tau_{select}$. A statistic quantifying the ethics of how well the method assigns regimes to patients in the trial is $R_{treat} = \{N^{-1} \sum_{n=1}^{N} u^{true}(\tau_{[n]}) - u_{min}\}/(u_{max} - u_{min})$, where $u^{true}(\tau_{[n]})$ 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.
### Table 3

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

<table>
<thead>
<tr>
<th>Scenario</th>
<th>1-day schedule</th>
<th>2-day schedule</th>
<th>( R_{select} )</th>
<th>( R_{treat} )</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Dose 1</td>
<td>Dose 2</td>
<td>Dose 3</td>
<td>Dose 1</td>
</tr>
<tr>
<td>1</td>
<td>52.2</td>
<td>57.5</td>
<td>62.9</td>
<td>52.2</td>
</tr>
<tr>
<td></td>
<td>% Sel</td>
<td>5</td>
<td>8</td>
<td>38</td>
</tr>
<tr>
<td></td>
<td># Pats</td>
<td>11.6</td>
<td>9.8</td>
<td>14.5</td>
</tr>
<tr>
<td>2</td>
<td>59.0</td>
<td>53.7</td>
<td>48.1</td>
<td>59.0</td>
</tr>
<tr>
<td></td>
<td>% Sel</td>
<td>39</td>
<td>7</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td># Pats</td>
<td>22.6</td>
<td>8.2</td>
<td>5.0</td>
</tr>
<tr>
<td>3</td>
<td>53.1</td>
<td>58.4</td>
<td>63.8</td>
<td>56.8</td>
</tr>
<tr>
<td></td>
<td>% Sel</td>
<td>3</td>
<td>5</td>
<td>16</td>
</tr>
<tr>
<td></td>
<td># Pats</td>
<td>11.1</td>
<td>9.1</td>
<td>12.9</td>
</tr>
<tr>
<td>4</td>
<td>58.6</td>
<td>54.6</td>
<td>49.7</td>
<td>55.4</td>
</tr>
<tr>
<td></td>
<td>% Sel</td>
<td>54</td>
<td>12</td>
<td>5</td>
</tr>
<tr>
<td></td>
<td># Pats</td>
<td>23.1</td>
<td>9.0</td>
<td>5.1</td>
</tr>
<tr>
<td>5</td>
<td>52.9</td>
<td>63.6</td>
<td>50.2</td>
<td>52.9</td>
</tr>
<tr>
<td></td>
<td>% Sel</td>
<td>8</td>
<td>34</td>
<td>6</td>
</tr>
<tr>
<td></td>
<td># Pats</td>
<td>13.0</td>
<td>16.7</td>
<td>6.2</td>
</tr>
<tr>
<td>6</td>
<td>53.5</td>
<td>48.1</td>
<td>56.5</td>
<td>53.5</td>
</tr>
<tr>
<td></td>
<td>% Sel</td>
<td>21</td>
<td>4</td>
<td>23</td>
</tr>
<tr>
<td></td>
<td># Pats</td>
<td>17.2</td>
<td>7.3</td>
<td>11.0</td>
</tr>
<tr>
<td>7</td>
<td>35.3</td>
<td>34.2</td>
<td>33.0</td>
<td>35.3</td>
</tr>
<tr>
<td></td>
<td>% Sel</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td># Pats</td>
<td>5.0</td>
<td>1.6</td>
<td>0.5</td>
</tr>
<tr>
<td>8</td>
<td>39.9</td>
<td>37.8</td>
<td>35.6</td>
<td>39.9</td>
</tr>
<tr>
<td></td>
<td>% Sel</td>
<td>1</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td># Pats</td>
<td>5.8</td>
<td>4.0</td>
<td>4.5</td>
</tr>
</tbody>
</table>

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.

### Table 4

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

<table>
<thead>
<tr>
<th>Scenario</th>
<th>( R_{select} )</th>
<th>( R_{treat} )</th>
<th>( R_{select} )</th>
<th>( R_{treat} )</th>
<th>( R_{select} )</th>
<th>( R_{treat} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.82 (1)</td>
<td>0.54</td>
<td>0.78 (1)</td>
<td>0.45</td>
<td>0.85 (0)</td>
<td>0.50</td>
</tr>
<tr>
<td>2</td>
<td>0.85 (1)</td>
<td>0.75</td>
<td>0.85 (1)</td>
<td>0.84</td>
<td>0.85 (0)</td>
<td>0.51</td>
</tr>
<tr>
<td>3</td>
<td>0.81 (1)</td>
<td>0.54</td>
<td>0.78 (1)</td>
<td>0.48</td>
<td>0.83 (0)</td>
<td>0.50</td>
</tr>
<tr>
<td>4</td>
<td>0.80 (1)</td>
<td>0.69</td>
<td>0.80 (0)</td>
<td>0.77</td>
<td>0.80 (0)</td>
<td>0.51</td>
</tr>
<tr>
<td>5</td>
<td>0.74 (1)</td>
<td>0.54</td>
<td>0.65 (1)</td>
<td>0.48</td>
<td>0.77 (0)</td>
<td>0.40</td>
</tr>
<tr>
<td>6</td>
<td>0.76 (2)</td>
<td>0.62</td>
<td>0.75 (3)</td>
<td>0.62</td>
<td>0.77 (0)</td>
<td>0.55</td>
</tr>
<tr>
<td>7</td>
<td>0.87 (100)</td>
<td>0.81</td>
<td>1.00 (100)</td>
<td>0.83</td>
<td>0.94 (98)</td>
<td>0.50</td>
</tr>
<tr>
<td>8</td>
<td>0.54 (96)</td>
<td>0.54</td>
<td>0.54 (96)</td>
<td>0.59</td>
<td>0.35 (72)</td>
<td>0.50</td>
</tr>
</tbody>
</table>

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.

### 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 outcomes may occur long after administration. This is very different from the additive hazard model, with a component
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

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