Dose Finding for Drug Combination in Early Cancer Phase I Trials in the Presence of a Baseline Binary Covariate using Conditional Escalation with Overdose Control

Mourad Tighiouart, Sungjin Kim

Biostatistics and Bioinformatics Research Center, Cedars-Sinai Medical Center
8700 Beverly Blvd, Los Angeles, CA 90048

Abstract

We present a Bayesian adaptive design for dose finding of a combination of two drugs in cancer phase I clinical trials that takes into account patients heterogeneity thought to be related to treatment susceptibility. The goal is to estimate the maximum tolerated dose (MTD) as a curve for continuous dose levels of the two agents for patient's specific baseline covariate value. Parametric models are used to describe the relationship between the doses, baseline covariate, and the probability of dose limiting toxicity (DLT). Trial design proceeds using univariate escalation with overdose control, where at each stage of the trial, we seek a dose of one agent using the current posterior distribution of the MTD of this agent given the current dose of the other agent and the next patient's baseline covariate value. At the end of the trial, we estimate MTD curves as functions of Bayes estimates of the model parameters. We evaluate design operating characteristics in terms of safety of the trial and percent of dose recommendation at dose combination neighborhoods around the true MTD by comparing the design that uses the covariate to the one that ignores the baseline characteristic.

Key Words: Cancer Phase I trials; Maximum tolerated dose; Escalation with overdose control; Drug combination; Dose limiting toxicity; Continuous dose; baseline covariate.

1. Introduction

The combination of several cytotoxic and biologic agents in drug development and cancer treatment can help reduce tumor resistance to chemotherapy by targeting different signaling pathways simultaneously and improve tumor response when using additive or synergistic drugs [1]. Although the majority of cancer phase I trials use drug combinations of several agents, most of them are designed to estimate the MTD of one drug for fixed dose levels of the other drugs. This approach may provide a single tolerable dose for the combination but it may be suboptimal in terms of therapeutic effects in subsequent efficacy studies.

Trials where the dose levels of at least two agents are allowed to vary yield more than one MTD, or even an infinite number of MTDs in the case of continuous does levels. Estimating the resulting set of MTDs by designing a safe trial is the main goal of phase I trials with dose combinations of several agents. Statistical methodologies for designing such trials have been studied extensively in the past decade [2-13]. These methods assume that the patient population is homogeneous in terms of treatment tolerance and every patient should be treated at a dose combination corresponding to a pre-defined

target probability of DLT. As a result, no allowance is made for individual patient differences in susceptibility to treatment. For single agent trials, strategies of drug allocation that accommodate individual patient needs based on pharmacokinetics and the genetics of drug metabolism have been used in [14-16]. Statistical designs permitting individualized MTD determination in single agent cancer phase I trials have also been proposed and implemented in real trials by a number of authors [17-21]. For drug combinations, an additional layer of complexity in specifying the dose-toxicity relationship given a baseline covariate is needed. Using the notation in [11], the general problem can be stated as follows. Let A_i, $i = 1,\ldots,k$ be k drugs and $S_i \subset \mathrm{R}+$ be the set of all possible doses of drug A_i. Denote by $x = (x_1,\ldots, x_k)$ a dose combination of the k drugs, z the baseline covariate value, and $S = S_1 \times \ldots \times S_k$. Consider a dose-toxicity model

$$\mathrm{Prob}\left(\mathrm{DLT} \mid \mathrm{dose} = \boldsymbol{x}, z\right) = F(\boldsymbol{x}, z, \boldsymbol{\xi}), \tag{1.1}$$

where F is a known link function and $\boldsymbol{\xi} \in \mathrm{R}^{\mathrm{d}}$ is an unknown parameter. The MTD for a patient with baseline covariate z is defined as the set C_z of dose combinations $\boldsymbol{x}$ such that the probability of DLT for a patient with baseline covariate z given dose $\boldsymbol{x}$ equals to a target probability of DLT θ:

$$C_z = \left\{\boldsymbol{x} \in S : F(\boldsymbol{x}, z, \boldsymbol{\xi}) = \theta\right\}. \tag{1.2}$$

In this manuscript, we extend the design described by Tighiouart et al. [11] using escalation with overdose control (EWOC) principle [21-23] by treating cohorts of two patients simultaneously and accounting for patients' baseline binary covariate. We consider a simplified form of model (2.1) by assuming that patients with different covariate values will have parallel MTD curves. This assumption is mathematically convenient and allows us to use parsimonious models due to the small sample size constraints in cancer phase I trials.

2. Model

2.1 Dose-Toxicity Model

Consider the problem of identifying tolerable dose combinations of two cytotoxic agents A and B given a patient with a binary baseline covariate value of z. We consider the dose-toxicity model of the form

$$\mathrm{Prob}\left(\delta = 1 \mid x, y, z\right) = F\left(\beta_0 + \beta_1 x + \beta_2 y + \beta_3 z + \beta_4 xy\right) \tag{2.1}$$

where δ is the indicator of DLT, $\delta = 1$ if a patient given the dose combination (x, y) exhibits DLT within one cycle of therapy, and $\delta = 0$ otherwise, $x \in [X_{min}, X_{max}]$ is the dose level of agent A, $y \in [Y_{min}, Y_{max}]$ is the dose level of agent B, z is a binary baseline covariate, and F is a known cumulative distribution function. Suppose that the doses of agents A and B are continuous and standardized to be in the interval [0, 1]. We will assume that that the probability of DLT increases with the dose of any one of the agents when the other one is held constant for $z = 0, 1$. A necessary and sufficient condition for this property to hold is to assume $\beta_i > 0,\ i = 1, 2$ and the interaction term β_4 is

nonnegative. The MTD for a patient with baseline covariate z is defined as the set C_z of combinations (x^*, y^*) such that

$$\text{Prob}\left(\delta = 1 \mid x^*,\ y^*, z\right) = \theta. \tag{2.2}$$

The target probability of DLT θ is set relatively high when the DLT is a reversible or non-fatal condition, and low when it is life threatening. Using (2.1) and (2.2), the MTD C_z is

$$C_z = \left\{ \left(x^*, y^*\right) \in [0,1]^2 : y^* = \frac{F^{-1}(\theta) - \beta_0 - \beta_1 x^* - \beta_3 z}{\beta_2 + \beta_4 x^*} \right\} \tag{2.3}$$

We reparameterize model (2.1) in terms of parameters that are easily interpreted by the clinicians. These are ρ_{00}, the probability of DLT at the minimum available doses of agents A and B for a patient with covariate value $z = 0$, $\Gamma_{A|0,0}$, the MTD of drug A when the level of drug B is Y_{min} and $z = 0$, $\Gamma_{A|0,1}$, the MTD of drug A when the level of drug B is Y_{min} and $z = 1$, $\Gamma_{B|0,0}$, the MTD of drug B when the level of drug A is X_{min} and $z = 0$, and the interaction parameter β_4. If MTD estimates of single agent trials using A and B are available, then these estimates can be used to approximate the prior mean and variances of the parameters $\Gamma_{A|0,0}$, $\Gamma_{A|0,1}$, and $\Gamma_{B|0,0}$. We will assume that $0 \le \Gamma_{A|0,0}, \Gamma_{A|0,1}, \Gamma_{B|0,0} \le 1$, i.e., the MTD of each agent when the other one is held at its minimum available dose in the trial is within the range of available doses in the trial. This is the reparameterization used in Tighiouart et al. [11] in the absence of a covariate. It follows that

$$\begin{cases} \beta_0 = F^{-1}\left(\rho_{00}\right) \\ \beta_1 = \left(\left(F^{-1}(\theta) - F^{-1}\left(\rho_{00}\right)\right) / \Gamma_{A|0,0}\right. \\ \beta_2 = \left(\left(F^{-1}(\theta) - F^{-1}\left(\rho_{00}\right)\right) / \Gamma_{B|0,0}\right. \\ \beta_3 = \left(\left(F^{-1}(\theta) - F^{-1}\left(\rho_{00}\right)\right) \times \left(\Gamma_{A|0,0} - \Gamma_{A|0,1}\right) / \Gamma_{A|0,0}\right. . \end{cases} \tag{2.4}$$

The MTD (2.3) can be expressed in terms of these new parameters as

$$C_z = \left\{ \left(x^*, y^*\right) : y^* = \frac{\left(\left(F^{-1}(\theta) - F^{-1}\left(\rho_{00}\right)\right)\left(1 - \left(x^* + \left(\Gamma_{A|0,0} - \Gamma_{A|0,1}\right) z\right) / \Gamma_{A|0,0}\right)\right.}{\left(\left(F^{-1}(\theta) - F^{-1}\left(\rho_{00}\right)\right) / \Gamma_{B|0,0} + \beta_4 x^*\right.} \right\}. \tag{2.5}$$

Let $D_n = \{(x_i, y_i, z_i, \delta_i), i = 1, \ldots, n\}$ be the data after enrolling n patients in the trial. The likelihood function under this reparameterization is

$$\begin{aligned} L\left(\rho_{00}, \Gamma_{A|0,0}, \Gamma_{A|0,1}, \Gamma_{B|0,0}, \beta_4; D_n\right) = \prod_{i=1}^{n} & \left(H\left(\rho_{00}, \Gamma_{A|0,0}, \Gamma_{A|0,1}, \Gamma_{B|0,0}, \beta_4; x_i, y_i, z_i\right)\right)^{\delta_i} \\ & \times \left(1 - H\left(\rho_{00}, \Gamma_{A|0,0}, \Gamma_{A|0,1}, \Gamma_{B|0,0}, \beta_4; x_i, y_i, z_i\right)\right)^{1-\delta_i}, \end{aligned} \tag{2.6}$$

where

$$H\left(\rho_{00},\Gamma_{A|0,0},\Gamma_{A|0,1},\Gamma_{B|0,0},\beta_4;x_i,y_i,z_i\right) = F\left(F^{-1}\left(\rho_{00}\right)+\frac{\left(F^{-1}\left(\theta\right)-F^{-1}\left(\rho_{00}\right)\right)}{\Gamma_{A|0,0}}x_i+\frac{\left(F^{-1}\left(\theta\right)-F^{-1}\left(\rho_{00}\right)\right)}{\Gamma_{B|0,0}}y_i + \frac{\left(F^{-1}\left(\theta\right)-F^{-1}\left(\rho_{00}\right)\right)\times\left(\Gamma_{A|0,0}-\Gamma_{A|0,1}\right)}{\Gamma_{A|0,0}}z_i+\beta_4 x_i y_i\right). \quad (2.7)$$

2.2 Prior and Posterior Distributions

Equations (2.4) imply that $0 < \rho_{00} < \theta$ since $\beta_i > 0,\ i = 1, 2$. We consider the priors ρ_{00}/θ ~ beta(a_0, b_0), $\Gamma_{A|0,0}$ ~ beta(a_1, b_1), $\Gamma_{A|0,1}$ ~ beta(a_2, b_2), $\Gamma_{B|0,0}$ ~ beta(a_3, b_3), β_4 ~ gamma(a, b) with mean $E(\beta_4) = a/b$ and variance $Var(\beta_4) = a/b^2$. Vague priors for these parameters are achieved by taking $a_j = b_j = 1$, $j = 0, 1, 2, 3$. The prior mean for β_4 is selected by setting $E(\beta_4) = 8\left(F^{-1}(\theta) - F^{-1}\left(E(\rho_{00})\right)\right)/E\left(\Gamma_{A|0,0}\right)E\left(\Gamma_{B|0,0}\right)$, see [11] for the rationale behind this choice. A vague prior is achieved by selecting a large variance for β_4,. Using Bayes rule, the posterior distribution of the model parameters is proportional to the product of the likelihood and prior distribution

$$\pi\left(\rho_{00},\Gamma_{A|0,0},\Gamma_{A|0,1},\Gamma_{B|0,0},\beta_4;D_n\right) \propto \prod_{i=1}^{n} H\left(\rho_{00},\Gamma_{A|0,0},\Gamma_{A|0,1},\Gamma_{B|0,0},\beta_4;x_i,y_i,z_i\right)^{\delta_i} \times\left(1-H\left(\rho_{00},\Gamma_{A|0,0},\Gamma_{A|0,1},\Gamma_{B|0,0},\beta_4;x_i,y_i,z_i\right)\right)^{1-\delta_i} \times\pi\left(\rho_{00}\right)\pi\left(\Gamma_{A|0,0}\right)\pi\left(\Gamma_{A|0,1}\right)\pi\left(\Gamma_{B|0,0}\right)\pi\left(\beta_4\right). \quad (2.8)$$

Features of this posterior distribution are estimated using WinBUGS[24] and JAGS.

2.3 Trial Design

The algorithm for dose escalation is similar to the ones described in [11, 25]. It uses the EWOC principle [21-23] where at each stage of the trial, we seek a dose of one agent using the current posterior distribution of the MTD of the agent given the current dose of the other agent and the next patient's baseline covariate value. For instance, if agent A is held constant at level x, the dose of agent B is y such that the posterior probability that y exceeds the MTD of agent B given the dose of agent $A = x$ and covariate value z is bounded by a feasibility bound α. Cohorts of two patients are enrolled simultaneously receiving different dose combinations. Specifically, the design proceeds as follows.

1. Each patient in the first cohort of two patients receives the same dose combination $(x_1, y_1) = (x_2, y_2) = (0, 0)$. Let $D_2 = \{(x_1, y_1, z_1, \delta_1), (x_2, y_2, z_2, \delta_2)\}$.
2. In the second cohort of two patients, patient 3 receives dose (x_1, y_3) and patient 4 receives dose (x_4, y_2), where y_3 is the α-th percentile of $\pi(\Gamma_{B|A=x1,\ Z=z3} \mid D_2)$ and x_4 is the α-th percentile of $\pi(\Gamma_{A|B=y2,\ Z=z4} \mid D_2)$. Here, $\pi(\Gamma_{B|A=x1,\ Z=z3} \mid D_2)$ is the posterior distribution of the MTD of agent B given that the level of agent A is x_1 and the

baseline covariate value of patient 3 is z_3, given the data D_2. $\pi(\Gamma_{A|B=y2,\ Z=z4} \mid D_2)$ is defined similarly.

3. In the i-th cohort of two patients, if i is even, patient $(2i-1)$ receives dose $(x_{2i\text{-}3}, y_{2i\text{-}1})$ and patient $2i$ receives dose $(x_{2i}, y_{2i\text{-}2})$, where $y_{2i-1}=\Pi^{-1}_{\Gamma_{B|A=x(2i-3),Z=z(2i-1)}}(\alpha|D_{2i-2})$ and $x_{2i}=\Pi^{-1}_{\Gamma_{A|B=y(2i-2),Z=z\,(2i)}}(\alpha|D_{2i-2})$. If i is odd, then patient $(2i-1)$ receives dose $(x_{2i\text{-}1}, y_{2i\text{-}3})$, patient $2i$ receives dose $(x_{2i\text{-}2}, y_{2i})$,where $x_{2i-1}=\Pi^{-1}_{\Gamma_{A|B=y(2i-3),Z=z(2i-1)}}(\alpha|D_{2i-2})$ and $y_{2i}=\Pi^{-1}_{\Gamma_{B|A=x(2i-2),Z=z(2i)}}(\alpha|D_{2i-2})$.
4. Repeat step 3 until n patients are enrolled to the trial subject to the following stopping rule.

Stopping rule:
We stop enrollment to the trial if $P(P(\text{DLT}|(x, y) = (0,0), z) \geq \theta+\delta_1 \mid \text{data}) > \delta_2$, i.e. if the posterior probability that the probability of DLT at the minimum available dose combination in the trial exceeds the target probability of DLT is high for $z = 0, 1$. δ_1 and δ_2 are design parameters chosen to achieve desirable model operating characteristics.

At the end of the trial, we estimate the MTD curve using Bayes estimates of the parameters defining this curve. For example, an estimate of the MTD curve for $z = 0, 1$ is obtained using (2.5) as

$$\hat{C}_z = \left\{ \left(x^*, y^*\right) \in [0,1]^2 : y^* = \frac{\left(\left(F^{-1}(\theta) - F^{-1}(\hat{\rho}_{00})\right)\left(1 - \frac{\left(x^* + \left(\hat{\Gamma}_{A|0,0} - \hat{\Gamma}_{A|0,1}\right)z\right)}{\hat{\Gamma}_{A|00}}\right)}{\frac{\left(\left(F^{-1}(\theta) - F^{-1}(\hat{\rho}_{00})\right)\right)}{\hat{\Gamma}_{B|0,0}} + \hat{\beta}_4 x^*} \right\} \quad (2.9)$$

where $\hat{\rho}_{00}$, $\hat{\Gamma}_{A|00}$, $\hat{\Gamma}_{A|01}$, $\hat{\Gamma}_{B|00}$, and $\hat{\beta}_4$ are the posterior medians given the data D_n.

3. Simulation Studies

3.1 Simulation Set-up and Scenarios

We evaluate design operating characteristics by assuming a logistic link function $F(u) = (1 + exp(-u))^{-1}$ for the working and true models. DLT responses are generated assuming a logistic link function. We present two scenarios for the true MTD curve. In all cases, the target probability of DLT is fixed at $\theta = 0.33$, the true value for ρ_{00} is 0.05, the interaction coefficient β_4 is 10, and the trial sample size is $n = 40$ patients with 20 patients in each group. In all scenarios $a_j = b_j = 1$, $j = 0, 1, 2, 3$ which correspond to uniform priors for ρ_{00}/θ, $\Gamma_{A|0,0}$, $\Gamma_{A|0,1}$, and $\Gamma_{B|0,0}$, and a vague prior for β_4 was selected by taking $E(\beta_4) = 30$ and $Var(\beta_4) = 900$. The first scenario corresponds to $\Gamma_{A|0,0} = \Gamma_{B|0,0} = 0.4$ and $\Gamma_{A|0,1} = 0.8$, and the corresponding true MTD curves are shown by the solid lines in Figure 1(a,b,c). In the second scenario, $\Gamma_{A|0,0} = \Gamma_{B|0,0} = 0.7$ and $\Gamma_{A|0,1} = 0.8$. The true MTD curves are displayed by solid lines in Figure 2(a,b,c). Other scenarios were studied but are not included here due to space limitation. For each scenario, $m = 1000$ trials were simulated using the logistic link function for the working and true model.

3.2 Design Operating Characteristics

3.2.1 Comparison of three Designs

In order to assess the performance of this method when designing a prospective trial, we evaluate its operating characteristics by comparing the following three designs.

a) Design using a covariate; patients are accrued to the trial sequentially and the dose combinations given to the next cohort of patients is calculated assuming model (2.1).
b) Design ignoring the covariate; patients are accrued to the trial sequentially and the dose combinations given to the next cohort of patients is calculating assuming model (2.1) without the covariate, i.e., as in [11, 25].
c) Design using separate trials; in each group, patients are accrued to the trial sequentially and model (2.1) without the covariate is implemented in each group.

We evaluate and compare the performance of these methods by assessing the safety of the trial designs and the efficiency of the estimate of the MTD curve.

3.2.2 Safety

We assess trial safety by reporting the average percent of DLTs across all $m = 1000$ trials and the percent of trials that have a DLT rate exceeding $\theta + \delta$, for $\delta = 0.05$ and 0.1. The threshold $\theta + 0.1$ is usually considered to be an indication of an excessive DLT rate.

3.2.3 Efficiency

We present an estimate of the MTD curve using the average posterior medians of the model parameters. The estimate for $z = 0, 1$ is

$$\overline{C}_z = \left\{ \left(x^*, y^*\right) : y^* = \frac{\left(\left(F^{-1}(\theta) - F^{-1}\left(\overline{\rho}_{00}\right)\right)\left(1 - \left(x^* + \left(\overline{\Gamma}_{A|0,0} - \overline{\Gamma}_{A|0,1}\right)z\right) / \overline{\Gamma}_{A|0,0}\right)\right)}{\left(\left(F^{-1}(\theta) - F^{-1}\left(\overline{\rho}_{00}\right)\right) / \overline{\Gamma}_{B|0,0} + \overline{\beta}_4 x^*\right.} \right\} \quad (3.1)$$

where $F(\cdot)$ is the logistic function and $\bar{\rho}_{00}$, $\bar{\Gamma}_{A|0,0}$, $\bar{\Gamma}_{A|0,1}$, $\bar{\Gamma}_{B|0,0}$, and $\bar{\beta}_4$ are the average posterior medians of the parameters ρ_{00}, $\Gamma_{A|0,0}$, $\Gamma_{A|0,1}$, $\Gamma_{B|0,0}$, and β_4 from all $m = 1000$ trials, respectively. The next measure of efficiency is the pointwise average relative minimum distance from the true MTD curve to the estimated MTD curve. For $i = 1, \ldots, m$, let C_i be the estimated MTD curve and C_{true} be the true MTD curve. For every point $(x, y) \in C_{true}$, let

$$d^{(i)}_{(x,y)} = sign\left(y' - y\right) \times \min_{\left\{\left(x^*, y^*\right) : \left(x^*, y^*\right) \in C_i\right\}} \left(\left(x - x^*\right)^2 + \left(y - y^*\right)^2\right)^{\frac{1}{2}} \quad (3.2)$$

where y' is such that $(x, y') \in C_i$. This is the minimum relative distance of the point (x, y) on the true MTD curve to the estimated MTD curve C_i. If the point (x, y) is below C_i, then $d^{(i)}_{(x,y)}$ is positive. Otherwise, it is negative. Let

$$d_{(x,\,y)} = m^{-1} \sum_{i=1}^{m} d^{(i)}_{(x_i, y_i)}. \quad (3.3)$$

This is the pointwise average relative minimum distance from the true MTD curve to the estimated MTD curve and can be interpreted as the pointwise average bias in estimating the MTD. Let $\Delta(x, y)$ be the Euclidian distance between the minimum dose combination (0, 0) and the point (x, y) on the true MTD curve and $0 < p < 1$. The last measure of efficiency we consider is

$$P_{(x,\ y)} = m^{-1}\sum_{i=1}^{m} I\left(|\, d^{(i)}_{(x_i, y_i)}\, | \leq p\Delta(x, y)\right). \tag{3.4}$$

This is the pointwise percent of trials for which the minimum distance of the point (x, y) on the true MTD curve to the estimated MTD curve C_i is no more than $(100 \times p)$% of the true MTD. This statistic is equivalent to drawing a circle with center (x, y) on the true MTD curve and radius $p\Delta(x, y)$ and calculating the percent of trials with MTD curve estimate C_i falling inside the circle. This will give us the percent of trials with MTD recommendation within $(100 \times p)$% of the true MTD for a given tolerance p.

Table 1. Average percent of DLTs and percent of trials with DLT rate exceeding $\theta + \delta$ for each design under two scenarios.

Scenario ($\Gamma_{A\|0,0}$, $\Gamma_{B\|0,0}$, $\Gamma_{A\|,01}$)	Design	Average % DLT	% Trials: DLT rate > θ + 0.05	% Trials: DLT rate > θ + 0.10
(0.4, 0.4, 0.8)	With covariate	27	0.00	0.00
	Parallel trials	25	0.03	0.00
	Ignoring covariate	27	0.00	0.00
(0.7, 0.7, 0.8)	With covariate	24	0.00	0.00
	Parallel trials	22	0.02	0.00
	Ignoring covariate	24	0.00	0.00

3.3 Results

3.3.1 Trial Safety

Table 1 shows that for each design and under the two scenarios, the average percent of DLTs varies between 22% and 27%. In general, the average DLT rate tends to be lower when the true MTD curve is farther away from the minimum dose combination. Furthermore, the percent of trials with an excessive rate of DLT as defined by a DLT rate exceeding θ + 0.1 is 0.00 using the three designs under the two scenarios. Further simulations (results not shown) under scenarios where $\Gamma_{A|0,0}$ and $\Gamma_{B|0,0}$ are different show that this rate does not exceed 2%. Based on these findings, we conclude that the methodology is safe in general.

3.3.2 Trial Efficiency

Figures 1(a,b) and 2(a,b) show the true and estimated MTD curves for each group of patients. The dashed line in figures 1(c) and 2(c) is the estimated MTD curve obtained from the model without a baseline covariate. The estimated MTD curves shown by dashed lines were obtained using (3.1) and DLT responses were simulated using the true logistic model. Figures 1(a) and 1(b) show that the estimated MTD curves are much closer to the true MTD curves when accounting for a significant baseline covariate relative to parallel trial designs. When ignoring the covariate, the estimated MTD curve tends to be in between the true MTD curves, see Figures 1(c) and 2(c). This shows that

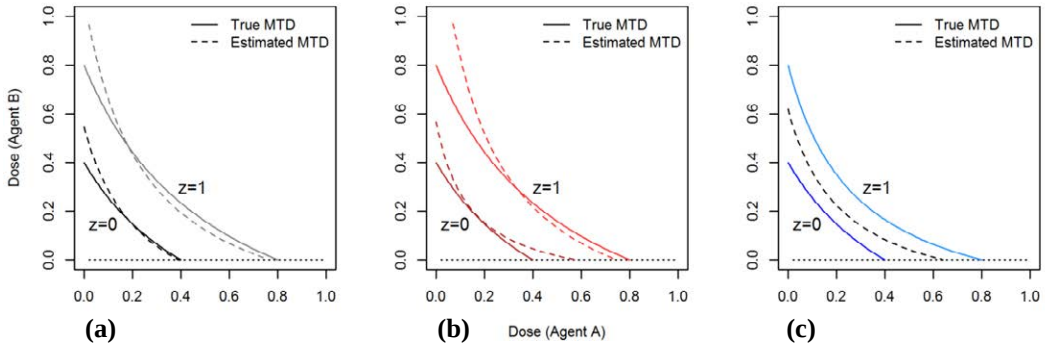

Figure 1. True and estimated MTD curves for each group of patients from m =1000 simulated trials under the first scenario with $\Gamma_{A|0,0} = \Gamma_{B|0,0} = 0.4$ and $\Gamma_{A|0,1} = 0.8$ when using (a) proposed design with a binary baseline covariate, (b) parallel trials, and (c) design ignoring a binary baseline covariate. The dashed lines represent estimated MTD curves and the solid lines indicate true MTD curves.

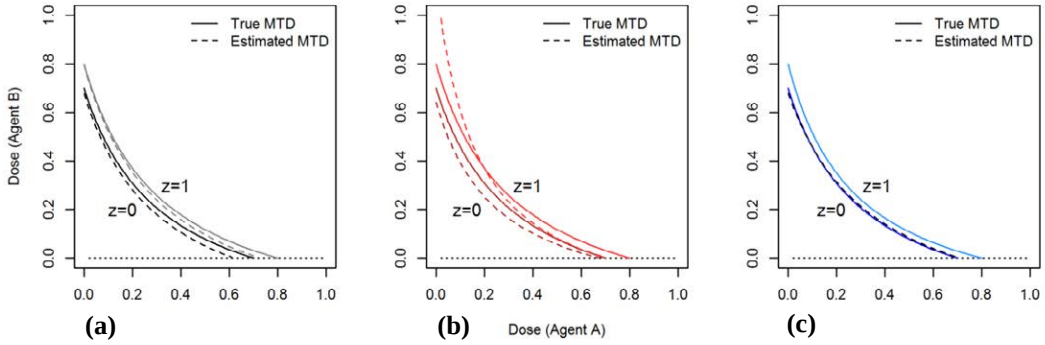

Figure 2. True and estimated MTD curves for each group of patients from m =1000 simulated trials under the second scenario with $\Gamma_{A|0,0} = \Gamma_{B|0,0} = 0.7$ and $\Gamma_{A|0,1} = 0.8$ when using (a) proposed design with a binary baseline covariate, (b) parallel trials, and (c) design ignoring a binary baseline covariate. The dashed lines represent estimated MTD curves and the solid lines indicate true MTD curves.

when the two MTD curves are well separated, not accounting for a baseline covariate results in suboptimal MTD curve estimation for one group of patients and a too toxic MTD curve recommendation for the other group.

Figures 3 and 4 display the pointwise average relative minimum distance from the true MTD curve to the estimated MTD curve as defined by (3.3). This is a measure of pointwise average bias of the estimate of the MTD. In scenario 1, the maximum average bias with our proposed design with a covariate is about 0.04 for $z = 0$ and 0.05 for $z = 1$, which correspond to 6.7% and 4.4% of the distance from the minimum dose combination (0, 0) to the true MTD dose combinations (0.4, 0.4) and (0.8, 0.8), respectively (Figure 3(a)). Similar measures of the average bias are found in Figure 3(b) when using parallel trials. In scenario 2 where the true MTD curves for $z = 0$ and $z = 1$ are very close, the pointwise average bias is very close when using designs with (Figure 4 (a)) and without covariate (Figure 4(c)) and slightly higher with the parallel trials (Figure 4(b)). Other scenarios (results not shown) also show that the maximum average bias with our

proposed design with a covariate is no more than 12% of the distance from (0, 0) to the corresponding true MTD dose combination.

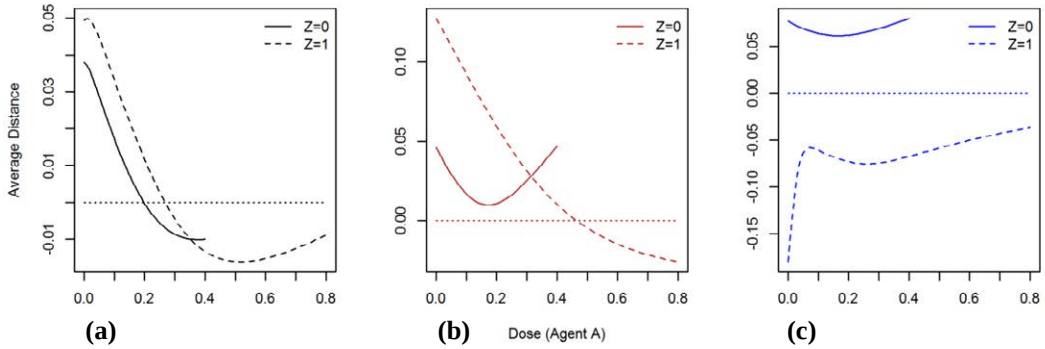

Figure 3. Pointwise average bias for each group of patients (dashed lines for patients with the covariate value of z=1; solid lines for patients with the covariate value of z=0) from m =1000 simulated trials under the first scenario with $\Gamma_{A|0,0} = \Gamma_{B|0,0} = 0.4$ and $\Gamma_{A|0,1} = 0.8$ when using (1) proposed design with a binary baseline covariate, (b) parallel trials, and (c) design ignoring a binary baseline covariate.

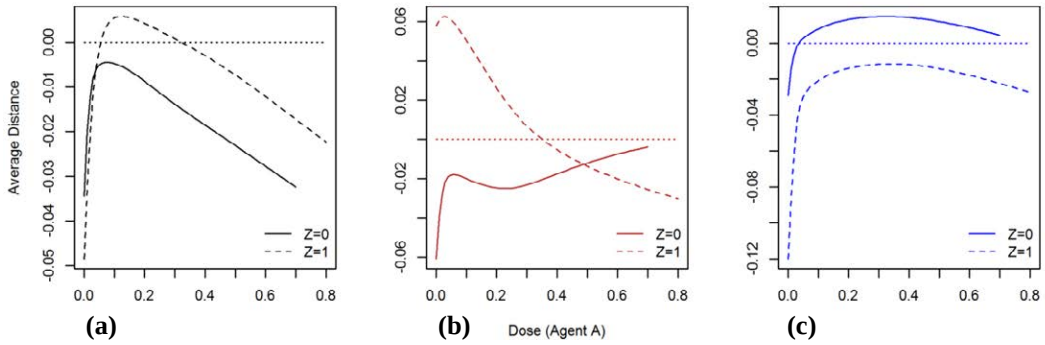

Figure 4. Pointwise average bias for each group of patients (dashed lines for patients with the covariate value of z=1; solid lines for patients with the covariate value of z=0) from m =1000 simulated trials under the second scenario with $\Gamma_{A|0,0} = \Gamma_{B|0,0} = 0.7$ and $\Gamma_{A|0,1} = 0.8$ when using (1) proposed design with a binary baseline covariate, (b) parallel trials, and (c) design ignoring a binary baseline covariate.

Figures 5 and 6 show the pointwise percent of trials for which the minimum distance from the true MTD curve to the estimated MTD curve is no more than $(100 \times p)$% of the true MTD for p = 0.2. This can be interpreted as the percent of MTD recommendation for a given tolerance p. In scenario 1, the percent of trials with correct MTD recommendation varies between 67% and 100% and 66% and 98% with the designs accounting for a significant baseline covariate as shown in Figures 5(a) and 5(b), respectively while with the design ignoring a covariate the percent of MTD recommendation varies more widely between 37% and 99% (Figure 5(c)). Under scenario 2, the percent recommendation becomes slightly higher varying between 71% and 100% and 76% and 100% with the designs with and without a baseline covariate as shown in Figures 6(a) and 6(c), respectively while it is slightly lower varying between 60% and 100% with the parallel trials (Figure 6(b)). Under scenarios where $\Gamma_{A|0,0}$ and $\Gamma_{B|0,0}$ are different (results not

shown) the percent recommendation with our proposed design varies more widely between 32% and 100%.

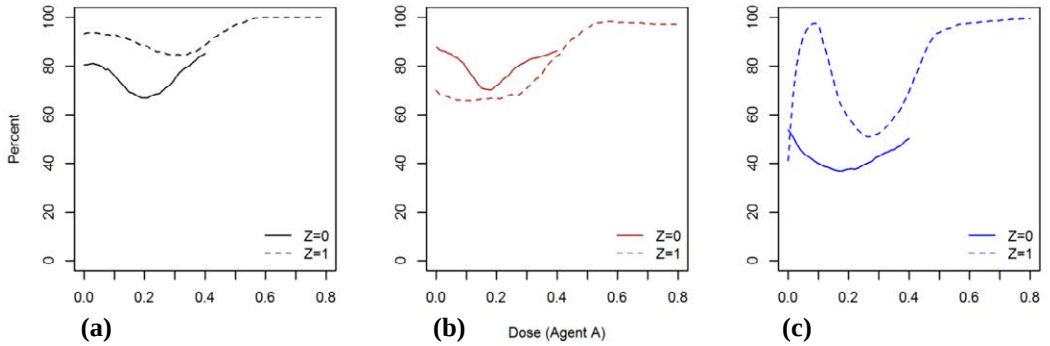

Figure 5. Pointwise percent MTD recommendation for tolerance p = 0.2 for each group of patients (dashed lines for patients with the covariate value of z=1; solid lines for patients with the covariate value of z=0) from m =1000 simulated trials under the first scenario with $\Gamma_{A|0,0} = \Gamma_{B|0,0}$ = 0.4 and $\Gamma_{A|0,1}$ = 0.8 when using (1) proposed design with a binary baseline covariate, (b) parallel trials, and (c) design ignoring a binary baseline covariate.

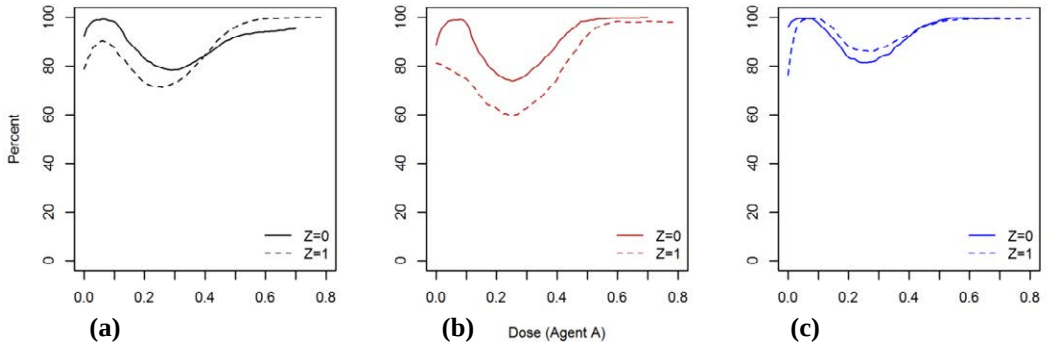

Figure 6. Pointwise percent MTD recommendation for tolerance p = 0.2 for each group of patients (dashed lines for patients with the covariate value of z=1; solid lines for patients with the covariate value of z=0) from m =1000 simulated trials under the second scenario with $\Gamma_{A|0,0} = \Gamma_{B|0,0}$ = 0.7 and $\Gamma_{A|0,1}$ = 0.8 when using (1) proposed design with a binary baseline covariate, (b) parallel trials, and (c) design ignoring a binary baseline covariate.

The percent recommendation increases as we move away from the minimum available dose combination. Based on these results and others from scenarios not shown here, we conclude that ignoring a practically important baseline covariate results in a lower MTD recommendation rate relative to a design accounting for this covariate and including a non-significant covariate in the model results in a slightly higher bias (still negligible) and a small reduction in percent of MTD recommendation. Therefore, we stand to lose little if we include a practically not important covariate in the model.

4. Discussion

We described Bayesian adaptive designs for cancer phase I clinical trials using two drugs with continuous dose levels in the presence of a binary baseline covariate. The goal is to estimate the MTD curve in the two-dimensional Cartesian plane for a patient's specific baseline covariate value. The methodology extends the single agent trial design with a baseline covariate and two agents design without a covariate. In each case, vague priors were used to quantify the toxicity profile of each agent *a priori*. We used an algorithm for dose escalation where cohorts of two patients are enrolled simultaneously and the patients receive different dose combinations. We studied design operating characteristics of the method under a number of practical scenarios (only two of them are included due to space limitation) by comparing this method with a design that ignores the baseline covariate and designs using parallel trials. In all simulations, we used a sample size of n = 40 patients. We found that in general, the methodology is safe in terms of the probability that a prospective trial will result in an excessively high number of DLTs when accounting for a significant covariate. We used several measures to assess the efficiency of the estimate of the MTD. In the presence of a practically significant baseline covariate, the design with a covariate had a smaller pointwise average bias and a higher percent of MTD recommendation relative to a design which ignores the covariate or when using parallel trials. When the two true MTD curves are very close, then including a baseline covariate in the model results in a slightly higher bias (still negligible) and a small reduction in percent of MTD recommendation relative to a design that ignores this covariate. Therefore, we stand to lose little if we include a practically not important covariate in the model. We plan to study the properties of this method under model misspecification and extend the approach for pre-specified discrete dose combination of the two agents in future work. Other extensions include accommodating late onset toxicity and efficacy studies.

Acknowledgment

This work is supported in part by the National Institute of Health Grant Number 1 R01CA188480-01A1 (M.T, S.K), the National Center for Research Resources, Grant UL1RR033176, and is now at the National Center for Advancing Translational Sciences, Grant UL1TR000124 (M.T), and 2 P01 CA098912 (M.T).

References

1. Frey E, III., Karon M, Levin RH, Freireich EJ, Taylor RJ, Hananian J, Selawry O, Holland JF, Hoogstraten B, Wolman IJ, Abir E, Sawitsky A, Lee S, Mills SD, Burgert EO, Spurr CL, Patterson RB, Ebaugh FG, James GWI, Moon JH. The Effectiveness of Combinations of Antileukemic Agents in Inducing and Maintaining Remission in Children with Acute Leukemia. *Blood* 1965; **26**: 642-656.
2. Thall PF, Millikan RE, Mueller P, Lee SJ. Dose-finding with two agents in phase I oncology trials. *Biometrics* 2003; **59**: 487-496.
3. Wang K, Ivanova A. Two-dimensional dose finding in discrete dose space. *Biometrics* 2005; **61**: 217-222.

4. Yin GS, Yuan Y. A Latent Contingency Table Approach to Dose Finding for Combinations of Two Agents. *Biometrics* 2009; **65**: 866-875.
5. Yin GS, Yuan Y. Bayesian dose finding in oncology for drug combinations by copula regression. *Journal of the Royal Statistical Society Series C-Applied Statistics* 2009; **58**: 211-224.
6. Braun TM, Wang SF. A Hierarchical Bayesian Design for Phase I Trials of Novel Combinations of Cancer Therapeutic Agents. *Biometrics* 2010; **66**: 805-812.
7. Wages NA, Conaway MR, O'Quigley J. Continual Reassessment Method for Partial Ordering. *Biometrics* 2011; **67**: 1555-1563.
8. Wages NA, Conaway MR, O'Quigley J. Dose-finding design for multi-drug combinations. *Clinical Trials* 2011; **8**: 380-389.
9. Sweeting MJ, Mander AP. Escalation strategies for combination therapy Phase I trials. *Pharm Stat* 2012; **11**: 258-266.
10. Shi Y, Yin G. Escalation with overdose control for phase I drug-combination trials. *Stat Med* 2013: in press.
11. Tighiouart M, Piantadosi S, Rogatko A. Dose finding with drug combinations in cancer phase I clinical trials using conditional escalation with overdose control. *Stat Med* 2014; **33**: 3815-3829.
12. Riviere MK, Yuan Y, Dubois F, Zohar S. A Bayesian dose-finding design for drug combination clinical trials based on the logistic model. *Pharm Stat* 2014; **13**: 247-257.
13. Mander AP, Sweeting MJ. A product of independent beta probabilities dose escalation design for dual-agent phase I trials. *Stat Med* 2015; **34**: 1261-1276.
14. Decoster G, Stein G, Holdener EE. Responses and Toxic Deaths in Phase-I Clinical-Trials. *Annals of Oncology* 1990; **1**: 175-181.
15. Ratain MJ, Mick R, Schilsky RL, Siegler M. Statistical and Ethical Issues in the Design and Conduct of Phase-I and Phase-Ii Clinical-Trials of New Anticancer Agents. *Journal of the National Cancer Institute* 1993; **85**: 1637-1643.
16. Ratain MJ, Mick R, Janisch L, Berezin F, Schilsky RL, Vogelzang NJ, Kut M. Individualized dosing of amonafide based on a pharmacodynamic model incorporating acetylator phenotype and gender. *Pharmacogenetics* 1996; **6**: 93-101.
17. O'Quigley J, Shen LZ, Gamst A. Two-sample continual reassessment method. *J Biopharm Stat* 1999; **9**: 17-44.
18. Babb JS, Rogatko A. Patient specific dosing in a cancer phase I clinical trial. *Stat Med* 2001; **20**: 2079-2090.
19. O'Quigley J, Paoletti X. Continual reassessment method for ordered groups. *Biometrics* 2003; **59**: 430-440.
20. Tighiouart M, Cook-Wiens G, Rogatko A. Incorporating a Patient Dichotomous Characteristic in Cancer Phase I Clinical Trials Using Escalation with Overdose Control. *Journal of Probability and Statistics* 2012; **2012**: 10.
21. Tighiouart M, Rogatko A. Dose Finding with Escalation with Overdose Control (EWOC) in Cancer Clinical Trials. *Statistical Science* 2010; **25**: 217-226.
22. Babb J, Rogatko A, Zacks S. Cancer Phase I clinical Trials: efficient dose escalation with overdose control. *Stat Med* 1998; **17**: 1103-1120.
23. Tighiouart M, Rogatko A, Babb JS. Flexible Bayesian methods for cancer phase I clinical trials. Dose escalation with overdose control. *Stat Med* 2005; **24**: 2183-2196.
24. Lunn DJ, Thomas A, Best N, Spiegelhalter D. WinBUGS - A Bayesian modelling framework: Concepts, structure, and extensibility. *Statistics and Computing* 2000; **10**: 325-337.
25. Tighiouart M, Piantadosi S, Rogatko A. Dose finding for drug combination in early cancer phase I trials using conditional escalation with overdose control. In JSM

Proceedings, Biopharmaceutical section, Alexandria, VA: American Statistical Association. 2014.