Design of Phase I/II Drug Combination Cancer Trials Using Conditional Continual Reassessment Method and Adaptive Randomization

Matthieu Clertant

Department of Mathematics and Statistics, Lancaster University, Lancaster, U.K.

E-mail: m.clertant@lancaster.ac.uk

Mourad Tighiouart

Biostatistics and Bioinformatics Research Center, Cedars-Sinai Medical Center, Los Angeles, USA.

Summary. We propose a design for a phase I/II cancer clinical trial in the context of drug combinations. The goal is to determine therapeutically efficacious dose combinations that are safe. In stage I, dose allocation is carried out according to univariate continual reassessment method. During this stage, only toxicity outcomes are used to guide dose escalation. In the second stage, toxicity and efficacy outcomes are used to update sequentially the joint model for the probability of toxicity and efficacy. At each step of the algorithm, an admissible set of safe doses is estimated using the posterior distribution of the model parameters. The next cohort of patients is allocated to dose combinations selected by an adaptive randomization in the admissible set. Since more than one dose combinations can be recommended at the end of the trial, we propose a new estimator to assess the reliability of the recommended doses. We study the performance of the method by deriving the operating characteristics with two versions of the joint toxicity and efficacy model using scenarios under model misspecification.

Keywords: Drug Combination; Dose limiting toxicity; Drug efficacy; Contingency table; CRM; Adaptive randomization.

1. Introduction

In phase I cancer clinical trials, the primary objective is to determine the maximum tolerated dose (MTD) among the doses included in the study. The MTD is defined as the dose with dose limiting toxicity (DLT) probability closest to a prespecified target θ_T. Numerous model based designs have been proposed for determining the MTD since the introduction of the continual reassessment method (CRM) (O'Quigley et al., 1990). For cytotoxic agents, the usual assumption of a complete ordering is reasonable when only one agent is allowed to vary in the study. In the case of drug combination treatment where more than one drug is allowed to vary, the monotonicity assumption on each marginal induces a partial ordering on the set of dose combinations. Several model based designs have been studied under this setting with the goal of estimating an MTD. Some of these methods recommend a single MTD (Yin and Yuan (2009a), Wages et al. (2011), Shi and Yin (2013), Riviere et al. (2014)) while other approaches propose a set

of MTDs at the end of the trial, see e.g., Thall et al. (2003); Wang and Ivanova (2005); Braun and Wang (2010); Mander and Sweeting (2015); Tighiouart et al. (2014, 2016, 2017).

For early phase studies where efficacy is ascertained in a short period of time such as one or two cycles of therapy, efficacy outcomes in addition to DLT are used to guide the dose allocation to successive cohorts of patients. The efficacy measurements may consist of tumor response or biomarker modulation if one or both drugs are biologic or immunotherpy. For single agent trials, sequential designs for modeling the joint probability of toxicity and efficacy and estimating the optimal dose have been studied extensively in the literature, see e.g. Murtaugh and Fisher (1990); Thall and Russell (1998); Braun (2002); Ivanova (2003); Thall and Cook (2004a); Chen et al. (2015); Sato et al. (2016). These methods are designed to identify one safe dose that maximizes the probability of treatment response or a prespecified utility function. The recent widespread use of combination therapy in cancer treatment using mixtures of cytotoxic, biologic, immunotherapy, and radiation therapy lead to various extensions of the previous designs to accomodate dual endpoint trials with two or more drugs allowed to vary during the trial (Yuan and Yin (2011); Wages and Conaway (2014); Cai et al. (2014), Riviere et al. (2015)). In general, these methods use a two-stage design approach where a set of tolerable dose combinations is first estimated, then patients are allocated to efficacious doses in the second stage according to the revised probabilities of DLT and treatment response. At the end of the trial, a dose combination is recommended for future randomized studies unless the trial stops early for safety or lack of efficacy.

In this manuscript, we study the performance of a two-stage design for drug combination of two cytotoxic agents where dose allocation is guided by DLT outcome in stage I and by using the outcomes DLT and efficacy in stage II. Unlike previous approaches that recommend a single dose for future studies, our method can recommend either one dose or a set of dose combinations at the end of the trial for future randomized phase II/III studies. Recommending more than one dose combination avoids the paradigm of specifying a utility function that trades the risk of toxicity with the probability of response and may lead to alternative optimal combinations that use desirable levels of one or both drugs. For instance, when treating metastatic castrate resistant prostate cancer patients with the combination cabazitaxel and cisplatin, and if two dose combinations are recommended, then the one with the lowest level of cisplatin may be selected due to the severe toxicities associated with cisplatin such as nephrotoxicity. During stage I, dose escalation/de-escalation is guided using only toxicity outcomes due to the high uncertainty about the model parameters. In the second stage, toxicity probabilities are updated sequentially. It becomes possible to control the safety of the trial while optimizing the allocation of patients according to the efficacy outcomes. The marginal model for stage I is extended by using a model of the contingency table for toxicity and efficacy outcomes. Two version are proposed: a Morgenstern model, already studied for single agent early phase dose finding studies (Thall and Cook, 2004b) and a Positive Dependent Model that assumes a positive association between the toxicity and efficacy outcomes. Such an assumption is natural in phase I cancer trials for cytotoxic agents. Since our method recommends a set of doses, we introduce a statistic that measures the reliability of the estimator for a prospective trial. This statistic can also be used for

phase I trials based on single DLT outcome whenever more than one MTD is selected at the end of the trial.

The manuscript is organized as follows. In Section 2, we introduce joint models for toxicity and efficacy outcomes based on copula models. The dose allocation algorithm is described in Section 3. Design operating characteristics are studied in Section 4 under a large number of scenarios for the marginal probabilities of toxicity and efficacy and models for generating the pairs toxicity and efficacy. Some concluding remarks are found in Section 5.

2. A Joint Model for Toxicity and Efficacy Outcomes

2.1. Marginal Models for toxicity and efficacy outcomes

Let $\mathcal{A}_1 = \{a_i, i = 1, \ldots, p\}$ and $\mathcal{B}_1 = \{b_j, i = 1, \ldots, q\}$ be the ordered dose levels of two synergistic drugs A and B, respectively. Denote by $\mathcal{D} = \mathcal{A}_1 \times \mathcal{B}_1$ the set of dose combinations available in the trial and T and E the binary indicators of toxicity and efficacy. In the remainder of the manuscript, we use the terms toxicity and DLT interchangeably. The class of models for the dose-toxicity relationship studied in Tighiouart et al. (2014, 2017) will be adopted to model the marginal probabilities of toxicity and efficacy. Specifically, for a patient treated with dose combination $(x, y) \in \mathcal{D}$, we have

$$P(T = 1|x, y, \beta_T, \eta_T) = F_T\left(\beta_{T,0} + \beta_{T,1}\, x + \beta_{T,2}\, y + \eta_T\, xy\right) \tag{1}$$

$$P(E = 1|x, y, \beta_E, \eta_E) = F_E\left(\beta_{E,0} + \beta_{E,1}\, x + \beta_{E,2}\, y + \eta_E\, xy\right), \tag{2}$$

where F_T, F_E are known cumulative distribution functions, and η_T, η_E are non-negative interaction coefficients. We will assume that the probabilities of toxicity and efficacy increase with the dose of any one of the agents when the other one is held constant. A necessary and sufficient condition for this to hold is $\beta_{T,1}, \beta_{T,2}, \beta_{E,1}, \beta_{E,2}$ are all positive. Suppose that the dose levels of the two drugs are standardized to be in the interval $[0, 1]$, $a_1 = b_1 = 0, a_p = b_q = 1$. Building on the work of Tighiouart et al. (2014, 2017), we reparameterize model (1) in terms of the probabilities of toxicity at dose combinations $(0, 0), (0, 1), (1, 0)$, denoted by $\rho_{T,00}$, $\rho_{T,01}$ and $\rho_{T,10}$, respectively. These parameters can be easily interpreted by the clinicians and they facilitate prior distributions specification. We then have

$$\begin{cases} \beta_{T,0} &= F_T^{-1}(\rho_{T,00}) \\ \beta_{T,1} &= F_T^{-1}(\rho_{T,10}) - F_T^{-1}(\rho_{T,00}) \\ \beta_{T,2} &= F_T^{-1}(\rho_{T,01}) - F_T^{-1}(\rho_{T,00}) \end{cases} \tag{3}$$

The marginal model for efficacy (2) is similarly reparameterized in terms of the probabilities of efficacy $\rho_{E,00}, \rho_{E,01}, \rho_{E,10}$ at dose combinations$(0, 0), (0, 1), (1, 0)$, respectively. Denotes by $\rho_T = (\rho_{T,00}, \rho_{T,01}, \rho_{T,10}), \rho_E = (\rho_{E,00}, \rho_{E,01}, \rho_{E,10}), \rho = (\rho_T, \rho_E)$, and $\eta = (\eta_T, \eta_E)$.

Table 1. Contingency Table for Toxicity/Efficacy Outcomes

	$1-P_T$	P_T
$1-P_E$	Π_{00}	Π_{10}
P_E	Π_{01}	Π_{11}

2.2. Joint toxicity and efficacy models

Denote by Π_{ik} the joint probability of toxicity and efficacy for a patient treated with dose combination (x, y), $i = 0, 1, k = 0, 1$ and $P_T = P(T = 1|(x, y))$ and $P_E = P(E = 1|(x, y))$. In this section, we omit the dose combination notation (x, y) and parameters of the marginals (ρ, η) to facilitate the description of the joint model for toxicity and efficacy with the understanding that the two marginals are conditional on the dose and parameters of models (1) and (2).

There exist several models for bivariate binary outcomes (Joe, 1997). In particular, copula models link the marginal distributions of toxicity and efficacy in a 2×2 contingency table (1) to form the joint distribution, see e.g, (Nelsen, 1999). For single agent phase I/II trials, the Morgenstern or Gumbel copula model was used in (Thall and Cook, 2004b) to model the joint distribution of toxicity and efficacy. This model was also used in dose combination phase I trials in (Yin and Yuan, 2009b) by collapsing the probabilities of DLT attributed to either one or both drugs. Let $f(\gamma) = (e^\gamma - 1)/(e^\gamma + 1), \gamma \in \mathbb{R}$. The Morgenstern model is defined by

$$\begin{cases} \Pi_{11} &= P_T P_E \times \{1 + (1 - P_T)(1 - P_E) \times f(\gamma)\} \\ \Pi_{10} &= P_T(1 - P_E) \times \{1 - P_T(1 - P_E) \times f(\gamma)\} \\ \Pi_{01} &= (1 - P_T)P_E \times \{1 - (1 - P_T)P_E \times f(\gamma)\} \\ \Pi_{00} &= 1 - (\Pi_{11} + \Pi_{10} + \Pi_{01}) \end{cases} \tag{4}$$

This is a flexible model since it accommodates the case of independence between toxicity and efficacy outcomes when $\gamma = 0$, positive association between these two outcomes for $\gamma > 0$, and negative association when $\gamma < 0$. We note that the Morgenstern model can be seen as a deviation from the independent assumption between efficacy and toxicity. For fixed γ, the probability of the joint event $\{T = i, E = k\}$ obtained in the independent case is multiplied by a linear function of the probability of the event $\{T = 1-i, E = 1-k\}$ under the independent case. However the morgenstern model can induce a tight range for the joint probability at certain dose combinations. For example, if the toxicity probability is $P_T = 0.25$ and the efficacy probability is $P_E = 0.8$, then the probability Π_{01} varies in the interval $[0.57, 0.63]$ as γ varies in $(-\infty, \infty)$. For such marginals, the probability of the joint event Π_{01} "observing an efficacy outcome without toxicity" can in theory belong to $[0.55, 0.75]$. Indeed, for any marginals P_E and P_T, the probability Π_{11} is bounded by the Fréchet-Hoeffding bounds $m = \max(0, P_E + P_T - 1) \leq \Pi_{11} \leq M = \min(P_E, P_T)$, see (Nelsen, 1999). Since for many cytotoxic agents, there exists a positive correlation between efficacy and toxicity, we modify model (4) by restricting γ to be in the interval $[0, 1]$ so that Π_{11} achieves the Fréchet-Hoeffding lower and upper bounds as follows. A linear transformation of $[P_T \times P_E, M]$ into the interval $[0, 1]$ is made by identifying the value 0 with the independent case where the probability Π_{11} is equal to $P_E \times P_T$. Let

$\gamma \in [0, 1]$. The positive dependent model (PDM) is defined by

$$\begin{cases} \Pi_{11} &= P_T P_E + (M - P_T P_E)\,\gamma \\ \Pi_{10} &= P_T(1 - P_E) - (M - P_T P_E)\,\gamma \\ \Pi_{01} &= (1 - P_T)P_E - (M - P_T P_E)\,\gamma \\ \Pi_{00} &= 1 - (\Pi_{11} + \Pi_{10} + \Pi_{01}) \end{cases} \tag{5}$$

The PDM should be used if there is *a priori* belief that toxicity and efficacy events are positively correlated, as is the case for many cytotoxic agents. In the simulation studies, we will evaluate the robustness of the model when the data are generated from a model with a negative association between toxicity and efficacy.

Let θ_T, θ_E be the target probabilities of toxicity and efficacy, respectively. The goal of the phase I/II trial is to determine dose combination $(x, y) \in \mathcal{G}$ that satisfy $P(T = 1|(x, y)) \leq \theta_T + \epsilon_T$ and $P(E = 1|(x, y)) > \theta_E$ where $\epsilon_T \geq 0$ is a tolerance parameter pre-specified in collaboration with the clinician.

3. Algorithm

At the start of the trial, the uncertainty about the nine model parameters is high. We therefore propose a two-stage design where in stage 1, N_1 patients are allocated to dose combinations according to univariate CRM and the algorithm described in (Tighiouart et al., 2017) using toxicity data only $(x_i, y_i, T_i), i = 1, \ldots, N_1$. In the second stage, successive cohorts of m patients are allocated to "safe" dose combinations that are likely to have high probability of efficacy using toxicity and efficacy data (x_i, y_i, T_i, E_i) and adaptive randomization.

3.1. Stage I

Let Λ^T be a prior distribution for the parameters (ρ_T, η_T). Using Bayes rule, the posterior distribution given $D_{T,n} = \{(x_j, y_j, T_j), j = 1, \ldots, n\}$ is

$$\Lambda_n(\rho_T, \eta_T|D_{T,n}) \propto \prod_{j=1}^{n} L(T_j|x_j, y_j, \rho_T, \eta_T) \times \Lambda^T(\rho_T, \eta_T), n = 1, \ldots,\ N_1, \tag{6}$$

where

$$L(T_j|x_j, y_j, \rho_T, \eta_T) = \sum_{i=0}^{1} P(T_j = i|x_j, y_j, \rho_T, \eta_T) \times I_{\{T_j=i\}} \tag{7}$$

is the likelihood of a single observation (x_j, y_j, T_j). During this phase, cohorts of two patients are treated simultaneously and only moves in one direction of the range of dose combinations are allowed. Thus, a patient in the current cohort can be treated at a dose (x, y) if and only if a patient in the previous cohort was treated at a dose on the same horizontal or vertical line in our dose range: $(x^{'}, y)$ with $x^{'} \in \mathcal{A}_1$ or $(x, y^{'})$ with $y^{'} \in \mathcal{B}_1$. An estimator working on vertical or horizontal line can be used to allocate the next cohort of patients. Let $\hat{X}$ and $\hat{Y}$ be plug in estimators as described for the CRM

((O'Quigley et al., 1990)). The posterior estimators of the parameters are $(\hat{\rho}_n^T, \hat{\eta}_n^T) = E_{\Lambda^T}[(\rho_T, \eta_T)|D_{T,n}]$. We then have

$$\hat{X}_n(y) = \underset{x\in\mathcal{A}_1}{\arg\min}|H_T(x, y, \hat{\rho}_n^T, \hat{\eta}_n^T) - \theta_T| \text{ and } \hat{Y}_n(x) = \underset{y\in\mathcal{B}_1}{\arg\min}|H_T(x, y, \hat{\rho}_n^T, \hat{\eta}_n^T) - \theta_T|\,, \quad (8)$$

where the function H_T is given by

$$\begin{aligned} H_T(x, y, \hat{\rho}_T, \hat{\eta}_T) = F_T[F_T^{-1}(\hat{\rho}_{T,00}) + \left\{F_T^{-1}(\hat{\rho}_{T,10}) - F_T^{-1}(\hat{\rho}_{T,00})\right\} \times x \\ + \left\{F_T^{-1}(\hat{\rho}_{T,01}) - F_T^{-1}(\hat{\rho}_{T,00})\right\} \times y + \hat{\eta}_T \times xy. \end{aligned}$$

CRM algorithm:

- In the first cohort, the two patients receive the minimum dose combination $(0, 0)$
- In the $k-$cohort of two patients, if i is even,
  $(x_{2k+1}, y_{2k+1}) = (\hat{X}_{2(k-1)+1}(y_{2(k-1)+1}), y_{2(k-1)+1})$ and $(x_{2k}, y_{2k}) = (x_{2(k-1)}, \hat{Y}_{2(k-1)+1}(x_{2(k-1)}))\,,$
  and if k is odd,
  $(x_{2k+1}, y_{2k+1}) = (x_{2(k-1)+1}, \hat{Y}_{2(k-1)+1}(x_{2(k-1)+1}))$ and $(x_{2k}, y_{2k}) = (\hat{X}_{2(k-1)+1}(y_{2(k-1)}), y_{2(k-1)})\,.$

Since the number of dose levels of each agent is not very large in our simulation studies, $p = 6$ and $q = 5$, dose skipping is not allowed. However, for larger number of dose levels or when using continuous dose levels, the maximum size of the jump can be set as a function of the dose range of either agent as in Tighiouart et al. (2017).

3.2. Stage II

Let $\Lambda(\rho, \eta, \gamma) = \Lambda^T(\rho_T, \eta_T) \times \Lambda^E(\rho_E, \eta_E) \times \Lambda^C(\gamma)$ be a prior distribution for the parameters (ρ, η, γ). The posterior distribution given $D_n = \{(x_j, y_j, T_j, E_j), j = N_1, N_1 + 1, \ldots, N\}$ is

$$\Lambda_n(\rho, \eta, \gamma) \propto \prod_{j=1}^{n} L(T_j, E_j|x_j, y_j, \rho, \eta, \gamma) \times \Lambda(\rho, \eta, \gamma), n = N_1, \ldots,\ N, \quad (9)$$

where

$$L(T_j, E_j|x_j, y_j, \rho, \eta, \gamma) = \sum_{i=0}^{1}\sum_{k=0}^{1} \Pi_{ik|x_j,y_j,\rho,\eta,\gamma} \times I_{\{E=i\}} \times I_{\{T=k\}} \quad (10)$$

is the likelihood of a single observation (x_j, y_j, T_j, E_j), $\Pi_{ik|x_j,y_j,\rho,\eta,\gamma}$ is given by (4) or (5), and N is the total sample size from both stages. At the end of stage I, a set of admissible doses $\mathcal{A}_{N_1}$ is determined

$$\mathcal{A}_{N_1} = \{(x, y) \in \mathcal{D} : \theta_T - \epsilon_L < \hat{P}_{N_1}(T = 1|(x, y)) < \theta_T + \epsilon_T\}, \quad (11)$$

where ϵ_L is a design parameter and

$$\hat{P}_{N_1}(T = 1|(x, y)) = E_{\Lambda_{N_1}}[P(T = 1|x, y, \rho, \eta, \gamma)]. \quad (12)$$

For each dose combination $(x, y) \in \mathcal{A}_{N_1}$, we estimate the probability of efficacy

$$\hat{P}_{N_1}(E = 1|(x, y)) = E_{\Lambda_{N_1}}[P(E = 1|x, y, \rho, \eta, \gamma)] \tag{13}$$

and every patient in a cohort of m patients is allocated to dose combination $(x^*, y^*) \in \mathcal{A}_{N_1}$ with probability

$$\pi_{N_1}(x^*, y^*) = \frac{\hat{P}_{N_1}(E = 1|(x^*, y^*))}{\sum_{(x,y)\in\mathcal{A}_{N_1}} \hat{P}_{N_1}(E=1|(x,y))}. \tag{14}$$

The set of admissible doses $\mathcal{A}_{N_1}$ in (11) consists of dose combinations with estimated probability of toxicity in a neighborhood of the target probability of toxicity θ_T. Patients in the first cohort of m patients at the start of stage II are allocated to doses in $\mathcal{A}_{N_1}$ using adaptive randomization according to the probability distribution $\pi_{N_1}(\cdot)$ defined in (14) so that these patients are likely to be allocated to dose combinations that are likely to have high probability of efficacy. After the toxicity and efficacy status of these patients are resolved, the posterior distribution of the model parameters is updated using (9) to Λ_{N_1+m} and the set of admissible doses is updated to $\mathcal{A}_{N_1+m}$. The next cohort of m patients are allocated to dose combinations in $\mathcal{A}_{N_1+m}$ using adaptive randomization with updated probability distribution $\pi_{N_1+m}(\cdot)$ obtained as in (14). The algorithm continues this way enrolling successive cohorts of m patients until a total of N patients are enrolled to the trial subject to one of the stopping rules defined in Section 3.3.

In a real trial, new patients may be available for enrollment while the toxicity or efficacy outcomes of some patients are not resolved. We note that our model works under this circumstance by adjusting the likelihood. Denote by Z a missing data. The likelihood for a single observation (T, E, x, y) is then:

$$L(T, E|x, y, \rho, \eta, \gamma) = \sum_{i=0}^{1}\sum_{k=0}^{1} \Pi_{ik|x,y,\rho,\eta,\gamma} \times \left(I_{\{T=i\}} + I_{\{T=Z\}}\right) \times \left(I_{\{E=k\}} + I_{\{E=Z\}}\right). \tag{15}$$

3.3. Stopping rules

For both phases of the study, stopping rules based on Bayesian tests at each dose combination are introduced. Let $P_{T,(x,y)}$ and $P_{E,(x,y)}$ be the toxicity and efficacy probabilities at a dose (x, y) and U a uniform prior in $[0, 1]$ on these probabilities. Let $n_{(x,y)}$ and $n^1_{(x,y)}$ the number of patients treated at dose (x, y) and the number of patient having experienced a DLT at dose (x, y), respectively. During stage I, the trial is stopped if the minimum dose combination $(0, 0)$ is too toxic:

Stopping rule for Stage I: We stop enrollment to the trial after n patients if:

$$P_U\left(P_{T,(0,0)} > \theta_T + \epsilon_T \,\middle|\, (n_{(0,0)}, n^1_{(0,0)})\right) > \delta_T. \tag{16}$$

The stopping rule for stage II is more complex since it deals with two objectives: localizing both the set $\mathcal{H}$ of dose combinations that are too toxic and the set $\mathcal{L}$ of dose combinations that are less efficacious define by

$$\mathcal{H} = \{(x,y) \in \mathcal{D} : P_{T,(x,y)} > \theta_T + \epsilon_T\} \text{ and } \mathcal{L} = \{(x,y) \in \mathcal{D} : P_{E,(x,y)} < \theta_E\}. \quad (17)$$

For $n = N_1, N_1+1, \ldots, N-1$, let $\tilde{\mathcal{H}}_n$ and $\tilde{\mathcal{L}}_n$ be the estimated sets of doses that are too toxic and less efficacious, respectively:

$$\tilde{\mathcal{H}}_n = \{(x,y) \in \mathcal{D} : P_U\left(P_{T,(x,y)} > \theta_T + \epsilon_T \,\middle|\, (n_{(x,y)}, n^1_{(x,y)})\right) > \delta_T, \quad (18)$$

$$\tilde{\mathcal{L}}_n = \{(x,y) \in \mathcal{D} : P_U\left(P_{E,(x,y)} < \theta_E \,\middle|\, (n_{(x,y)}, n^1_{(x,y)})\right) > \delta_E. \quad (19)$$

The monotonicity assumptions for the probabilities of toxicity and efficacy described in Section 2.1 imply that for any dose $(x,y) \in \tilde{\mathcal{H}}_n$, then $(x',y') \in \tilde{\mathcal{H}}_n$ whenever $x' \geq x$ and $y' \geq y$. Similarly, for any dose $(x,y) \in \tilde{\mathcal{L}}_n$, then $(x',y') \in \tilde{\mathcal{L}}_n$ whenever $x' \leq x$ and $y' \leq y$. The idea here is to stop the trial if there are no "good" doses in $\mathcal{D}$, i.e, doses that have an acceptable level of toxicity with a desirable level of efficacy.
Stopping rule for Stage II: We stop enrollment to the trial after n patients if:

$$\mathcal{D} = \{(x,y) \in \mathcal{D} : \exists\,(x',y') \in \tilde{\mathcal{H}}_n,\, x > x' \text{ and } y > y'\} \cup \{(x,y) \in \mathcal{D} : \exists\,(x',y') \in \tilde{\mathcal{L}}_n,\, x < x' \text{ and } y < y'\}, \quad (20)$$

which means that there exists no more "good" doses in the set of available dose combinations in the trial $\mathcal{D}$, based on the estimated sets $\tilde{\mathcal{H}}_n$ and $\tilde{\mathcal{L}}_n$.

The design parameters δ_T, δ_E in (16), (18), and (19) are selected to achieve good operative characteristics.

3.4. Final selection

At the end of the trial, we recommend a set of doses or a single dose to be used in future randomized Phase II or III studies. Our goal is to estimate the set $\mathcal{G}$ of good doses defined by $\mathcal{G} = \mathcal{D} \setminus (\mathcal{H} \cup \mathcal{L})$, see (17). An estimator of the set $\mathcal{G}$ is

$$\hat{\mathcal{G}}_N = \mathcal{D} \setminus \Big[\{(x,y) \in \mathcal{D} : \hat{P}_N(T=1|(x,y)) > \theta_T + \epsilon_T\} \cup \{(x,y) \in \mathcal{D} : \hat{P}_N(E=1|(x,y)) < \theta_E\}\Big], \quad (21)$$

where $\hat{P}_N(T=1|(x,y))$ and $\hat{P}_N(E=1|(x,y))$ are given by (12) and (13), respectively. Thus, we exclude doses with low probability of efficacy and high probability of toxicity.

If the goal of the trial is to recommend a single dose, we can use the $0-1$ loss function l and determine the optimal decision rule $\Delta^\star_N$ selecting one dose in $\hat{\mathcal{G}}_N$. Let

$$l(\rho,\eta,\gamma,\Delta_N) = \begin{cases} 1\,, & \text{if } P(E=1|(x,y) = \Delta_N, \rho, \eta, \gamma) < \theta_E \\ 0\,, & \text{otherwise,} \end{cases} \quad (22)$$

the decision rule in $\hat{\mathcal{G}}_N$ minimizing the risk function $R(\rho,\eta,\gamma,\Delta_N) = E_{\Lambda_N}\left[l(\rho,\eta,\gamma,\Delta_N)\right]$ is:

$$\Delta_N^\star = \underset{(x,y)\in\hat{\mathcal{G}}_N}{\arg\max}\, P_{\Lambda_N}\left[P(E=1|(x,y),\rho,\eta,\gamma) > \theta_E\right]. \tag{23}$$

The estimator $\Delta_N^\star$ maximizes the probability to select the most efficaciousious dose in $\hat{\mathcal{G}}_N$. Properties of this estimator will be studied in the next section. Other loss functions can be chosen in collaboration with the clinicians.

4. Simulation Studies

4.1. Simulation set-up and scenarios

We assume that drug A has $p = 6$ levels and drug B has $q = 5$ levels. The trial sample size is $N = 60$ patients with $N_1 = 20$ enrolled in stage I. Stage II enrolls consecutive cohorts of $m = 5$ patients. The target probability of toxicity is $\theta_T = 0.33$ and the target probability of efficacy is $\theta_E = 0.5$. The link functions for the marginal models are the logistic function $F_T(u) = F_E(u) = (1 + e^{-u})^{-1}$.

We take vague priors on the model parameters (ρ, η) by assuming that $\rho_{T,01}, \rho_{T,10}$ are i.i.d $\sim U(0, 1)$, and conditional on $(\rho_{T,01}, \rho_{T,10})$, $\rho_{T,00}/\min(\rho_{T,01}, \rho_{T,10}) \sim U(0, 1)$ and η_T has gamma distribution with mean 20 and variance 500 as in Tighiouart et al. (2017). The same prior is placed on the marginal model parameters for efficacy (ρ_E, η_E). The prior distribution for the association parameter γ is taken as a normal distribution with mean 0 and variance 100 when the working model is the Morgenstern model (4) and a uniform distribution in $[0, 1]$ when the working model is the PDM given in (5). The tolerance parameter is fixed at $\epsilon_T = 0.1$. During stage II, the set of admissible doses $\mathcal{A}_n$ may be empty if the design parameter ϵ_L is too small. We therefore start with $\epsilon_L = 0.1$ and increase it by increments of 0.05 until $\mathcal{A}_n$ contains at least one dose. The design parameters δ_T and δ_E for the stopping rules are $\delta_T = 0.7, \delta_E = 0.9$. A sensitivity analysis on these parameters is included in Table 7.

We considered nine scenarios for the marginal probabilities of toxicity and efficacy shown in Tables 2 and 3. For each scenario of the marginals, the pairs (T, E) are generated using the Morgenstern model with $\gamma = -2, 0, 2$ and the PDM model with $\gamma = 0, 0.9$. This gives us a total of $9 \times 4 = 36$ scenarios for the true model. These scenarios reflect various number and location of "good" dose combinations, e.g, near the minimum dose combination for scenario IX, middle of the dose range for scenario III-V, towards the highest dose combination for scenario VI, and cases of no set of good doses for scenarios VII and VIII. Only scenario IV is fully generated under the logistic model. The true efficacy probabilities for scenario I and II are also generated under the logistic model. All the other scenarios are misspecified in order to assess the robustness of our model. For each scenario, $M = 2000$ trial replicates are generates to evaluate the performance of the models and dose allocation algorithm.

4.2. Operating characteristics

For each scenario, we report several statistics to evaluate safety and efficiency of the methodology.

- TOX is the average % of DLTs observed across the M trials.

Table 2. Scenarios for the marginal probabilities P_T and P_E.

Scenario		Toxicity						Efficacy					
		1	2	3	4	5	6	1	2	3	4	5	6
I	1	01	03	07	13	17	38	01	02	03	06	11	19
	2	02	05	09	24	36	47	04	08	15	27	42	59
	3	08	19	35	45	50	60	17	31	48	67	81	90
	4	**18**	**36**	44	53	59	68	**52**	**68**	83	92	96	98
	5	**27**	45	50	62	71	80	**81**	91	96	98	99	1.00
II	1	01	03	07	13	**17**	**38**	05	11	23	41	**62**	**79**
	2	02	05	09	**24**	**36**	47	07	16	33	**55**	**75**	89
	3	08	19	35	45	50	60	10	23	44	68	85	94
	4	18	36	44	53	59	68	14	31	57	79	92	97
	5	27	45	50	62	71	80	19	41	69	87	95	98
III	1	01	02	05	19	28	49	01	03	08	12	15	28
	2	04	08	**21**	**35**	46	58	05	25	**55**	**75**	85	90
	3	14	25	**37**	51	68	75	09	32	**68**	85	87	95
	4	23	36	44	53	70	79	15	37	73	86	92	96
	5	31	50	50	59	79	92	21	40	75	90	93	98
IV	1	02	03	06	09	**15**	**24**	05	11	23	41	**62**	**79**
	2	04	07	11	**18**	**28**	**42**	07	16	33	**55**	**75**	**89**
	3	07	12	20	**32**	47	62	10	23	44	**68**	85	94
	4	13	22	**35**	50	66	79	14	31	**57**	79	92	97
	5	22	36	52	68	81	89	19	41	69	87	95	98
V	1	02	03	05	10	16	20	01	02	04	06	11	20
	2	03	04	07	**12**	**19**	**25**	02	05	24	**55**	**70**	**75**
	3	05	12	**32**	47	63	70	04	32	**78**	85	90	92
	4	07	**20**	**40**	55	70	85	09	**55**	**82**	86	92	93
	5	10	**29**	47	60	75	90	13	**58**	84	89	93	94
VI	1	01	02	04	05	09	11	02	13	16	17	18	22
	2	02	05	06	07	14	23	06	16	17	19	31	35
	3	04	06	08	11	23	33	08	15	23	28	40	44
	4	05	08	13	24	**35**	**41**	11	20	30	34	**55**	**62**
	5	06	14	22	**30**	**38**	53	15	23	37	**52**	**65**	72

- EFF is the average % of observed responses across the M trials.
- $A_{\mathcal{G}}$, $A_{\mathcal{L}}$, $A_{\mathcal{H}}$ are the average % of patients allocated to doses in sets $\mathcal{G}, \mathcal{L}$ and $\mathcal{H}$ across the M trials, respectively.

When more than one dose combination is recommended at the end of the trial, the percent selection of each dose combination is a useful summary of trial efficiency but these percentages are hard to interpret. For example, Table 4 shows the percent selection of each dose combination with levels $(1,4),(1,5),(2,4)$ in the true set of good doses. These percentages are $32\%, 78\%$, and 70%, respectively. While two of these percentages are reasonably high, they do not reflect the reliability of the estimate $\hat{\mathcal{G}}_N$ of the set of good doses $\mathcal{G}$ given in (21). We therefore introduce a statistic that measures the reliability of an estimated set of doses as follows. For any set $\mathcal{E} \subset \mathcal{G}$, let

$$S_{\mathcal{E}} = \frac{1}{\sum_{j=1}^{M} |\hat{\mathcal{G}}_{N,j}|} \times \sum_{j=1}^{M} |\mathcal{E} \cap \hat{\mathcal{G}}_{N,j}|, \text{ and } S^1_{\mathcal{E}} = \frac{1}{M} \times \sum_{j=1}^{M} I\{\Delta^\star_N \in \mathcal{E}\}, \tag{24}$$

where $\hat{\mathcal{G}}_{N,j}$ is the estimated set of "good" doses from the j^{th} trial and $|\mathcal{T}|$ denotes the cardinal of a set $\mathcal{T}$. The following statistics are used to summarize trial efficiency:

- $S_{\mathcal{G}}$, $S_{\mathcal{L}}$, $S_{\mathcal{H}}$ are the percent selection in sets $\mathcal{G}, \mathcal{L}$ and $\mathcal{H}$, respectively, when the set of doses $\hat{\mathcal{G}}_N$ is recommended at the end of the trial.
- $S^1_{\mathcal{G}}$, $S^1_{\mathcal{L}}$, $S^1_{\mathcal{H}}$ are the percent selection in sets $\mathcal{G}, \mathcal{L}$ and $\mathcal{H}$, respectively, when a single dose is selected with the decision rule $\Delta^\star_N$ at the end of the trial.
- NS is the average of number of dose selected in $\hat{\mathcal{G}}_N$.

Note that $S_{\mathcal{G}}$ can be interpreted as an estimate of the probability that a dose in $\hat{\mathcal{G}}_N$ belongs to G, the set of good doses. Indeed, the expression $S_{\mathcal{G}}$ in (24) is a weighted probability that any dose in $\hat{\mathcal{G}}_N$ is a good dose with weights given by $|\hat{\mathcal{G}}_{N,i}| / \sum_{j=1}^{M} |\hat{\mathcal{G}}_{N,j}|$. Finally, two additional statistics are reported to study the properties of the stopping rules:

- ET is the percent of trials that are stopped early.
- N_{et} is the average of number of patients enrolled when the trial stops.

4.3. Results

Table 4 shows the percent selection and allocation of patients to each dose combination under scenario 1 and $\gamma = 0$, that is assuming independence between toxicity and efficacy when the working model is the PDM. We can see that the percent of patients allocated to too toxic doses is small whereas the two good dose combinations with levels $(1,5),(2,4)$ are selected with high probability. However, since the final selection uses $\hat{\mathcal{G}}_N$, it is hard to interpret these percentages when recommending a group of doses to the clinician as discussed in Section 4.2. Percent selection and allocation for the other scenarios are not shown.

Table 3. Scenarios for the marginal probabilities P_T and P_E.

Scenario		Toxicity						Efficacy					
		1	2	3	4	5	6	1	2	3	4	5	6
VII	1	05	10	19	35	45	55	01	05	08	15	22	42
	2	08	18	28	42	51	70	02	07	10	19	23	45
	3	12	22	35	45	53	72	04	10	13	20	28	50
	4	20	34	45	55	65	75	08	15	18	24	37	56
	5	28	37	47	57	70	82	14	22	34	45	55	70
VIII	1	30	35	41	60	68	72	01	03	08	13	17	29
	2	42	55	57	63	72	75	05	10	12	15	20	35
	3	53	60	70	75	80	83	08	15	16	20	25	40
	4	65	67	74	77	85	90	14	22	24	28	32	44
	5	71	79	81	85	90	95	23	25	31	36	43	50
IX	1	30	**40**	55	65	68	74	32	**55**	60	62	64	70
	2	**43**	53	57	66	72	75	**75**	85	87	89	91	93
	3	57	62	72	74	81	83	80	87	90	91	92	93
	4	65	69	74	77	85	92	81	88	91	92	93	95
	5	71	79	81	85	91	98	90	92	95	96	98	99

Table 4. Percent selection and allocation.

Scenario		Toxicity						Efficacy					
		1	2	3	4	5	6	1	2	3	4	5	6
I	1	01	03	07	13	17	38	01	02	03	06	11	19
	2	02	05	09	24	36	47	04	08	15	27	42	59
	3	08	19	35	45	50	60	17	31	48	67	81	90
	4	**18**	**36**	44	53	59	68	**52**	**68**	83	92	96	98
	5	**27**	45	50	62	71	80	**81**	91	96	98	99	1.00
		Percent allocation						Percent selection using $\hat{\mathcal{G}}_N$					
		1	2	3	4	5	6	1	2	3	4	5	6
PDM	1	3	2	0	0	2	4	0	0	0	0	0	0
	2	2	4	6	8	5	2	0	0	0	1	12	12
	3	0	9	15	4	1	0	0	3	42	17	3	1
	4	**3**	**11**	4	1	0	0	**32**	**70**	15	1	0	0
	5	**7**	5	1	0	0	0	**78**	30	3	0	0	0

Operating characteristics under scenarios I–VI with $\gamma = 0$ for the true model are presented in Table 5. The summary statistics are conditional on the event that the trial did not stop early. The average percent of DLTs is below the target $\theta_T = 0.33$ for all scenarios. Similarly, the average percent of efficacy is around the target ($\theta_E = 50\%$) except for the scenario VI where the good dose combinations are near the highest dose combinations. The mean value of EFF across scenarios I to V is 48%. The average percent of patients allocated to highly toxic doses (doses with probability more than $\theta_T + 0.1$) never exceeds 19%. The percent selections when recommending either one or more than one dose are reasonably high for phase I/II trials. In particular, the estimator $\hat{\mathcal{G}}_N$ that recommends a set of dose combinations achieves 68% recommendation on the average for the six scenarios when the working model is Morgenstern. For scenarios IV and V, the percent selection is 80% or more. This is not very surprising since the true set of good doses contains 7 dose combinations. The fact that scenario IV was generated under the true model but scenario V is completely misspecified does not seem to affect the performance of the method. We also note that the summary statistics obtained using the two working models Morgenstern and PDM when toxicity and efficacy are generated independently are very similar with the largest difference of $S_{\mathcal{G}} = 3\%$ obtained under scenario VI.

In Table 6, we present the operating characteristics when toxicity and efficacy are generated from the Morgenstern model $\gamma_{\text{Morg}} = -2, 0, 2$ and the PDM model $\gamma_{\text{PDM}} = 0.9$. These statistics are presented for scenarios III and IV and averaged across the six scenarios I–VI denoted by AV$(\cdot)$. The performance of these models are quite close when the pair (T, E) is generated using the Morgenstern model with $\gamma_{\text{Morg}} = 0, 2$. When $\gamma_{\text{Morg}} = -2$ the PDM has a slightly lower percent selection relative to the Morgenstern model on the average. In particular, under scenario III the percent selection of the PDM

Table 5. Operating characteristics when $\gamma = 0$.

Model	Scenario	TOX	EFF	$A_{\mathcal{G}}$	$A_{\mathcal{H}}$	$A_{\mathcal{L}}$	$S^1_{\mathcal{G}}$	$S^1_{\mathcal{H}}$	$S^1_{\mathcal{L}}$	$S_{\mathcal{G}}$	$S_{\mathcal{H}}$	$S_{\mathcal{L}}$	NS
Morg.	I	28.5	46.9	22	18	58	61	33	5	57.1	24.8	18.2	3
	II	28.1	46.5	31	16	56	73	24	2	72.4	17.7	12	3.3
	III	30.1	43.8	34	18	55	42	32	37	47.8	33	36.5	2
	IV	25.6	52.2	47	9	45	83	15	1	81.2	7.3	11.5	5.5
	V	24.1	55.5	59	11	28	72	26	1	83	13.3	3.7	5.1
	VI	21.1	36	25	2	74	61	6	32	64.2	2.7	33.2	2.3
PDM.	I	28.3	46.2	21	18	60	59	36	4	56.2	25.6	18.1	3.2
	II	28.2	46.7	31	17	56	74	24	2	71.7	17.3	13	3.5
	III	30.1	44.2	34	19	55	42	33	35	48.6	32	36.9	2.2
	IV	25.9	52.6	47	9	43	81	18	1	80.4	7.6	11.9	5.8
	V	24.4	56	59	12	28	72	27	0	82.1	13.3	4.6	5.4
	VI	21	35.9	25	2	74	62	6	31	61.2	2.4	36.3	2.5

Table 6. Operating characteristics for selected values of γ and selected scenarios.

Scenario	Model	TOX	EFF	$A_{\mathcal{G}}$	$A_{\mathcal{H}}$	$A_{\mathcal{L}}$	$S^1_{\mathcal{G}}$	$S^1_{\mathcal{H}}$	$S^1_{\mathcal{L}}$	$S_{\mathcal{G}}$	$S_{\mathcal{H}}$	$S_{\mathcal{L}}$	NS
AV ($\gamma_{\text{Morg}} = 0$)	Morg.	26.2	46.8	36.3	12.3	52.7	65.3	22.7	13	67.6	16.5	19.2	3.5
	Pos.	26.3	46.9	36.2	12.8	52.7	65	24	12.2	66.7	16.4	20.2	3.8
AV ($\gamma_{\text{Morg}} = 2$)	Morg.	26.3	46.8	36.3	12.5	52.5	65.7	22.3	12.7	67.6	16.4	18.9	3.5
	Pos.	26.2	46.6	35.3	12.3	53	65.7	22	13	66.4	16.1	20.3	3.6
AV ($\gamma_{\text{Morg}} = -2$)	Morg.	26.3	46.6	36.3	12.3	52.7	65.5	22.5	13	67.2	16.6	19.4	3.5
	Pos.	26.4	46.8	35.8	12.7	52.8	63.8	24.2	14.3	65.7	16.8	21.0	3.8
III ($\gamma_{\text{Morg}} = -2$)	Morg.	30.0	43.8	33.0	19.0	56.0	43.0	31.0	37	47.4	33.6	36.5	2.1
	Pos.	30.1	43.8	34.0	19.0	55.0	38.0	36.0	41.0	46.5	33.3	38.6	2.3
AV ($\gamma_{\text{PDM}} = 0.9$)	Morg.	26.1	46.5	35.2	12.3	53.8	66.8	22.2	10.3	68.6	16.4	17.0	3.5
	Pos.	26.1	46.7	35.8	11.8	52.8	68.3	21.7	9.3	69.8	15.0	17.1	3.6
III ($\gamma_{\text{PDM}} = 0.9$)	Morg.	29.7	43.7	34.0	16.0	56.0	49.0	26.0	28.0	53.1	30.4	27.8	1.9
	Pos.	30.0	44.6	34.0	17.0	54.0	57.0	23.0	23.0	55.6	27.8	27.3	2.0
IV ($\gamma_{\text{PDM}} = 0.9$))	Morg.	25.5	51.7	45.0	9.0	46.0	84.0	15.0	0.0	81.4	8.2	10.4	5.5
	Pos.	25.5	51.8	46.0	9.0	45.0	87.0	14.0	0.0	83.1	7.1	9.8	5.6

when recommending a single dose at the end of the trial $S^1_{\mathcal{G}}$ is 38% compared to 43% for the Morgenstern model. This is not surprising since the PDM assumes nonnegative association between toxicity and efficacy. On the other hand, when (T, E) is generated using the PDM with $\gamma_{\text{PDM}} = 0.9$, the percent selection using the PDM as the working model is slightly better than the Morgenstern model on the average but is substantially higher under scenario III when recommending a single dose, 57% compared to 49%. This shows that the PDM can be an improvement over the Morgenstern model in capturing strong association between toxicity and efficacy.

The performance of the stopping rules described in (16) and (20) are studied under desirable scenarios II, IV, and VI where the set of good doses is located somewhere in the middle of the dose combination range and under extreme scenarios where $\mathcal{G} = \emptyset$ (scenarios VII and VIII) and scenario IX where two good doses are near the minimum dose combination, see Table 3. Table 7 gives the percent of trials that are stopped early (ET) and the corresponding average sample size N_{et} when the working model is Morgenstern and the pairs (T, E) are generated using a Morgenstern model with $\gamma = 2$. Under scenarios II, IV, and VI, the estimated probability of stopping the trial early is small for $\delta_T, \delta_E \geq 0.7$ as expected. Although scenarios VII and VIII are cases where no good dose exists in $\mathcal{D}$, the probability of stopping the trial is much higher under scenario VIII compared to VII. The reason is that under scenario VIII, only four dose combinations have an acceptable level of toxicity and only the last dose combination at level $(6, 5)$ has an acceptable level of efficacy. In other words, the sets $\mathcal{H}^{\text{c}}$ and $\mathcal{L}^{\text{c}}$ are well separated. On the other hand, under scenario VII, half the number of doses are acceptable whereas four dose combinations satisfy $P_E \geq \theta_E$, i.e, the sets $\mathcal{H}^{\text{c}}$ and $\mathcal{L}^{\text{c}}$ are not well separated. Nevertheless, the percent selection of each dose combination under scenarios VII and VIII are very low (data not shown), consistent wit the ones obtained when $\gamma = 0$. In scenario IX, two doses are good but very close of the limit in term toxicity probability. This is why the probability of stopping the trial early is 0.71 when $\delta_T = 0.7$ and $\delta_E = 0.9$. In summary, the choice of the parameters δ_T and δ_E is a compromise between the rate of wrongly terminating a trial and the rate of true positive decision. These parameters should be chosen in close interaction with the clinicians. In our study, we used $(0.7, 0.9)$. We obtain 94.85% of good decision for the more toxic and non-efficacious scenario VIII and an average of 2.3% of wrong decision for scenarios II, IV and VI.

5. Conclusion

In this article, we proposed a two-stage design for dose finding in early phase I/II cancer trials with drug combinations of two cytotoxic agents. We explored two different models for the joint probability of toxicity and efficacy based on a class of copula models with varying degrees of association between toxicity and efficacy outcomes. Unlike other approaches, the method can recommend a single dose or a set of dose combinations at the end of the trial. In the latter case, we proposed a summary statistic to assess the reliability of the set estimator. The statistic gives an estimate of the probability that any given dose among the set of recommended doses at the end of the trial is safe and efficacious. This statistic can also be used for phase I drug combination trials with DLT

Table 7. Sensitivity analysis of the design parameters δ_T and δ_E.

Scenario	II		IV		VI		VII		VIII		IX	
(δ_T,δ_E)	ET	N_{et}	ET	N_{et}	ET	N_{et}	ET	N_{et}	ET	N_{et}	ET	N_{et}
$(0.6, 0.6)$	5.00	51.7	3.40	49.5	46.95	46.5	40.05	46.0	97.45	22.6	90.75	24.4
$(0.7, 0.7)$	0.90	51.5	0.50	40.8	21.30	49.0	29.50	47.2	95.25	24.2	80.40	26.7
$(0.8, 0.8)$	0.35	49.4	0.20	30.5	13.05	48.7	18.30	47.4	88.15	27.0	68.50	27.8
$(0.85, 0.85)$	0.10	54.5	0.15	30.7	6.25	48.5	13.70	47.2	82.55	27.6	60.10	27.8
$(0.9, 0.9)$	0.00	NaN	0.00	NaN	1.15	45.8	2.20	51.2	51.45	33.7	36.75	30.9
$(0.85, 0.95)$	0.00	NaN	0.10	22.5	1.00	48.4	8.20	48.5	80.10	28.4	49.70	27.9
$(0.7, 0.95)$	0.00	NaN	0.15	30.7	1.65	49.6	14.45	48.7	93.85	25.3	62.20	28.3
$(0.7, 0.9)$	0.20	48.2	0.20	32.5	6.10	48.7	23.25	47.5	94.85	24.4	70.55	27.8

outcome when more than one MTD is recommended.

Extensive simulation studies under misspecified scenarios show that the method is safe in terms of percent of patients allocated to toxic doses and that the trial stops early very frequently when the minimum dose combination is too toxic. The percent of recommendation of good doses is reasonably high for phase I/II trials with the lowest level obtained when the model is extremely misspecified (scenario III). In the extreme scenarios where no safe and efficacious dose combination exists among the set of available doses in the trial, the probability that the trial stops early depends on how well the sets of non-toxic and efficacious doses are separated. When these two sets are not well separated, the probability of stopping the trial early for futility is low but the trial is still safe. In any case, the average percent recommendation of a good dose is essentially 0 for scenarios VII and VIII (data not shown). We conclude that our approach is useful when clincians are interested in identifying more than one dose combination for future randomized studies. We are currently working on extending this work when the monotonicity assumptions described in Section 2.1 do not hold, e.g., when one or more drugs are biologic or immunotherapy. Other extensions include modeling late onset toxicity or efficacy or cases where efficacy takes more than three cycles to resolve. In this case, careful modeling of repeated toxicity outcome should be accounted for with possible dose reductions before the efficacy outcome is assessed if a DLT occurs.

6. Acknowledgments

This work is supported in part by the National Institute of Health Grant Number R01 CA188480-01A1, the National Center for Research Resources, Grant UL1RR033176, and is now at the National Center for Advancing Translational Sciences, Grant UL1TR000124, and P01 CA098912.

References

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

Braun, T. M. and Wang, S. (2010) A hierarchical bayesian design for phase i trials of novel combinations of cancer therapeutic agents. *Biometrics*, **66**, 805–812.

Cai, C., Yuan, Y. and Ji, Y. (2014) A Bayesian dose finding design for oncology clinical trials of combinational biological agents. *Applied Statistics*, **63**, 159–173.

Chen, Z., Yuan, Y., Li, Z., Kutner, M., Owonikoko, T., Curran, W., Khuri, F. and Kowalski, J. (2015) Dose escalation with over-dose and under-dose controls in phase I/II clinical trials. *Contemporary Clinical Trials*, **43**, 133–141.

Ivanova, A. (2003) A new dose-finding design for bivariate outcomes. *Biometrics*, **59**, 1001–1007.

Joe, H. (1997) *Multivariate models and dependence concepts*. London: Chapman & Hall.

Mander, A. P. and Sweeting, M. J. (2015) A product of independent beta probabilities dose escalation design for dual-agent phase i trials. *Statistics in Medicine*, **34**, 1261–1276.

Murtaugh, P. and Fisher, L. (1990) Bivariate binary models of efficacy and toxicity in dose-ranging trials. *Communications in Statistics. Theory and Methods*, **19**, 2003–2020.

Nelsen, R. B. (1999) *An Introduction to copulas*. New York: Springer.

O'Quigley, J., Pepe, M. and Fisher, L. (1990) Continual reassessment method: a practical design for phase 1 clinical trials in cancer. *Biometrics*, 33–48.

Riviere, M., Yuan, Y., Dubois, F. and Zohar, S. (2015) A Bayesian dose-finding design for clinical trials combining a cytotoxic agent with a molecularly targeted agent. *Journal of the Royal Statistical Society: Series C*, **64**, 215–229.

Riviere, M. K., Yuan, Y., Dubois, F. and S, Z. (2014) A Bayesian dose-finding design for drug combination clinical trials based on the logistic model. *Pharmaceutical Statistics*, **13**, 247–257.

Sato, H., Hirakawa, A. and Hamada, C. (2016) An adaptive dose-finding method using a change-point model for molecularly targeted agents in phase I trials. *Statistics in Medicine*, **35**, 4093–4109.

Shi, Y. and Yin, G. (2013) Escalation with overdose control for phase i drug-combination trials. *Statistics in Medicine*, **32**, 4400–4412.

Thall, P. and Cook, J. (2004a) Dose-finding based on efficacytoxicity trade-offs. *Biometrics*, **60**, 684–693.

Thall, P. and Russell, K. (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. and Cook, J. D. (2004b) Dose-finding based on efficacy–toxicity trade-offs. *Biometrics*, **60**, 684–693.

Thall, P. F., Millikan, R. E., Mueller, P. and Lee, S.-J. (2003) Dose-finding with two agents in phase i oncology trials. *Biometrics*, **59**, 487–496.

Tighiouart, M., Li, Q., Piantadosi, S. and Rogatko, A. (2016) A bayesian adaptive design for combination of three drugs in cancer phase i clinical trials. *American Journal of Biostatistics*, **6**, 1–11.

Tighiouart, M., Li, Q. and Rogatko, A. (2017) A Bayesian adaptive design for estimating the maximum tolerated dose curve using drug combinations in cancer phase I clinical trials. *Statistics in Medicine*, **36**, 280–290.

Tighiouart, M., Piantadosi, S. and Rogatko, A. (2014) Dose finding with drug combinations in cancer phase I clinical trials using conditional escalation with overdose control. *Statistics in Medicine*, **33**, 3815–3829.

Wages, N. A. and Conaway, M. R. (2014) Phase i/ii adaptive design for drug combination oncology trials. *Statistics in medicine*, **33**, 1990–2003.

Wages, N. A., Conaway, M. R. and O'Quigley, J. (2011) Continual reassessment method for partial ordering. *Biometrics*, **67**, 1555–1563.

Wang, K. and Ivanova, A. (2005) Two-dimensional dose finding in discrete dose space. *Biometrics*, **61**, 217–222.

Yin, G. and Yuan, Y. (2009a) Bayesian dose finding in oncology for drug combinations by copula regression. *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, **58**, 211–224.

— (2009b) A latent contingency table approach to dose finding for combinations of two agents. *Biometrics*, **65**, 866–875.

Yuan, Y. and Yin, G. (2011) Bayesian phase I/II adaptively randomized oncology trials with combined drugs. *Annals of Applied Statistics*, **5**, 924–942.