Escalation with overdose control for drug combination using time to toxicity data in cancer phase I trials

Márcio A. Diniz[1], Jose L. Jimenez[2], and Mourad Tighiouart [1]
[1]Cedars-Sinai Medical Center, Los Angeles, California, U.S.A.
[2]Politecnico di Torino, Turin, Italy

Abstract
Bayesian adaptive methods for phase I trials such as the continual reassessment method (CRM) and escalation with overdose control (EWOC) do not take into account the time for a patient exhibits dose limiting toxicity (DLT), since the toxicity outcome is treated as a binary variable. This characteristic is proven to increase the study duration and makes it almost non-feasible in practice. Despite there are several methodologies to overcome this limitation in single agent trials, there only a few approaches specifically designed for drug combinations that model the time to toxicity. In this paper, we propose a Bayesian adaptive EWOC design that incorporates both the status of DLT and the time to toxicity using a cure rate model in clinical trials with drug combinations.

1. Introduction

Cancer phase I clinical trials are the first step when investigating the safety of a potential drug combination between cytotoxic and biological agents. Patients are usually recruited sequentially and the dose in which a new cohort of patients is allocated depends on the doses and dose limiting toxicity (DLT) status of all the previous patients enrolled in the trial. The target those combination (x, y) is referred to as the maximum tolerated dose (MTD), and is defined as the dose that produces DLTs in a given fraction of patients:

$$P(DLT|(x, y)) = \theta. \tag{1.1}$$

There are two phase I designs that perhaps are the most popular: the continual reassessment method (CRM) [1], with its modifications [2, 3, 4, 5, 6], and escalation with overdose control (EWOC) [7], with its extensions [8, 9, 10, 11, 12, 13].

These designs assume that the toxicities will be resolved before enrolling the next patient in the trial. Such assumption if followed increases the trial duration making the study non-feasible if it has cycles of therapy longer than 2 months, which is not rare in treatments that involve radiotherapy. In this way, ad-hoc solutions are to recommend the same dose used in the last patient or escalate the dose based on the partial information without modeling such uncertainty.

Several methodologies have been developed to address this limitation for a single agent. Cheung et al. [14, 15] proposed the time-to-event CRM (TITE-CRM), which is an extension

the continual reassessment method (CRM) to allow late-onset toxicities by incorporating different weights in the likelihood function. Mauguen et al. [16] developed a similar approach using escalation with overdose control (EWOC). Liu et al. [17] presents a CRM design showing that the late onset toxicities are informative missing data. Tighiouart et al. [18] discussed an EWOC design that models time to toxicity using a proportional hazards model to describe the partial information available.

In the last decade, statisticians have studied phase I designs for drug combinations. There are still some points of controversy such as the selection of the MTD. Several authors [19, 20, 21, 22, 23, 24] proposed methods that select a unique MTD whereas others [25, 26, 27, 28, 29, 30, 31] proposed methods that select multiple MTDs.

Nonetheless, most of them assume that toxicities will be observed in a short window of time before enrolling a new patient. There are few exceptions: Wages et al. [32] generalized TITE-CRM for two agents and Liu et al. [33] did the same based on their previous work [17].

In this article, we propose a Bayesian dose-finding design that selects multiple MTDs modeling the status of DLT and the time it takes for a patient to exhibit DLT in drug combination trials with late-onset toxicities. We describe the two-dimensional dose-toxicity surface using the the cure rate model [34], which takes into account that a certain fraction of patients in the trial may never experience a DLT.

The rest of the article is structured as follows. In section 2, we present the dose-toxicity model and the dose escalation algorithm. In section 3, we describe the simulated scenarios, the design operating characteristics and the results. Last, in section 4 we present some concluding remarks.

2. Method

2.1 Dose-toxicity model

Following Chen et al. [34], let N denote the number of toxicities for a patient after the initial treatment assuming a Poisson distribution with mean $\lambda > 0$. Also let W_i denote the random time for the i^{th} toxicity to happen, that is, W_i can be viewed as an incubation time for the i^{th} toxicity.

The variables W_i, $i = 1, 2, \ldots$ are assumed to be i.i.d. with a common distribution function $F(t|\mu) = 1 - S(t|\mu)$, where μ is a vector of parameters associated with F. The time to toxicity is defined by the random variable $T = \min\{W_i, 0 \leq i \leq N\}$, where $P(W_0 = \infty) = 1$ and N is independent of the sequence $W_1, W_2, \ldots$. The survival function

for T is given by,

$$\begin{aligned} S_T(t) &= P(T > t) \\ &= P(N = 0) + P(W_1 > t, \ldots, W_n > T, N \geq 1) \\ &= exp(-\lambda) + \sum_{k=1}^{\infty} S(t)^k \frac{\lambda^k}{k!} exp(-\lambda) \\ &= exp(-\lambda + \lambda S(t)) \\ &= exp(-\lambda F(t)), \end{aligned} \tag{2.1}$$

where S_T is not a proper survival since

$$S_T(\infty) \equiv P(N = 0) = exp(-\lambda) \neq 0. \tag{2.2}$$

We also see from equation (2.2) that the cure fraction is given by $exp(-\lambda)$. As $\lambda \to \infty$, the cure fraction tends to 0, whereas as $\lambda \to 0$, the cure fraction tends to 1. The hazard function for all the patients is given by

$$h_T(t) = \lambda f(t). \tag{2.3}$$

Note that $h_T(t)$ is not a hazard function corresponding to a probability function since $S_T(t)$ is not a proper survival function.

The survival and hazard functions for the patients who experienced DLT is defined as

$$S_{DLT}(t) = P(T > t | N \geq 1) = \frac{exp(-\lambda F(t)) - exp(-\lambda)}{1 - exp(-\lambda)}, \tag{2.4}$$

$$h_{DLT}(t) = \frac{exp(-\lambda F(t))}{exp(-\lambda F(t)) - exp(-\lambda)} h_{ALL}(t), \tag{2.5}$$

where S_{DLT} and $h_{DLT}(t)$ are proper survival and hazard functions, respectively.

In this way, if $Z \sim$ Bernoulli(p) is the indicator of DLT in the window $[0, \tau]$, where p is the probability of patient exhibits DLT within one cycle of therapy, follows that

$$p = 1 - P(N = 0) = 1 - exp(-\lambda). \tag{2.6}$$

The probability of DLT can be modeled as

$$p = G(\beta_0 + \beta_1 x + \beta_2 y + \beta_3 xy), \tag{2.7}$$

where $x \in [X_{min}, X_{max}]$ is the dose level of agent A, $y \in [Y_{min}, Y_{max}]$ is the dose level of agent B, and G is a known cumulative distribution function. Suppose that the doses of agents A and B are continuous and standardized to be in the interval $[0, 1]$.

It is assumed that the probability of DLT increases with the dose of any one of the agents when the other one is held constant. A sufficient condition for this property to hold is to assume $\beta_1 > 0$ and $\beta_2 > 0$ and the interaction term β_3 is non-negative.

The relationship between p and λ presented in (2.6) implies that,

$$\lambda = -log(1-p). \tag{2.8}$$

Thus, the regression coefficients β affects the cure fraction for the patients who did not experienced DLT as can be seen by equation (2.7), and the hazard function for the patients who experienced DLT as can be seen by equation (2.5) and (2.8).

The MTD is defined as any dose combination (x^*, y^*) such that

$$\text{Prob}(Z = 1|x^*, y^*) = \theta. \tag{2.9}$$

The target probability of DLT θ is set relatively high when the DLT is a reversible or non-fatal condition, and low when it is life threatening. Using (2.7) and (2.9), then

$$C = \left\{ (x^*, y^*) \in [0,1]^2 : y^* = \frac{G^{-1}(\theta) - \beta_0 - \beta_1 x^*}{\beta_2 + \beta_3 x^*} \right\}. \tag{2.10}$$

The model (2.7) is reparametrized in terms of parameters that clinicians can easily interpret. Then we define ρ_{10}, the probability of DLT when the levels of drugs A and B are 1 and 0, respectively, ρ_{01}, the probability of DLT when the levels of drugs A and B are 0 and 1, respectively, ρ_{00}, the probability of DLT when the levels of drugs A and B are both 0, and the interaction parameter η. It can be shown that

$$\begin{aligned} \beta_0 &= G^{-1}(\rho_{00}), \\ \beta_1 &= G^{-1}(\rho_{10}) - G^{-1}(\rho_{00}), \\ \beta_2 &= G^{-1}(\rho_{01}) - G^{-1}(\rho_{00}), \\ \beta_3 &= \eta. \end{aligned} \tag{2.11}$$

The MTD set (2.10) becomes

$$C = \left\{ (x^*, y^*) : y^* = \frac{(G^{-1}(\theta) - G^{-1}(\rho_{00})) - (G^{-1}(\rho_{10}) - G^{-1}(\rho_{00}))x^*}{(G^{-1}(\rho_{01}) - G^{-1}(\rho_{00})) + \eta x^*} \right\}. \tag{2.12}$$

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

$$\begin{aligned} L(\mu, \beta|\mathbf{D_n}) = &\left(\prod_{i=1}^{n} S(t_i|\mu)^{N_i - z_i} (N_i f(t_i|\mu))^{z_i} \right) \times \\ &exp\left\{ \sum_{i=1}^{n} [N_i log(\lambda_i) - log(N_i!)] - n\lambda_i \right\}. \end{aligned} \tag{2.13}$$

Summing out the unobserved latent vector $\mathbf{N}$, the complete-data likelihood given in (2.13) reduces to

$$L(\mu, \beta|\mathbf{D_n}) = \left(\prod_{i=1}^{n} (\lambda_i f(t_i|\mu))^{z_i} exp\{-\lambda_i (1 - S(t_i|\mu))\} \right). \tag{2.14}$$

There are several choices for the distributions F and G. The time to toxicity distribution F could be, for example, exponential, Weibull, log-normal and log-logistic while the dose-toxicity link could be any link for binary regression.

2.2 Prior and posterior distributions

In this section, we specify the prior distribution we place over the parameters we need to estimate: ρ_{01}, ρ_{10}, ρ_{00}, η, and μ.

Equation (2.11) implies that $0 < \rho_{00} < \min(\rho_{01}, \rho_{10})$ because β_1 and β_2 are positive. We assume that ρ_{01}, ρ_{10}, η, μ are independent *a priori* with: $\rho_{01} \sim \text{beta}(a_1, b_1)$, $\rho_{10} \sim \text{beta}(a_2, b_2)$, conditional on (ρ_{01}, ρ_{10}), $\rho_{00}/\min(\rho_{01}, \rho_{10}) \sim \text{beta}(a_3, b_3)$, and $\eta \sim \text{gamma}(a_4, b_4)$. If F is defined as Weibull(γ, κ) distribution, then $\mu = (\gamma, \kappa)$ such that $\gamma \sim \text{gamma}(a_5, b_5)$ and $\kappa \sim \text{gamma}(a_6, b_6)$.

Using Bayes rule, it is possible to show that the posterior distribution if the model parameters is proportional to the product of the likelihood functions times the prior distribution of each of the model parameters.

$$\pi(\rho_{01}, \rho_{10}, \rho_{00}, \eta, \mu|D_n) \propto \pi(\rho_{01}, \rho_{10}, \rho_{00}, \eta, \mu) \times \prod_{i=1}^{n} (\lambda_i f(t_i|\mu))^{z_i} \exp\{-\lambda_i (1 - S(t_i|\mu))\} \tag{2.15}$$

where λ_i is defined in (2.8) and (2.7).

We employed JAGS [35] to estimate the posterior distribution of these parameters and to estimate the design operating characteristics of the phase I clinical trial described in the following section.

2.3 Trial design

In this section we describe a dose escalation/de-escalation algorithm similar to the one followed by [31, 36] yet different given the time to event nature of our design. It follows the EWOC scheme, where at each evaluation, the posterior probability of overdosing a future patient is bounded by a feasibility bound α [7, 31]. For a given cohort of two patients, one receives a new dose of agent A for a given dose of agent B that was previously assigned. The other patient receives a new dose of agent B for a given dose of agent A that was previously assigned. However, instead of treating two patients at a time, we will treat patients in a continuous way and the posterior probability of overdosing will calculated after every enrolled patient. Specifically, the design proceeds as follows:

1. In the first cohort of patients, both receive the minimum dose combination available

$$(x_1, y_1) = (x_2, y_2) = (0, 0).$$

   Let be $D_n = \{(x_i, y_i, z_i), i = 1, \cdots, n\}\}$.

2. In the second cohort of patient, patient 3 receives dose (x_3, y_3) and patient 4 receives dose (x_4, y_4), where $y_3 = y_1$, $x_4 = x_2$, x_3 is the α-th percentile of $\pi(\Gamma_{A|B=y_1}|D_2)$, and y_4 is the α-th percentile of $\pi(\Gamma_{B|A=x_2}|D_3)$. Note that, the recommended dose combination for patient 3 is calculated using the dataset D_2, the recommended dose combination but for patient 4 is calculated using the dataset D_3.

3. In the i-th cohort of two patients,

   (a) If i is odd, then patient $(2i-1)$ receives dose (x_{2i-3}, y_{2i-1}) and patient $2i$ receives dose (x_{2_i}, y_{2i-2}), where

$$y_{2i-1} = \Pi^{-1}_{\Gamma_{B|A=x_{2i-3}}(\alpha|D_{2i-2})} \text{ and } x_{2i} = \Pi^{-1}_{\Gamma_{A|B=y_{2i-2}}(\alpha|D_{2i-1})}.$$

   (b) If i is even, then patient $(2i-1)$ receives dose (x_{2i-1}, y_{2i-3}) and patient $2i$ receives dose (x_{2_i-2}, y_{2i}), where

$$x_{2i-1} = \Pi^{-1}_{\Gamma_{A|B=y_{2i-3}}}(\alpha|D_{2i-2}) \text{ and } y_{2i} = \Pi^{-1}_{\Gamma_{B|A=x_{2i-2}}}(\alpha|D_{2i-1}).$$

4. Repeat step 3 until n patients are enrolled subject to the following stopping rule: if the α-th of $\pi(\Gamma_{A|B=y}|D)$ or $\pi(\Gamma_{B|A=x}|D)$ is less than 0 or greater than 1, the recommended dose for the next patient is 0 or 1 respectively. Also, any dose escalation is restricted to be no more than a pre-specified fraction of the dose range of the corresponding agent (e.g., 0.1).

2.3.1 Stopping rule

We stop the enrollment to the trial if $P(P(DLT|(x, y) = (0, 0)) > \theta + \xi_1|\text{data}) > \xi_2$, i.e., if the posterior probability that the probability of DLT at the minimum dose combination available in the trial exceeds the target probability of DLT plus ξ_1 is higher than ξ_2. ξ_1 and ξ_2 are design parameters chosen to achieve good operating characteristics.

At the end of the trial, we estimate the MTD using (2.12) as

$$\widehat{C} = \left\{(x^*, y^*) : y^* = \frac{(G^{-1}(\theta) - F^{-1}(\hat{\rho}_{00})) - (G^{-1}(\hat{\rho}_{10}) - G^{-1}(\hat{\rho}_{00}))x^*}{(G^{-1}(\hat{\rho}_{01}) - G^{-1}(\hat{\rho}_{00})) + \hat{\eta}x^*}\right\}, \qquad (2.16)$$

where $\hat{\rho}_{00}$, $hat\rho_{01}$, $\hat{\rho}_{10}$ and $\hat{\eta}$ are the posterior medians given the data D_n.

3. Simulation studies

3.1 Set-up and scenarios

We studied four scenarios for the true continuous MTD curve. In all cases, the time to toxicity distribution F is Weibull(1, 1.5), the dose-toxicity link is Logit, the target probability of DLT is fixed at $\theta = 0.33$, and the trial sample size is $n = 40$ patients. Scenario (1), $(\rho_{00}, \rho_{10}, \rho_{01}, \eta) = (10^{-7}, 3 \times 10^{-6}, 3 \times 10^{-6}, 10)$, shows two drugs that are very safe within the range of available doses in the trial but the true MTD curve lies near the upper-right corner of the xy plane. In scenario (2), $(0.01, 0.2, 0.9, 20)$, the MTD of agent A when agent B is at its minimum dose level is within the range of doses of drug A but the MTD of agent B when drug A is at its minimum dose level is above the maximum dose level of agent B. Scenario (3), $(0.001, 0.6, 0.01, 10)$, is a case where drug A is safe but the MTD of agent B when drug A is at its minimum dose level is just above 0.8 and scenario (4) (0.01, 0.2, 0.9, 100) presents two drugs that are highly interactive with $\eta = 100$.

Priors for ρ_{01}, ρ_{10}, ρ_{00}, η are the same chosen in Tighiouart et al. [31]. In addition, priors for γ, κ are chosen as gamma$(1, 1)$ and gamma$(1, 1)$.

3.2 Operating Characteristics

We evaluate the performance of the two model Time to Event and Binary assessing the safety of the trial designs as well as the efficiency of the estimate of the MTD curve based on 1000 simulated trials.

3.2.1 Safety

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

3.2.2 Efficiency

We present an estimate of the MTD curve using the average posterior medians of the model parameters. Under the reparameterization, the estimate is

$$\hat{C} = \left\{ (x^*, y^*) : y^* = \frac{(G^{-1}(\theta) - G^{-1}(\hat{\rho}_{00})) - (G^{-1}(\hat{\rho}_{10}) - G^{-1}(\hat{\rho}_{00}))x^*}{(G^{-1}(\hat{\rho}_{01}) - G^{-1}(\hat{\rho}_{00})) + \hat{\eta} x^*} \right\}, \tag{3.1}$$

where $F(.)$ is the logistic function and $\hat{\rho}_{00}, \hat{\rho}_{01}, \hat{\rho}_{10}, \hat{\eta}$ are the average posterior medians of the parameters $\rho_{00}, \rho_{01}, \rho_{10}, \eta$ from all $m = 1000$ trials. The next measure of efficiency is the pointwise average relative minimum distance from the true MTD curve to the estimated MTD curve. For $i = 1, \ldots, m$, let C_i be the estimated MTD curve and C_{true} be the true

MTD curve. For every point $(x, y) \in C_{true}$, let

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

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

$$d_{(x,y)} = \frac{1}{m}\sum_{i=1}^{m} d^{(i)}_{(x,y)}. \tag{3.3}$$

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

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

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

3.3 Results

Table 1 presents the average percent of DLTs varies between 9.05% and 40.5% for Binary model and 11.38% and 34.14% for Time to Event model. In addition, the percent of trials with an excessive DLT rate is significantly less when the uncertainty of the partial information is modeled than it is ignored.

Figure 1 shows the plots of the true and estimated MTD curves obtained using (3.1). The two designs are similar in the four scenarios.

Figure 2 displays the pointwise average relative minimum distance from the true MTD curve to the estimated MTD curved under the four scenarios (1)-(4) as defined in (3.3). In the scenarios (2) - (4), the average bias is similar for the two model. The binary model presents slight smaller bias only scenario (1).

Figure 3 contains the pointwise percent of MTD recommendation for tolerance p = 0.2 as defined in (3.4). In scenario (1), the binary model is better than time to toxicity model; in scenario (2) and (3), time to event model present better results, and in scenario (4) both models are similar to each other.

Table 1: Average DLT rate and % trials: DLT rate ¿ + 0.10 under scenarios (1)-(4)

Scenario	Design	DLT Rate (DLT Rate ¿ 0.43)
1	Time to Event	11.38 (0)
	Binary	9.05 (0)
2	Time to Event	29.85 (0.8)
	Binary	37.54 (15.4)
3	Time to Event	24.01 (0)
	Binary	32.01 (3.4)
4	Time to Event	34.14 (2.5)
	Binary	40.5 (27.2)

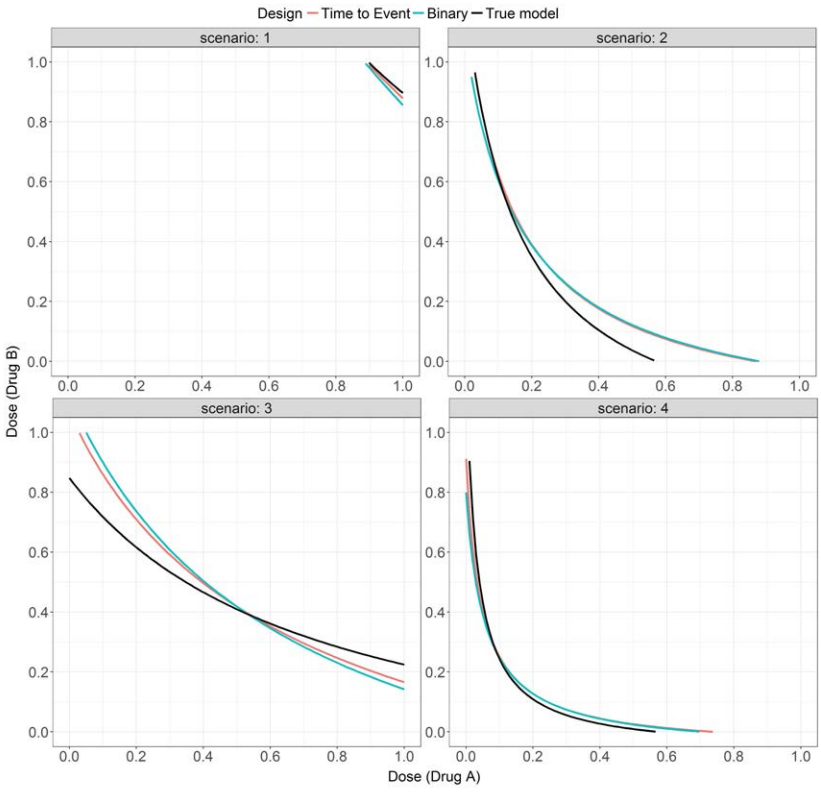

Figure 1: True and estimated MTD curves under scenarios (1)-(4)

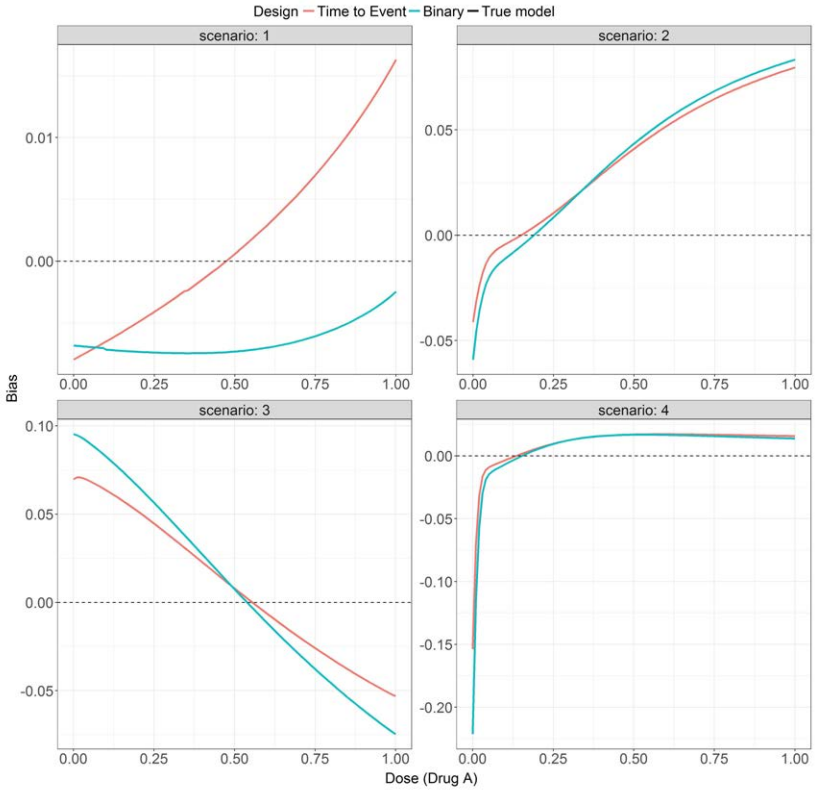

Figure 2: Pointwise average relative minimum distance from the true MTD curve to the estimated MTD curve under scenarios (1)-(4)

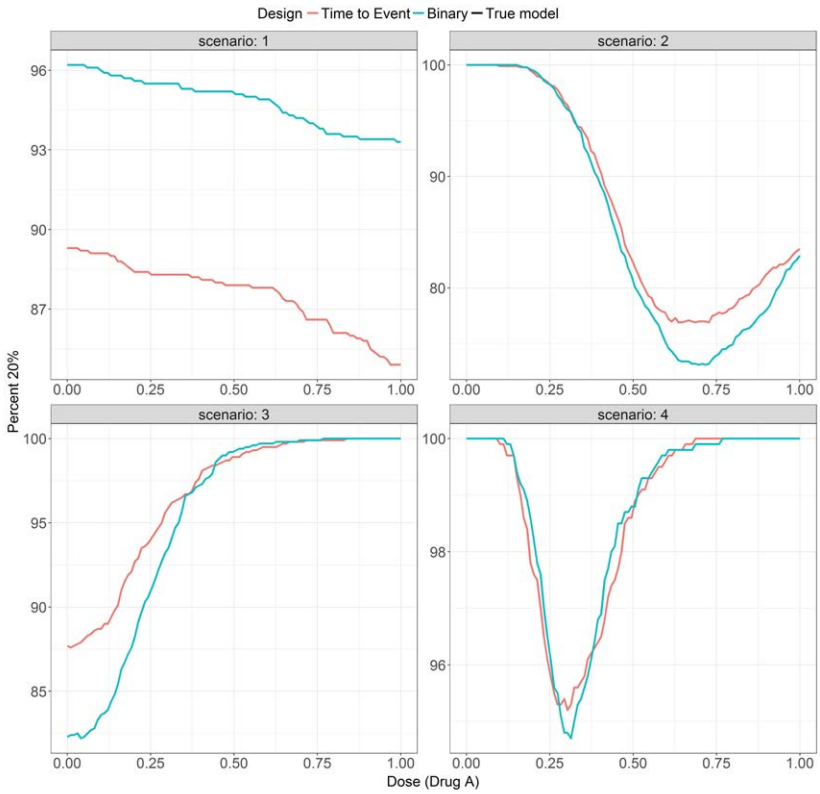

Figure 3: Pointwise percent of MTD recommendation for p = 0.2 under scenarios (1)-(4)

4. Concluding Remarks

We described Bayesian adaptive designs for cancer phase I clinical trials using two drugs with continuous dose levels taking in account lower grade toxicities. We compared these results to Tighiouart et al. [31]. In each case, vague priors were used for quantifying the toxicity profile of each agent a priori. We studied design operating characteristics of the methodology under four practical scenarios. The binary and time to event designs are similar in efficiency, but time to event is safer. Further investigation under different scenarios is still needed.

5. 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. It has also received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Sklodowska-Curie grant agreement, No 633567.

References

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

[2] Faries, D. (1994) Practical modifications of the continual reassessment method for phase i cancer clinical trials. *Journal of biopharmaceutical statistics*, **4**, 147–164.

[3] Goodman, S. N., Zahurak, M. L., and Piantadosi, S. (1995) Some practical improvements in the continual reassessment method for phase i studies. *Statistics in medicine*, **14**, 1149–1161.

[4] Møller, S. (1995) An extension of the continual reassessment methods using a preliminary up-and-down design in a dose finding study in cancer patients, in order to investigate a greater range of doses. *Statistics in medicine*, **14**, 911–922.

[5] Piantadosi, S., Fisher, J., and Grossman, S. (1998) Practical implementation of a modified continual reassessment method for dose-finding trials. *Cancer chemotherapy and pharmacology*, **41**, 429–436.

[6] Storer, B. E. (2001) An evaluation of phase i clinical trial designs in the continuous dose–response setting. *Statistics in medicine*, **20**, 2399–2408.

[7] Babb, J., Rogatko, A., and Zacks, S. (1998) Cancer phase i clinical trials: efficient dose escalation with overdose control. *Statistics in medicine*, **17**, 1103–1120.

[8] Zacks, S., Rogatko, A., and Babb, J. (1998) Optimal bayesian-feasible dose escalation for cancer phase i trials. *Statistics & probability letters*, **38**, 215–220.

[9] Tighiouart, M., Rogatko, A., and Babb, J. S. (2005) Flexible bayesian methods for cancer phase i clinical trials. dose escalation with overdose control. *Statistics in medicine*, **24**, 2183–2196.

[10] Tighiouart, M. and Rogatko, A. (2012) Number of patients per cohort and sample size considerations using dose escalation with overdose control. *Journal of Probability and Statistics*, **2012**.

[11] 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**.

[12] Tighiouart, M., Cook-Wiens, G., and Rogatko, A. (2012) Escalation with overdose control using ordinal toxicity grades for cancer phase i clinical trials. *Journal of probability and statistics*, **2012**.

[13] Tighiouart, M., Rogatko, A., et al. (2010) Dose finding with escalation with overdose control (ewoc) in cancer clinical trials. *Statistical Science*, **25**, 217–226.

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

[15] Polley, M.-Y. C. (2011) Practical modifications to the time-to-event continual reassessment method for phase i cancer trials with fast patient accrual and late-onset toxicities. *Statistics in medicine*, **30**, 2130–2143.

[16] Mauguen, A., Le Deley, M., and Zohar, S. (2011) Dose-finding approach for dose escalation with overdose control considering incomplete observations. *Statistics in medicine*, **30**, 1584–1594.

[17] Liu, S., Yin, G., and Yuan, Y. (2013) Bayesian data augmentation dose finding with continual reassessment method and delayed toxicity. *The annals of applied statistics*, **7**, 1837.

[18] 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.

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

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

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

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

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

[24] 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.

[25] 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.

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

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

[28] 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.

[29] 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.

[30] 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.

[31] 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.

[32] Wages, N. A., Conaway, M. R., and O'Quigley, J. (2013) Using the time-to-event continual reassessment method in the presence of partial orders. *Statistics in medicine*, **32**, 131–141.

[33] Liu, S. and Ning, J. (2013) A bayesian dose-finding design for drug combination trials with delayed toxicities. *Bayesian analysis*, **8**, 703.

[34] Chen, M.-H., Ibrahim, J. G., and Sinha, D. (1999) A new bayesian model for survival data with a surviving fraction. *Journal of American Statistical Association*, **94**, 909–919.

[35] Plummer, M. et al. (2003) Jags: A program for analysis of bayesian graphical models using gibbs sampling. *Proceedings of the 3rd international workshop on distributed statistical computing*, vol. 124, p. 125, Vienna, Austria.

[36] 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.