## POLITECNICO DI TORINO
## Repository ISTITUZIONALE

Cancer phase I trial design using drug combinations when a fraction of dose limiting toxicities is attributable to one or more agents

*Terms of use:*

(Article begins on next page)

19 April 2024

# Cancer phase I trial design using drug combinations when a fraction of dose limiting toxicities is attributable to one or more agents

**Jose L. Jimenez**[*,1]**, Mourad Tighiouart** [2]**,** and **Mauro Gasparini** [1]

[1] Politecnico di Torino, Dipartimento di Scienze Matematiche, Corso Duca degli Abruzzi, 24, 10129, Turin, Italy

[2] Samuel Oschin Comprehensive Cancer Institute, 8700 Beverly Blvd., Los Angeles, CA 90048, U.S.A.

Drug combination trials are increasingly common nowadays in clinical research. However, very few methods have been developed to consider toxicity attributions in the dose escalation process. We are motivated by a trial in which the clinician is able to identify certain toxicities that can be attributed to one of the agents. We present a Bayesian adaptive design in which toxicity attributions are modeled via Copula regression and the maximum tolerated dose (MTD) curve is estimated as a function of model parameters. The dose escalation algorithm uses cohorts of two patients, following the continual reassessment method (CRM) scheme, where at each stage of the trial, we search for the dose of one agent given the current dose of the other agent. The performance of the design is studied by evaluating its operating characteristics when the underlying model is either correctly specified or misspecified. We show that this method can be extended to accommodate discrete dose combinations.

*Key words:* Attributable toxicity; Cancer phase I trials; Continual reassessment method; Copula type models; Drug combination

## 1 Introduction

Cancer phase I clinical trials constitute the first step in investigating a potentially promising combination of cytotoxic and biological agents. Due to safety and ethical concerns, patients are sequentially enrolled in the trial, and the dose combinations assigned to subsequent patients depend on dose combinations already given to previous patients and their dose limiting toxicity (DLT) status at the end of the first cycle of therapy. The main objective of these trials is to estimate a maximum tolerated dose (MTD) that will be used in future efficacy evaluation in phase II/III trials. The MTD is usually defined as any dose combination $(x, y)$ that will produce DLT in a prespecified proportion $\theta$ of patients,

$$\text{Prob}(\text{DLT}|\text{ dose} = (x, y)) = \theta. \tag{1}$$

The definition of DLT depends on the type of cancer and drugs under study, but it is usually defined as a grade 3 or 4 non-hematologic toxicity (see the National Cancer Institute CTCAE v4.03 for the definition of the different grades of toxicity). The pre-specified proportion of DLTs $\theta$, sometimes referred as target probability of DLT, also depends on the nature of the toxicity, but it usually take values between 0.2 and 0.4.

In the drug combination dose finding literature, designs that recommend a unique MTD (see Yin and Yuan (2009a,b); Wages *et al.* (2011a,b); Shi and Yin (2013); Riviere *et al.* (2014); Wheeler *et al.* (2017)) or multiple MTDs (see Thall *et al.* (2003); Wang and Ivanova (2005); Yuan and Yin (2008); Braun and Wang

---

*Corresponding author: e-mail: jose.jimenez@polito.it,

(2010); Mander and Sweeting (2015); Tighiouart *et al.* (2014, 2016, 2017)) have been studied extensively. Most of these methods use a parametric model for the dose-toxicity relationship

$$\text{Prob}(\text{DLT}|(x,y)) = F((x,y), \boldsymbol{\xi}), \tag{2}$$

where $(x,y)$ represents the drug combination of two agents, $F(.)$ is a known link function, e.g. a power model or a logistic model, and $\boldsymbol{\xi} \in R^d$ is a vector of $d$ unknown parameters. Non-parametric designs have been proposed in the past, both in single agent and drug combination settings Mander and Sweeting (2015); Gasparini and Eisele (2000); Whitehead *et al.* (2010). These designs unique assumption is monotonicity, which is imposed either through the prior distribution (see Gasparini and Eisele (2000); Whitehead *et al.* (2010)), or by choosing only monotonic contours when escalating (see Mander and Sweeting (2015)).

Let $S$ be the set of all dose combinations available in the trial, and $C(\boldsymbol{\xi})$ be the set of dose combinations $(x,y)$ such the probability of DLT equals a target risk of toxicity $\theta$. Hence,

$$C(\boldsymbol{\xi}) = \{(x,y) \in S : F((x,y), \boldsymbol{\xi}) = \theta\}. \tag{3}$$

Equation (3) is the traditional definition of MTD set. When $S$ is discrete, following Tighiouart *et al.* (2017), we can define the MTD as the set of dose combinations $(x,y)$ that satisfy

$$|F((x,y), \boldsymbol{\xi}) - \theta| \leq \delta, \tag{4}$$

since $C(\boldsymbol{\xi})$ may be empty, i.e., when the MTD is not in $S$. The threshold parameter $\delta, 0 < \delta < 1$, is pre-specified after close collaboration with the clinician.

This work is motivated by a cancer phase I trial a clinician at Cedars-Sinai Medical Center is planning. The trial involves the combination of Taxotere, a known cytotoxic agent, and Metformin, a diabetes drug, in advanced or metastatic breast cancer patients. According to the clinician, some DLTs can be attributable to either agent or both. For example, a grade 3 or 4 neutropenia can only be attributable to Taxotere and not Metformin. Furthermore, for ethical reasons, if a patient has a DLT attributable to Taxotere when treated with dose level $x_T$ of taxotere, then $x_T$ cannot be increased for the next patient in the trial (see the dose escalation restriction in Section 2.2). Very few methods have been developed to incorporate toxicity attribution in the dose escalation process. Yin and Yuan (2009a) proposed a design that models the joint probability of toxicity with a copula model known as the Gumbel model Murtaugh and Fisher (1990). This model allows the investigator to compute the probability of DLT when the DLT is exclusively attributed to one drug, the other one, or both. However, they require all toxicities to be attributable, which is rare in practice. Wheeler *et al.* (2017) proposed a semi-attributable toxicity design based on a trial with non-concurrent drug administration. In their design, one drug is administered at the beginning of the treatment cycle and the other drug is administered at a much later time point if and only if the patient did not experience DLT. If a DLT occurs before the second drug is administered, then the DLT is attributed to the first drug. However, if the DLT occurs after the second drug has been administered, then the DLT could be caused by any of the drugs and therefore is not attributable. Iasonos and O'Quigley (2016) propose a method that reduces the effect of the bias caused by toxicity attribution errors by using personalized scores instead of the traditional binary DLT outcome. Lee and Fan (2012) considered the toxicity attribution problem for ruled-based designs with non-overlapping toxicities.

In this article, we propose a Bayesian adaptive design for drug combinations that allows the investigator to attribute a DLT to one or both agents in an unknown fraction of patients, even when the drugs are given concurrently. The relationship between the dose combinations and the risk of toxicity is modeled using the same copula model used by Yin and Yuan (2009a). The design proceeds using a variation of the algorithm proposed in Tighiouart *et al.* (2017) where cohorts of two patients are allocated to dose

combinations where, at each stage of the trial, we search for the dose of one agent given the current dose of the other agent. Our approach differs from the methodologies of Yin and Yuan (2009a) and Tighiouart *et al.* (2017) in three aspects; (i) a non-negative fraction of DLTs are attributable to either one or both agents, (ii) the dose combination allocated to patients uses the CRM scheme as opposed to escalation with overdose control (EWOC) approach proposed by Babb *et al.* (1998), and (iii) if a current patient experiences DLT attributed drug $D_1$ at dose level $x_{D_1}$, then the dose level of agent $D_1$ cannot be more than $x_{D_1}$ for the next cohort of two patients. At the end of the trial, an estimate of the MTD curve is proposed as a function of Bayes estimates of the model parameters. Last, we show that our method can be easily adapted from a setting with continuous dose combinations to discrete dose combinations by rounding up the estimated MTD curve to the nearest discrete dose combinations.

The rest of the manuscript is organized as follows. In Section 2, we describe the model for the dose-toxicity relationship and the adaptive design to conduct the trial for continuous dose combinations. In Section 3, we study the performance of the method in terms of safety and efficiency of the estimate of the MTD set. In Section 4, we adapt our proposal to the setting of discrete dose combinations. In section 5, we conduct a model misspecification evaluation. Discussion and practical considerations of the method are discussed in Section 6.

## 2 Method

### 2.1 Dose-Toxicity Model

Let $X_{\min}, X_{\max}, Y_{\min}, Y_{\max}$, be the minimum and maximum doses available in a trial that combines drugs with continuous dose combination levels. The doses are standardized to be in a desired interval, e.g., [0.05, 0.3], so that $X_{\min} = Y_{\min} = 0.05$ and $X_{\max} = Y_{\max} = 0.3$. Let $F_\alpha(\cdot)$ and $F_\beta(\cdot)$ be parametric models for the probability of DLT of drugs $D_1$ and $D_2$, respectively. We specify the joint dose-toxicity relationship using the Gumbel copula model Murtaugh and Fisher (1990) as

$$
\begin{aligned}
\pi^{(\delta_1,\delta_2)} = \text{Prob}(\delta_1, \delta_2|x,y) = F_\alpha^{\delta_1}(x)\left[1 - F_\alpha(x)\right]^{1-\delta_1} \times \\
F_\beta^{\delta_2}(y)\left[1 - F_\beta(y)\right]^{1-\delta_2} + (-1)^{(\delta_1+\delta_2)} F_\alpha(x)\left[1 - F_\alpha(x)\right] F_\beta(y)\left[1 - F_\beta(y)\right] \frac{e^{-\gamma}-1}{e^{-\gamma}+1},
\end{aligned}
\tag{5}
$$

where $x$ is the standardized dose level of drug $D_1$, $y$ is the standardized dose level of agent $D_2$, $\delta_1$ is the binary indicator of DLT attributed to drug $D_1$, $\delta_2$ is the binary indicator of DLT attributed to drug $D_2$ and $\gamma$ is the interaction coefficient. We assume that the joint probability of DLT, when one of the drugs is held constant, is monotonically increasing; that is $\text{Prob}(\text{DLT}|x',y) \geq \text{Prob}(\text{DLT}|x,y)$ or $\text{Prob}(\text{DLT}|x,y') \geq \text{Prob}(\text{DLT}|x,y)$, where $x' > x$ and $y' > y$. A sufficient condition for this property to hold is to assume that $F_\alpha(\cdot)$ and $F_\beta(\cdot)$ are increasing functions with $\alpha > 0$ and $\beta > 0$. In this article we use $F_\alpha(x) = x^\alpha$ and $F_\beta(y) = y^\beta$. Using (5), if the DLT is attributed exclusively to drug $D_1$, then

$$
\pi^{(\delta_1=1,\delta_2=0)} = \text{Prob}(\delta_1 = 1, \delta_2 = 0|x,y) = x^\alpha(1-y^\beta) - x^\alpha\left(1-x^\alpha\right)y^\beta\left(1-y^\beta\right)\frac{e^{-\gamma}-1}{e^{-\gamma}+1}.
\tag{6}
$$

If the DLT is attributed exclusively to drug $D_2$, then

$$
\pi^{(\delta_1=0,\delta_2=1)} = \text{Prob}(\delta_1 = 0, \delta_2 = 1|x,y) = y^\beta(1-x^\alpha) - x^\alpha\left(1-x^\alpha\right)y^\beta\left(1-y^\beta\right)\frac{e^{-\gamma}-1}{e^{-\gamma}+1}.
\tag{7}
$$

If the DLT is attributed to both drugs $D_1$ and $D_2$, then

$$\pi^{(\delta_1=1,\delta_2=1)} = \text{Prob}(\delta_1 = 1, \delta_2 = 1 | x, y) = x^\alpha y^\beta + x^\alpha (1 - x^\alpha) y^\beta (1 - y^\beta) \frac{e^{-\gamma} - 1}{e^{-\gamma} + 1}. \quad (8)$$

Equation (6) represents the probability of DLT attributed to drug $D_1$ but not $D_2$. Note that there over 70 types of non-hematological and hematological adverse events (toxicities) listed in the Common Terminology Criteria for Adverse Events v4.0 (CTCAE) graded as 0, 1, 2, 3, and 4. In most cancer trials, a DLT is defined as any type of grade 3 or 4 non-hematologic or grade 4 hematologic toxicity. In the example described in the introduction, a grade 4 neutropenia can only be attributed to taxotere and not metformin regardless of the dose of metformin when these two drugs are used in combination. Although the dose level of drug $D_2$ appears (6) and changes the probability of DLT, it does not necessarily change the probability of DLT of all types of toxicities. Equations (7) and (8) are interpreted similarly.

Following Yin and Yuan (2009a), it is easy to see that the total probability of having a DLT is calculated as the sum of (6), (7) and (8). Hence,

$$\pi = \text{Prob}(\text{DLT}|x, y) = \pi^{(\delta_1=1,\delta_2=0)} + \pi^{(\delta_1=0,\delta_2=1)} + \pi^{(\delta_1=1,\delta_2=1)} =$$
$$x^\alpha + y^\beta - x^\alpha y^\beta - x^\alpha (1 - x^\alpha) y^\beta (1 - y^\beta) \frac{e^{-\gamma} - 1}{e^{-\gamma} + 1}. \quad (9)$$

We define the MTD as any dose combination $(x_*, y_*)$ such that $\text{Prob}(\text{DLT}|x_*, y_*) = \theta$. We set (9) equal to $\theta$ and re-write it as a 2nd degree polynomial in $y^\beta$, and solve for the solutions. This allows us to define the MTD set $C(\alpha, \beta, \gamma)$ as

$$C(\alpha, \beta, \gamma) = \left\{ (x_*, y_*) : y_* = \left[ \frac{-(1 - x_*^\alpha - \kappa) \pm \sqrt{(1 - x_*^\alpha - \kappa)^2 - 4\kappa(x_*^\alpha - \theta)}}{2\kappa} \right]^{\frac{1}{\beta}} \right\}, \quad (10)$$

where

$$\kappa = x_*^\alpha (1 - x_*^\alpha) \frac{e^{-\gamma} - 1}{e^{-\gamma} + 1}.$$

Let $T$ be the indicator of DLT, $T = 1$ if a patient treated at dose combination $(x, y)$ experiences DLT within one cycle of therapy that is due to either drug or both, and $T = 0$ otherwise. Among patients treated with dose combination $(x, y)$ who exhibit DLT, suppose that an unknown fraction $\eta$ of these patients have a DLT with known attribution, i.e. the clinician knows if the DLT is caused by drug $D_1$ only, or drug $D_2$ only, or both drugs $D_1$ and $D_2$. Let $A$ be the indicator of DLT attribution when $T = 1$. It follows that for each patient treated with dose combination $(x, y)$, there are five possible toxicity outcomes: $\{T = 0\}, \{T = 1, A = 0\}, \{T = 1, A = 1, \delta_1 = 1, \delta_2 = 0\}, \{T = 1, A = 1, \delta_1 = 0, \delta_2 = 1\}$ and $\{T = 1, A = 1, \delta_1 = 1, \delta_2 = 1\}$. This is illustrated in the chance tree diagram in Figure 1. Using equations (6),(7),(8),(9) and Figure 1, the contributions to the likelihood from each of the five observable outcomes are listed in Table 1. The likelihood function is defined as

$$L(\alpha, \beta, \gamma \mid \text{data}) = \prod_{i=1}^{n} \left[ \left( \eta \pi_i^{(\delta_{1_i}, \delta_{2_i})} \right)^{A_i} \left( \pi_i (1 - \eta) \right)^{1-A_i} \right]^{T_i} (1 - \pi_i)^{1-T_i}, \quad (11)$$

and the joint posterior probability distribution of the model parameters as

$$\text{Prob}(\alpha, \beta, \gamma \mid \text{data}) \propto \text{Prob}(\alpha, \beta, \gamma) \times L(\alpha, \beta, \gamma \mid \text{data}). \quad (12)$$

With equation (12) we can easily sample and obtain MCMC estimates of $\alpha$, $\beta$ and $\gamma$.

### 2.2 Trial Design

Dose escalation / de-escalation proceeds using the algorithm described in Tighiouart *et al.* (2017) but univariate continual reassessment method (CRM) is carried out to estimate the next dose instead of EWOC. In a cohort with two patients, the first one would receive a new dose of agent $D_1$ given the dose $y$ of agent $D_2$ that was previously assigned. The new dose of agent $D_1$ is defined as $x_{\text{new}} = \text{argmin}_u |\widehat{\text{Prob}}(\text{DLT}|u, y) - \theta|$, where $y$ is fixed and $\widehat{\text{Prob}}(\text{DLT}|u, y)$ is computed using equation (9) with $\alpha, \beta, \gamma$ replaced by their posterior medians. The other patient would receive a new dose of agent $D_2$ given the dose of agent $D_1$ that was previously assigned. Specifically, the design proceeds as follows:

1. Patients in the first cohort receive the same dose combination $(x_1, y_1) = (x_2, y_2) = (X_{\min}, Y_{\min})$.

2. In the $i$-th cohort of two patients,

   - If $i$ is even,
     - Patient $(2i-1)$ receives doses $(x_{2i-1}, y_{2i-1})$, where $x_{2i-1} = \underset{u}{\text{argmin}} \left| \widehat{\text{Prob}}(\text{DLT}|u, y_{2i-3}) - \theta \right|$, and $y_{2i-1} = y_{2i-3}$. If a DLT was observed in the previous cohort of two patients and was attributable to drug $D_1$, then $x_{2i-1}$ is further restricted to be no more than $x_{2i-3}$.
     - Patient $2i$ receives doses $(x_{2i}, y_{2i})$, where $y_{2i} = \underset{v}{\text{argmin}} \left| \widehat{\text{Prob}}(\text{DLT}|x_{2i-2}, v) - \theta \right|$, and $x_{2i} = x_{2i-2}$. If a DLT was observed in the previous cohort of two patients and was attributable to drug $D_2$, then $y_{2i}$ is further restricted to be no more than $y_{2i-2}$.
   - If $i$ is odd,
     - Patient $(2i-1)$ receives doses $(x_{2i-1}, y_{2i-1})$, where $y_{2i-1} = \underset{v}{\text{argmin}} \left| \widehat{\text{Prob}}(\text{DLT}|x_{2i-3}, v) - \theta \right|$, and $x_{2i-1} = x_{2i-3}$. If a DLT was observed in the previous cohort of two patients and was attributable to drug $D_2$, then $y_{2i-1}$ is further restricted to be no more than $y_{2i-3}$.
     - Patient $2i$ receives doses $(x_{2i}, y_{2i})$, where $x_{2i} = \underset{u}{\text{argmin}} \left| \widehat{\text{Prob}}(\text{DLT}|u, y_{2i-2}) - \theta \right|$, and $y_{2i} = y_{2i-2}$. If a DLT was observed in the previous cohort of two patients and was attributable to drug $D_1$, then $x_{2i}$ is further restricted to be no more than $x_{2i-2}$.

3. Repeat step 2 until the maximum sample size is reached subject to the following stopping rule.

4. We would stop the trial if, $\text{Prob}(\text{Prob}(\text{DLT}|x = X_{\min}, y = Y_{\min}) \geq \theta + \xi_1|\text{data}) > \xi_2$, i.e. if the posterior risk of toxicity at the lowest combination significantly is high. $\xi_1$ and $\xi_2$ are design parameters tuned to obtain the best operating characteristics.

In step 2 of the algorithm, any dose escalation is further restricted to be no more than a pre-specified fraction of the dose range of the corresponding agent. At the end of the trial, we obtain the MTD curve estimate $\hat{C} = C(\hat{\alpha}, \hat{\beta}, \hat{\gamma})$, where $\hat{\alpha}$, $\hat{\beta}$ and $\hat{\gamma}$ are the posterior medians of the parameters $\alpha, \beta$ and $\gamma$, given the data.

## 3 Simulation Studies

### 3.1 Simulation set up and Scenarios

In all simulated trials, the link functions $F_\alpha(x) = x^\alpha$ and $F_\beta(y) = y^\beta$ are used. To evaluate the performance of our proposal, the DLT outcomes are generated from the true model showed in (9). We used this model in 3 different scenarios to study the behavior of our design when the prior distribution of the model parameters is both well and poorly calibrated. Let $\alpha_{\text{true}}, \beta_{\text{true}}$ and $\gamma_{\text{true}}$ represent the true parameter values

we use in (9) to generate DLT outcomes. In each scenario we select different values for $\alpha_{\text{true}}$, $\beta_{\text{true}}$, but the prior distribution for $\alpha$ and $\beta$, $P(\alpha)$ and $P(\beta)$, as well as $\gamma_{\text{true}}$, do not vary. In scenario 1, we choose values for $\alpha_{\text{true}}$ and $\beta_{\text{true}}$ such that $\alpha_{\text{true}} < E[P(\alpha)]$ and $\beta_{\text{true}} < E[P(\beta)]$. In scenario 2, we choose values for $\alpha_{\text{true}}$ and $\beta_{\text{true}}$ such that $\alpha_{\text{true}} = E[P(\alpha)]$ and $\beta_{\text{true}} = E[P(\beta)]$. Last, in scenario 3, we choose values for $\alpha_{\text{true}}$ and $\beta_{\text{true}}$ such that $\alpha_{\text{true}} > E[P(\alpha)]$ and $\beta_{\text{true}} > E[P(\beta)]$. Figure 2 shows the MTD curves with the true parameter values described here and their contours at $\theta \pm 0.05$ and $\theta \pm 0.1$. We evaluate the effect of toxicity attribution in these 3 scenarios using 4 different values for $\eta$: 0, 0.1, 0.25 and 0.4. These values are reasonable because higher values of $\eta$ in practice are very rare. Data is randomly generated using the following procedure:

- For a given dose combination $(x, y)$, a binary indicator of DLT $T$ is generated from a Bernoulli distribution with probability of success computed using equation (9).

- If $\{T = 1\}$, we generate the attribution outcome $A$ using a Bernoulli distribution with probability of success $\eta$.

- If $\{T = 1, A = 1\}$, we attribute the DLT to drug $D_1$, $D_2$, or to both drugs with equal probabilities.

We assume that the model parameters $\alpha, \beta, \gamma$ and $\eta$ are independent *a priori*. We assign vague prior distributions to $\alpha$, $\beta$ and $\gamma$ following Yin and Yuan (2009a), where $\alpha \sim \text{Uniform}(0.2, 2)$, $\beta \sim \text{Uniform}(0.2, 2)$ and $\gamma \sim \text{Gamma}(0.1, 0.1)$. These prior distributions correspond to the ones used by Yin and Yuan (2009a) for the main analysis. The prior distribution for the fraction of attributable toxicities $\eta$ is set to be Uniform$(0, 1)$. With these prior distributions, the true parameter values for each scenario are as follows. In scenario 1, $\alpha = \beta = 0.9$ and $\gamma = 1$. In scenario 2, $\alpha = \beta = 1.1$ and $\gamma = 1$. Last, in scenario 3, $\alpha = \beta = 1.3$ and $\gamma = 1$. For each scenario, $m = 1000$ trials will be simulated. The target risk of toxicity is fixed at $\theta = 0.3$, the sample size is $n = 40$, and the values for $\xi_1$ and $\xi_2$ will be 0.05 and 0.8 respectively. All simulation are done using the software R version 3.3.1.

### 3.2 Design Operating Characteristics

We evaluate the performance of the design by assessing its safety and its efficiency in estimating the MTD curve.

For trial safety, we employ the average percent of DLTs, the percent of simulated trials with DLT rate greater than $\theta \pm 0.05$ and $\theta \pm 0.10$.

For efficiency, we employ the pointwise average relative minimum distance from the true MTD curve to the estimated MTD curve. This measure of efficiency is well described in Tighiouart *et al.* (2014, 2017) and can interpreted as a pointwise average bias in estimating the true MTD curve. We also consider 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 is no more than $(100 \times p)\%$ of the true MTD curve. This measurement will give us an estimate of the percent of trials with MTD recommendation within $(100\ p)\%$ of the true MTD. This measure of efficiency can be interpreted as the pointwise percent of correct MTD recommendation. In this paper we select $p = 0.1, 0.2$. For a detailed explanation of these measures of efficiency, see Tighiouart *et al.* (2014, 2017).

### 3.3 Results

In general, increasing the value of $\eta$ until 0.4 generates estimated MTD curves closer to the true MTD curve. Figure 3 shows the estimated MTD curves for each scenario as a function of $\eta$. In terms of safety, overall we observe that increasing the fraction of toxicity attributions $\eta$ reduces the average percent of toxicities and percent of trials with toxicity rates greater than $\theta + 0.05$ and $\theta + 0.10$. Table 2, shows the average percent of toxicities as well as the percent of trials with toxicity rates greater than $\theta + 0.05$ and $\theta + 0.1$ for scenarios 1-3.

Figure 4 shows the pointwise average bias of the 3 proposed scenarios for each value of $\eta$. Overall, increasing the value of $\eta$ until 0.4 reduces the pointwise average bias. In any case, the pointwise average bias is around 10% of the dose range of either drug and practically negligible for $\eta = 0.25, 0.4$. For instance, under scenario 3, the maximum absolute value of the pointwise average bias when $\eta = 0.40$ is about 0.01, which corresponds to 0.3% of the dose range, which is practically negligible.

Figure 5 shows the pointwise percent of MTD recommendation of the 3 proposed scenarios for each value of $\eta$. In general, increasing the value of $\eta$ increases the pointwise percent of MTD recommendation, reaching up to 80% of correct recommendation when $p = 0.2$, and up to 70% of correct recommendation when $p = 0.1$. Based on these simulation results, we conclude that in continuous dose setting the approach of partial toxicity attribution generates safe trial designs and efficient estimation of the MTD.

## 4    Discrete Dose Combinations

### 4.1    Approach

Dose escalation follows the same procedure described in section 2.2. The only difference is that, in step 2, the continuous doses recommended are rounded to the nearest discrete dose level. For a detailed explanation of this procedure see Tighiouart *et al.* (2017).

### 4.2    Illustration

We study the performance of our proposal in a discrete dose level setting where the probability of toxicity of each dose level is generated from the working model. We employ 6 scenarios with 4 dose levels respectively in each drug for scenarios 1 - 3, and 4 and 6 dose levels respectively in each drug for scenarios 4 - 6. The target probability of toxicity is always $\theta = 0.3$ and, for each scenario, we simulate $m = 1000$ trials using the same vague priors for $\alpha$, $\beta$ and $\gamma$ specified in section 3.1. The maximum sample size in all scenarios is again $n = 40$. The performance of the method is evaluated using the percent of MTD selection statistic proposed by Tighiouart *et al.* (2017).

In Table 3 we present the 6 mentioned scenarios we use to illustrate the implementation of our design with discrete dose levels. Moreover, in Figure 6 we show the dose-toxicity surface of these 6 scenarios, where we observe that all of them have a flat (near-constant) surface.

In Table 4 we show the percent of times that at least 25%, 50%, 75% or 100% of recommended MTDs belong to the true MTD set. Using vague prior distributions, the scenario where toxicity attribution has the strongest effect is scenario 2. In scenarios 1,4 and 5, we observe a slight effect but it does not make a big difference.

## 5    Model Misspecification

In the previous sections, all the simulated scenarios are generated with the model showed in (9). However, in practice we do not know the underlying model that generates the data and therefore we need to assess the performance of our design under model misspecification. We employ the same toxicity scenarios used by Yin and Yuan (2009a), which are shown in Table 5. Moreover, In Figure 7 we show the dose-toxicity surface of these scenarios. Scenario 1 presents a very constant surface gradient. The rest of the scenarios present surface gradients that vary as we increase the dose combination levels. However, scenarios 3, 4 and 6 vary more abruptly than scenarios 2 and 5. Scenario 6 is a particular case because the lowest dose combination level has a probability of DLT that is already higher than the target risk of toxicity $\theta + 0.1$. Therefore, for this scenario, instead of presenting the percent of correct recommendation we present the percent of times the trial is stopped due to safety using the stopping rule in Section 2.2 with $\xi_1 = 0.05$ and $\xi_2 = 0.8$. For each scenario, we simulate $m = 1000$ trials with a target risk of toxicity of $\theta = 0.30$, a sample size of $n = 40$ and we use the same prior distributions for $\alpha$, $\beta$ and $\gamma$ as in section 3.1.

In terms of safety, in general we observe that toxicity attributions reduce the average percent of toxicities and percent of trials with toxicity rates greater than $\theta + 0.05$ and $\theta + 0.10$. Table 6 shows the average percent of toxicities as well as the percent of trials with toxicity rates greater than $\theta + 0.05$ and $\theta + 0.1$.

In Table 7, we show the percent of times that at least 25%, 50%, 75% or 100% of recommended MTDs belong to the true MTD set. In scenario 1 we observe a positive effect of the toxicity attributions, improving the percent of times at least 75% and 100% of recommended MTDs belong to the true MTD set in to 5% when $\eta = 0.25$. In scenario 2 we observe a positive effect of the toxicity attributions improving the percent of times at least 25% and 50% of recommended MTDs belong to the true MTD up to 5% and 4% respectively. In scenarios 3, 4 and 5 we do not observe any positive effect when attributing toxicities. However, these scenarios are particularly difficult for our design given the rounding up procedure we follow with discrete dose combinations. In scenario 6 we observe a positive effect of the toxicity attributions, improving the percent of times the trial is stopped due to safety by almost 4% when $\eta = 0.40$. Based on these simulation results under model misspecification, we conclude that the partial toxicity attribution method has good operating characteristics in recommending dose combinations of which, at least 50% are the true MTDs; these percent of correct recommendations vary between 65% to 98% depending on the scenario. Moreover, there is a high probability of stopping the trial if there is evidence that the minimum dose combination in the trial has high probability of DLT.

However some of the scenarios showed in Table 5 have a set of true MTDs that include a large number of dose combinations. For this reason, we implemented our design in 6 extra scenarios taken from Yin and Yuan (2009a,b), and presented in Table S1 at the supplementary material, where a set of true MTDs contains a much more restricted number of dose combinations. Also, since the scenarios presented in Table 5 where generated with a logistic model, we selected the scenarios to observe how robust is our proposal in scenarios generated with other models, such us the Clayton Copula, and scenarios that were arbitrarily generated. Moreover, since we are using the same set of true MTDs as Yin and Yuan (2009a,b), we make a performance comparison in terms of percent of correct MTD selection, with the methodology proposed by Yin and Yuan (2009a,b).

In Tables S2 and S3, in the supplementary material, we present operating characteristics in term of safety and efficiency for each of the 6 scenarios. In general, we observe that the design behaves in a similar way as with the scenarios presented along this manuscript. In terms of safety, toxicity attributions reduce the average percent of toxicities and the percent of trials with toxicity rates greater than $\theta + 0.05$ and $\theta + 0.10$. In terms of efficiency, we only observe a positive effect in scenarios with a relatively flat dose-toxicity surface. In terms of performance comparison, our proposed method is competitive with other standard designs for drug combinations such us Yin and Yuan (2009a,b), and achieves better percent of correct MTD recommendation in 4 out of the 6 employed scenarios.

Another issue that is relevant to the methodology we present in this manuscript is the errors in the attribution of toxicities by the treating investigators. Our design does not include a parameter to control the uncertainty around the decision made by the investigator when attributing the the DLT, which could be an extension of this work. However, in the supplementary material, to assess the impact of these kind of errors, we have simulated 3 scenarios taken also from Yin and Yuan (2009a,b) where we introduce 10% and 50% of errors in the attribution of DLTs, and compare it to the case where we correctly attributes 100% of the DLTs. In Tables S4 and S5, we present the simulated results in terms of safety and efficiency. Overall we do not observe any major difference when incorrectly attributing 10% and 50% of the DLTs with respect to correctly attributing 100% of the DLTs.

## 6 Conclusions

In this paper we proposed a Bayesian adaptive design for cancer phase I clinical trials using drug combinations with continuous dose levels and attributable DLT in a fraction of patients.. A copula-type model was used to describe the relationship between dose combinations and probability of DLT. The trial design

proceeds by treating cohorts of two patients, each patient with a different dose combination estimated using univariate CRM for a better exploration of the space of doses. Treating cohorts of two patients will allow trial conduct to be completed in a reasonable amount of time. Although the two patients in a cohort are allocated to different dose combinations, 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 within our dose range, that is was treated with either dose x or dose y. The use of continuous dose levels is not uncommon in early phase trials, particularly when the drugs are given as infusions intravenously. For instance, we recently designed a drug combination trial of cabazitaxel and cisplatin delivered intravenously in advanced prostate cancer patients where the dose levels are continuous and the protocol was approved by the scientific review at Cedars-Sinai. For ethical reasons, we further imposed dose escalation restrictions for one of the drugs when a DLT is attributable to that drug.

We studied the operating characteristics of the design under various scenarios for the true location of the MTD curve. In general, we observed that the trial is safe and as the proportion of attributed toxicities increases, the average proportion of toxicities decreases when we attribute toxicities. To assess the efficiency when estimating the MTD curve, we employed the pointwise average bias and average percent selection. In general the method is efficient although the results varied depending on the proportion of attributed toxicities. Note that the operating characteristics were evaluated under vague prior distributions of the model parameters and no toxicity profiles of single agent trials were used *a priori*. We also showed how the method can be adapted to the setting of discrete dose combinations.

We also performed a model misspecification evaluation in scenarios with different dose-toxicity surfaces. We only observed a positive effect in terms of percent of correct MTD recommendation in scenarios with flat surfaces. In scenarios with non-flat dose-toxicity surfaces we observed a decline in performance of percent selection consistent with the findings by Riviere *et al.* (2014) when working with copula regression models. We also observed a positive effect in scenarios where the lowest dose combination has an excessively high probability of DLT. In this case, toxicity attributions improves the percent of times the trial was stopped due to safety. In all cases, safety of the trial is not compromised by accounting for a partial toxicity attribution. Clearly, there is a trade-off when increasing the fraction of DLT attribution to one or more drugs. The design is more conservative in future escalations, lowering the in-trial DLT percentages and reducing how quickly the MTD contour is reached, by favoring experimentation over recommendation.

Our design is practically useful when the two drugs do not have many overlapping toxicities, see e.g. Miles *et al.* (2002) for some examples of drug combination trials with these characteristics. In cases where we expect a high percent of overlapping DLTs, designs that do not distinguish between drug attribution listed in the introduction may be more appropriate. Our method relies on clinical judgment regarding DLT attribution. In many phase I trials, such decisions are subject to error classifications and a possible extension is to introduce a parameter to account for errors in toxicity attribution as in Iasonos and O'Quigley (2016) for single agent trials. We also plan to study the performance of this design using other link functions under different copula models, and extend this method to early phase cancer trials with late onset toxicity and by accounting for patient's baseline characteristic by extending the approaches in Tighiouart *et al.* (2014, 2012) to the drug combination setting.

## 7    Acknowledgments

## 8  Conflict of Interests Statement

The authors have declared no conflict of interest.

## References

Babb, J., Rogatko. A, and Zacks, S. (1998). Cancer phase i clinical trials: efficient dose escalation with overdose control. Statistics in medicine 1998; 17(10): 11031120.

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.

Cheung, Y.K. (2011). Dose finding by the continual reassessment method. CRC Press.

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

Gasparini, M. (2013). General classes of multiple binary regression models in dose finding problems for combination therapies. *Journal of the Royal Statistical Society: Series C (Applied Statistics)* **62**, 115–133.

Gasparini, M. and Eisele, J.(2000). A curve-free method for phase I clinical trials. *Biometrics* **56**, 609–615.

Iasonos, A., and O'Quigley, J. (2016). Phase I designs that allow for uncertainty in the attribution of adverse events. *Journal of the Royal Statistical Society: Series C (Applied Statistics)*.

Lee, B.L. and Fan, S.K. (2012). A two-dimensional search algorithm for dose-finding trials of two agents. *Journal of biopharmaceutical statistics* **22**, 802–818.

Mander, A. P., and Sweeting, M. J. (2015). A product of independent beta probabilities dose escalation design for dualagent phase I trials. *Statistics in medicine* **34**, 1261–1276.

Miles, D., Von Minckwitz, G. and Seidman, A. D. (2002). Combination versus sequential single-agent therapy in metastatic breast cancer. *The oncologist* 7 (Supplement 6), 13-19.

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

Ntzoufras, I. (2011). Bayesian modeling using WinBUGS. *John Wiley & Sons*.

O'Quigley, J., Paoletti, X. and Maccario, J. (2002) Non-parametric optimal design in dose finding studies. *Biostatistics* **3**, 51–56.

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

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

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

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., Cook-Wiens, G., and Rogatko, A. (2012). Incorporating a patient dichotomous characteristic in cancer phase I clinical trials using escalation with overdose control. Journal of Probability and Statistics, 2012.

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., Liu, Y. and Rogatko A. (2014). Escalation with overdose control using time to toxicity for cancer phase i clinical trials. *PloS one* **9**, e93070.

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., Conaway, M.R. and O'Quigley, J. (2011). Dose-finding design for multi-drug combinations. *Clinical Trials* **8**, 380–389.

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). TwoDimensional Dose Finding in Discrete Dose Space. *Biometrics* **61**, 217-222.

Wheeler, G. M., Sweeting, M. J., Mander, A. P., Lee, S. M., and Cheung, Y. K. K. (2017). Modelling semiattributable toxicity in dualagent phase I trials with nonconcurrent drug administration. *Statistics in medicine* **36**, 225–241.

Whitehead, J., Thygesen, H. and Whitehead, A. (2010). A bayesian dose-finding procedure for phase i clinical trials based only on the assumption of monotonicity. *Statistics in medicine* **29**, 1808–1824.

Yin, G. and Yuan, Y. (2009a). A latent contingency table approach to dose finding for combinations of two agents. *Biometrics* **65**, 866–875.

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

Yuan, Y., and Yin, G. (2008). Sequential continual reassessment method for twodimensional dose finding. *Statistics in medicine* **27**, 5664–5678.

Yuan, Y. and Yin, G. (2009). Bayesian dose finding by jointly modelling toxicity and efficacy as time-to-event outcomes. *Journal of the Royal Statistical Society: Series C (Applied Statistics)* **58**, 719–736.

**Table 1** Contributions to the likelihood function based on the observed outcomes: toxicity, attribution, attribution to drug 1 ($\delta_1$) and attribution to drug 2 ($\delta_2$) for each patient.

| Toxicity | Attribution | $\delta_1$ | $\delta_2$ | Likelihood |
|---|---|---|---|---|
| 0 | - | - | - | $1 - \pi = 1 - \left[x^\alpha + y^\beta - x^\alpha \times y^\beta - x^\alpha \left(1 - x^\alpha\right) y^\beta \left(1 - y^\beta\right) \frac{e^{-\gamma}-1}{e^{-\gamma}+1}\right]$ |
| 1 | 0 | - | - | $\pi \times (1 - \eta) = \left[x^\alpha + y^\beta - x^\alpha \times y^\beta - x^\alpha \left(1 - x^\alpha\right) y^\beta \left(1 - y^\beta\right) \frac{e^{-\gamma}-1}{e^{-\gamma}+1}\right] \times (1 - \eta)$ |
| 1 | 1 | 1 | 0 | $\pi \times \eta \times \frac{\pi^{(1,0)}}{\pi} = \eta \times \left[x^\alpha(1 - y^\beta) - x^\alpha \left(1 - x^\alpha\right) y^\beta \left(1 - y^\beta\right) \frac{e^{-\gamma}-1}{e^{-\gamma}+1}\right]$ |
| 1 | 1 | 0 | 1 | $\pi \times \eta \times \frac{\pi^{(0,1)}}{\pi} = \eta \times \left[y^\beta(1 - x^\alpha) - x^\alpha \left(1 - x^\alpha\right) y^\beta \left(1 - y^\beta\right) \frac{e^{-\gamma}-1}{e^{-\gamma}+1}\right]$ |
| 1 | 1 | 1 | 1 | $\pi \times \eta \times \frac{\pi^{(1,1)}}{\pi} = \eta \times \left[x^\alpha \times y^\beta + x^\alpha \left(1 - x^\alpha\right) y^\beta \left(1 - y^\beta\right) \frac{e^{-\gamma}-1}{e^{-\gamma}+1}\right]$ |

**Table 2** Operating characteristics summarizing trial safety in $m = 1000$ simulated trials.

| | | Average % of toxicities | % of trials with toxicity rate $> \theta + 0.05$ | % of trials with toxicity rate $> \theta + 0.10$ |
|---|---|---|---|---|
| Scenario 1 | $\eta = 0.00$ | 33.62 | 25.90 | 4.10 |
| | $\eta = 0.10$ | 32.67 | 22.60 | 4.80 |
| | $\eta = 0.25$ | 31.55 | 17.60 | 2.70 |
| | $\eta = 0.40$ | 30.70 | 13.30 | 2.00 |
| Scenario 2 | $\eta = 0.00$ | 30.64 | 9.40 | 0.90 |
| | $\eta = 0.10$ | 29.69 | 7.30 | 0.40 |
| | $\eta = 0.25$ | 28.76 | 5.00 | 0.20 |
| | $\eta = 0.40$ | 28.04 | 4.10 | 0.30 |
| Scenario 3 | $\eta = 0.00$ | 27.47 | 2.00 | 0.00 |
| | $\eta = 0.10$ | 26.80 | 1.80 | 0.00 |
| | $\eta = 0.25$ | 25.99 | 1.30 | 0.00 |
| | $\eta = 0.40$ | 25.37 | 0.70 | 0.00 |

**Table 3**  Dose limiting toxicity scenarios with $\theta = 0.3$ generated with the working model. In bold the dose combination levels that would compose the true MTD set.

| Dose level | 1 | 2 | 3 | 4 | 1 | 2 | 3 | 4 | 5 | 6 |
|---|---|---|---|---|---|---|---|---|---|---|
| | | | Scenario 1 | | | | | Scenario 4 | | |
| 4 | **0.39** | 0.46 | 0.52 | 0.58 | **0.39** | 0.43 | 0.47 | 0.51 | 0.55 | 0.58 |
| 3 | **0.31** | **0.38** | 0.46 | 0.52 | **0.30** | **0.35** | 0.40 | 0.44 | 0.48 | 0.52 |
| 2 | **0.22** | **0.31** | **0.38** | 0.46 | **0.22** | **0.27** | **0.32** | **0.37** | 0.41 | 0.46 |
| 1 | 0.13 | **0.22** | **0.31** | **0.39** | 0.13 | 0.19 | **0.24** | **0.29** | **0.34** | **0.39** |
| | | | Scenario 2 | | | | | Scenario 5 | | |
| 4 | **0.30** | **0.36** | 0.42 | 0.48 | **0.30** | **0.33** | **0.37** | 0.40 | 0.44 | 0.48 |
| 3 | **0.22** | **0.28** | **0.35** | 0.42 | **0.22** | **0.26** | **0.29** | **0.33** | **0.38** | 0.42 |
| 2 | 0.14 | **0.21** | **0.28** | **0.36** | 0.14 | 0.18 | **0.22** | **0.27** | **0.31** | **0.35** |
| 1 | 0.07 | 0.14 | **0.22** | **0.30** | 0.07 | 0.11 | 0.16 | 0.20 | **0.25** | **0.30** |
| | | | Scenario 3 | | | | | Scenario 6 | | |
| 4 | **0.23** | **0.27** | **0.33** | **0.39** | **0.23** | **0.25** | **0.28** | **0.32** | **0.35** | **0.39** |
| 3 | 0.16 | **0.21** | **0.26** | **0.33** | 0.16 | 0.18 | **0.22** | **0.25** | **0.29** | **0.33** |
| 2 | 0.09 | 0.14 | **0.21** | **0.27** | 0.09 | 0.12 | 0.16 | 0.19 | **0.23** | **0.27** |
| 1 | 0.04 | 0.09 | 0.16 | **0.23** | 0.04 | 0.07 | 0.11 | 0.14 | 0.19 | **0.23** |

**Table 4**  Percent of times that at least 25%, 50%, 75% or 100% of recommended MTDs belong to the true MTD set in $m = 1000$ simulated trials.

| | | | | | % of correct MTD recommendation for $\theta \pm 0.10$ | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | $\geq 25\%$ | $\geq 50\%$ | $\geq 75\%$ | 100% | | $\geq 25\%$ | $\geq 50\%$ | $\geq 75\%$ | 100% |
| $\eta = 0.00$ | | 91.40 | 87.30 | 83.70 | 83.70 | | 90.00 | 85.50 | **73.70** | 67.70 |
| $\eta = 0.10$ | | **92.50** | **87.80** | **83.90** | **83.90** | | **91.30** | **86.00** | 71.10 | 65.80 |
| $\eta = 0.25$ | Scenario 1 | 90.90 | 87.70 | 83.80 | 83.80 | Scenario 4 | 89.40 | 85.90 | 71.50 | **68.20** |
| $\eta = 0.40$ | | 90.90 | 87.70 | 83.80 | 83.80 | | 89.80 | 85.70 | 70.50 | 67.30 |
| $\eta = 0.00$ | | 78.10 | 78.10 | 73.60 | 73.60 | | 82.90 | 81.50 | **72.70** | **72.70** |
| $\eta = 0.10$ | | 79.80 | 79.80 | 73.90 | 73.90 | | 82.80 | 81.40 | 71.70 | 71.70 |
| $\eta = 0.25$ | Scenario 2 | 83.00 | 83.00 | 75.50 | 75.50 | Scenario 5 | 85.00 | 81.80 | 71.00 | 71.00 |
| $\eta = 0.40$ | | **83.50** | **83.50** | **76.20** | **76.20** | | **85.70** | **82.50** | 70.60 | 70.60 |
| $\eta = 0.00$ | | 99.10 | **99.00** | **97.00** | **97.00** | | **98.80** | **96.50** | **93.10** | **92.70** |
| $\eta = 0.10$ | | **99.30** | 98.60 | 95.10 | 95.10 | | 98.60 | 95.70 | 89.90 | 89.60 |
| $\eta = 0.25$ | Scenario 3 | 97.10 | 96.40 | 91.90 | 91.90 | Scenario 6 | 96.20 | 92.10 | 87.00 | 86.20 |
| $\eta = 0.40$ | | 95.90 | 95.10 | 89.50 | 89.50 | | 94.00 | 89.50 | 81.90 | 81.00 |

**Table 5**  Toxicity scenarios for the model misspecification evaluation. In bold the doses considered as true MTD set.

| Dose level | 1 | 2 | 3 | 4 | 1 | 2 | 3 | 4 |
|---|---|---|---|---|---|---|---|---|
| | | Scenario 1 | | | | Scenario 4 | | |
| 4 | **0.28** | 0.41 | 0.55 | 0.68 | 0.04 | 0.09 | 0.17 | **0.32** |
| 3 | **0.25** | **0.35** | 0.48 | 0.60 | 0.03 | 0.06 | 0.12 | **0.23** |
| 2 | **0.22** | **0.30** | **0.40** | 0.51 | 0.02 | 0.05 | 0.09 | 0.16 |
| 1 | 0.19 | **0.26** | **0.34** | 0.43 | 0.02 | 0.03 | 0.06 | 0.11 |
| | | Scenario 2 | | | | Scenario 5 | | |
| 4 | 0.17 | **0.29** | 0.45 | 0.62 | 0.12 | **0.26** | 0.48 | 0.71 |
| 3 | 0.14 | **0.23** | **0.35** | 0.50 | 0.09 | 0.19 | **0.36** | 0.57 |
| 2 | 0.12 | 0.18 | **0.27** | **0.38** | 0.07 | 0.14 | **0.26** | 0.43 |
| 1 | 0.09 | 0.14 | 0.19 | **0.27** | 0.05 | 0.10 | 0.18 | **0.30** |
| | | Scenario 3 | | | | Scenario 6 | | |
| 4 | **0.37** | 0.72 | 0.92 | 0.98 | 0.78 | 0.94 | 0.99 | 1.00 |
| 3 | **0.26** | 0.59 | 0.85 | 0.96 | 0.68 | 0.90 | 0.97 | 0.99 |
| 2 | 0.18 | 0.44 | 0.74 | 0.91 | 0.57 | 0.83 | 0.94 | 0.98 |
| 1 | 0.12 | **0.30** | 0.59 | 0.82 | 0.45 | 0.73 | 0.90 | 0.97 |

**Table 6**  Operating characteristics summarizing trial safety for model misspecification in $m = 1000$ simulated trials.

| | | Average % of toxicities | % of trials with toxicity rate $> \theta + 0.05$ | % of trials with toxicity rate $> \theta + 0.10$ |
|---|---|---|---|---|
| | $\eta = 0.00$ | 32.99 | 22.90 | 4.20 |
| | $\eta = 0.10$ | 32.19 | 18.50 | 2.90 |
| Scenario 1 | $\eta = 0.25$ | 31.43 | 15.80 | 2.60 |
| | $\eta = 0.40$ | 30.58 | 12.90 | 2.50 |
| | $\eta = 0.00$ | 29.85 | 6.60 | 0.20 |
| | $\eta = 0.10$ | 29.14 | 4.10 | 0.10 |
| Scenario 2 | $\eta = 0.25$ | 28.20 | 3.10 | 0.30 |
| | $\eta = 0.40$ | 27.90 | 2.30 | 0.00 |
| | $\eta = 0.00$ | 36.53 | 40.70 | 16.40 |
| | $\eta = 0.10$ | 35.13 | 33.90 | 12.50 |
| Scenario 3 | $\eta = 0.25$ | 33.94 | 28.60 | 11.00 |
| | $\eta = 0.40$ | 32.94 | 23.40 | 9.50 |
| | $\eta = 0.00$ | 22.43 | 0.00 | 0.00 |
| | $\eta = 0.10$ | 21.83 | 0.00 | 0.00 |
| Scenario 4 | $\eta = 0.25$ | 21.39 | 0.00 | 0.00 |
| | $\eta = 0.40$ | 20.87 | 0.00 | 0.00 |
| | $\eta = 0.00$ | 30.43 | 6.60 | 0.30 |
| | $\eta = 0.10$ | 29.48 | 3.30 | 0.10 |
| Scenario 5 | $\eta = 0.25$ | 28.60 | 3.60 | 0.00 |
| | $\eta = 0.40$ | 27.60 | 2.30 | 0.00 |

**Table 7**  Percent of times that at least 25%, 50%, 75% or 100% of recommended MTDs belong to the true MTD set in $m = 1000$ simulated trials under model misspecification. In scenario 6 we show the percent of times the trial is stopped due to safety reasons.

| | | \% of correct MTD recommendation for $\theta \pm 0.10$ | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | $\geq 25\%$ | $\geq 50\%$ | $\geq 75\%$ | $100\%$ | | $\geq 25\%$ | $\geq 50\%$ | $\geq 75\%$ | $100\%$ |
| $\eta = 0.00$ | | 82.90 | 75.60 | 55.40 | 55.40 | | **98.80** | **98.80** | **87.80** | **87.80** |
| $\eta = 0.10$ | Scenario 1 | 82.70 | 72.70 | 57.30 | 57.30 | Scenario 4 | 97.20 | 97.20 | 85.70 | 85.70 |
| $\eta = 0.25$ | | **83.30** | **75.80** | **60.40** | **60.40** | | 95.70 | 95.70 | 82.00 | 82.00 |
| $\eta = 0.40$ | | 80.60 | 73.10 | 57.60 | 57.60 | | 95.20 | 95.20 | 76.20 | 76.20 |
| $\eta = 0.00$ | | 74.70 | 71.00 | **58.20** | **45.70** | | **75.40** | **69.10** | **20.50** | **20.50** |
| $\eta = 0.10$ | Scenario 2 | 77.00 | 73.50 | 53.60 | 44.60 | Scenario 5 | 71.80 | 62.80 | 20.40 | 20.40 |
| $\eta = 0.25$ | | **79.60** | **75.00** | 50.00 | 41.20 | | 70.70 | 59.40 | 18.90 | 18.90 |
| $\eta = 0.40$ | | 77.30 | 73.10 | 47.90 | 37.80 | | 71.20 | 60.50 | 16.70 | 16.70 |
| $\eta = 0.00$ | | **76.90** | **65.30** | **23.30** | **23.30** | | | 83.60 | | |
| $\eta = 0.10$ | Scenario 3 | 72.50 | 61.80 | 21.90 | 21.90 | Scenario 6 | | 82.90 | | |
| $\eta = 0.25$ | | 66.40 | 57.30 | 18.60 | 18.60 | | | 84.80 | | |
| $\eta = 0.40$ | | 66.10 | 54.70 | 15.70 | 15.70 | | | **87.20** | | |