Designs of Early Phase Cancer Trials with Drug Combinations

Designs of Early Phase Cancer Trials with Drug Combinations

José L. Jiménez, Márcio Augusto Diniz, André Rogatko, and Mourad Tighiouart

1 Introduction

The primary objective of a phase I oncology trial is to estimate a maximum tolerable dose (MTD) of a new drug or combination of drugs for future efficacy evaluation in phase II/III trials. For the case of combination trials with two drugs, the MTD is any dose combination (x, y) of drugs A and B that produces DLT in a pre-specified proportion of patients θ [16].

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

The definition of DLT depends on the type of cancer and drugs used in the trial, but it is typically defined as a grade 3 or 4 non-hematologic or grade 4 hematologic toxicity. Different types and grades of toxicity are described in the Common Terminology Criteria for Adverse Events (CTCAE), an observer-rated toxicity grading system used in cancer clinical trials to assess the severity of various organ system toxicities associated with treatment [58]. Depending on the nature and severity of treatment-attributable toxicity, the target probability of DLT θ typically takes values between 0.1 and 0.4.

To model the probability of DLT, we assume a parametric model of the form

$$P(\text{DLT}|x, y) = F(x, y; \boldsymbol{\beta}), \tag{2}$$

J. L. Jiménez
Novartis Pharma A.G., Basel, Switzerland
e-mail: jose_luis.jimenez@novartis.com

M. A. Diniz · A. Rogatko · M. Tighiouart (✉)
Samuel Oschin Comprehensive Cancer Institute, Los Angeles, CA, USA
e-mail: Marcio.Diniz@cshs.org; Andre.Rogatko@cshs.org; mourad.Tighiouart@cshs.org

Y. Zhao, (Din) D.-G. Chen (eds.), *Modern Statistical Methods for Health Research*,
Emerging Topics in Statistics and Biostatistics,
https://doi.org/10.1007/978-3-030-72437-5_7

where $F(.)$ is a known link function (e.g., power model or a logistic model), and $\boldsymbol{\beta} \in R^p$ is a $p \times 1$ vector of unknown parameters. Non-parametric designs have also been proposed in the past, both for single agent and drug combination settings, see e.g., [15, 23, 27, 30, 64]. The common assumption in these designs is monotonicity (i.e., the probability of DLT increases with the dose of any one of the agents when the other one is held constant), which is imposed either through the prior distribution, or by choosing only monotonic contours when escalating.

Following [51, 53], the general phase I design for drug combinations can be stated as follows. Let S be the set of all dose combinations available in the trial and C be the set of combinations (x, y) that produce DLT in a proportion of patients that is equal to the target risk of DLT. Hence,

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

An alternative definition of the MTD is the set of dose combinations (x, y) that satisfy $|F(x, y; \boldsymbol{\beta}) - \theta| \leq \delta$, since the set C in (3) may be empty. This can happen, for example, when S is finite and the MTD is not part of the dose combinations available in the trial. The threshold parameter δ, $0 < \delta < 1$, referred to as "$100 \times \delta$-point window" in [5] must be pre-specified by the clinician.

Consecutive cohorts of one to three patients are enrolled in the trial, and the model parameters and estimated probabilities of toxicities are updated sequentially, using dose combinations allocated to all previously treated patients and their DLT outcomes. The next cohort of patients receives doses determined by minimizing the risk of exceeding the target probability of DLT according to some loss function. This general framework of dose finding for drug combinations was studied extensively in the last two decades, see e.g., [5, 12, 23, 27, 38, 42, 45, 50, 51, 53, 61, 62, 65, 67]. These methods are aimed at either identifying a single MTD or recommending more than one MTD combination for future efficacy studies. Approaches where a single MTD is selected may be sub-optimal because important dose combinations with similar acceptable DLT level and possibly with high probability of response may be missed. This could happen for two reasons. First, the discrete set of dose combinations is selected by the investigator based on prior experience with single agents. Therefore, when these agents are combined, the selected set may not include intermediate dose combinations with probability of DLT close to the target probability of DLT and the target probability of treatment response. Second, even if this discrete set includes dose combinations with probability of DLT close to the target, their probability of response may be very different and these approaches may recommend a dose combination with a low probability of response. Hence, approaches that recommend more than one MTD should be used for future efficacy studies using randomized or response-adaptive designs.

The main goal of early phase oncology phase I–II trials is to identify one or many dose combinations that are both safe and efficacious. In single-agent trials, where efficacy is evaluated within a short window of time (e.g., one or two cycles of therapy), one-stage sequential designs are frequently used by updating the joint probability of toxicity and efficacy after each cohort of patients [4, 7, 20, 31, 41,

44, 46]. This methodology has been extended to accommodate drug combinations [6, 19, 26, 39, 60, 67, 68]. However, if efficacy cannot be evaluated in a short time interval, two-stage designs are frequently employed. In the first stage, a set of maximum tolerated dose combinations is selected, and in the second stage, the set is tested for efficacy. The patient population used in the second stage may be different than that from the first stage [8, 24, 40]. For drug combination trials, different methods for two-stage designs have been proposed for binary efficacy endpoints [43, 47, 68, 70] and time-to-event efficacy endpoints [21].

In Sect. 2, we review some designs of drug combination trials focusing on continuous dose levels of both drugs. Estimation of doses to be allocated to the next cohort of patients uses the escalation with overdose control (EWOC) principle [1, 2, 13, 48, 49, 52, 54–56, 69, 71] and the continual reassessment method (CRM) [9, 14, 17, 29, 32–34]. A method that incorporates a covariate with the patients' baseline characteristics and settings where an unknown fraction of attributable DLTs will also be reviewed. In Sect. 3, we describe two-stage phase I/II designs based on the work of [47] and [21]. In each case, stage 1 follows the designs described in Sect. 2.1 to estimate the MTD curve. In stage 2, [21, 47] search for combinations along this MTD curve that maximize the probability of treatment response or median time to an event of interest. We conclude this chapter with a discussion including practical implementation of these designs and related ongoing research.

2 Designs for Phase I Clinical Trials

2.1 Phase I Model-based Designs for Drug Combinations

Model

Tighiouart et al. [51, 53] assumed that the doses (x, y) from a combination of two drugs A and B are continuous and standardized into the interval [0,1], with a dose-toxicity model of the form

$$P(Z = 1|x, y) = F(\beta_0 + \beta_1 x + \beta_2 y + \beta_3 xy), \tag{4}$$

where $\beta_1, \beta_2 > 0$, $\beta_3 \geq 0$, $Z = 1$ if a patient exhibits DLT within one cycle of therapy given the dose combination (x, y) and $Z = 0$ otherwise, and F is a cumulative distribution function (c.d.f). In particular, the logistic function, $F(u) = (1 + e^{-u})^{-1}$, has been used by several authors for single and drug combination trials, and the probit, normal, and complementary log-log link functions were used by Tighiouart et al. [51, 53] and Diniz et al. [12] to assess model misspecification.

Following (1) and (4), the MTD set is defined as

$$C = \left\{(x^*, y^*) \in S : y^* = \frac{F^{-1}(\theta) - \beta_0 - \beta_1 x^*}{\beta_2 + \beta_3 x^*}\right\}, \tag{5}$$

where (x^*, y^*) represents any dose combinations such that $P(Z = 1|x^*, y^*) = \theta$. In this context, the MTD is a hyperbola in the Cartesian plane (or a decreasing line if $\beta_3 = 0$).

Tighiouart et al. [51] reparameterized the model in (4) using the parameters $\rho_{10}, \rho_{01}, \rho_{00}$ corresponding to the probabilities of DLT when the levels of drugs A and B are 1 and 0, 0 and 1, and both 0, respectively. These parameters can be easily interpreted by clinicians, and they facilitate prior specifications since prior information on ρ_{01}, ρ_{10}, and ρ_{00} may be available from the previous trials. Moreover, this parametrization extends the one presented in [53], where it was assumed that the MTD of each drug is within the range of available doses of the corresponding agent. In this case, the MTD of each agent can lie outside the range of available doses in the trial when the other one is held at its minimum value.

The original parametrization can be recovered as follows:

$$\begin{aligned} \beta_0 &= F^{-1}(\rho_{00}), \\ \beta_1 &= F^{-1}(\rho_{10}) - F^{-1}(\rho_{00}), \\ \beta_2 &= F^{-1}(\rho_{01}) - F^{-1}(\rho_{00}), \\ \beta_3 &= \eta, \end{aligned} \tag{6}$$

such that $0 < \rho_{00} < \min(\rho_{01}, \rho_{10})$ since β_1 and β_2 are positive. The MTD set (5) becomes

$$C = \left\{ (x^*, y^*) \in S : y^* = \frac{F^{-1}(\theta) - F^{-1}(\rho_{00}) - (F^{-1}(\rho_{10}) - F^{-1}(\rho_{00}))x^*}{F^{-1}(\rho_{01}) - F^{-1}(\rho_{00}) + \eta x^*} \right\}. \tag{7}$$

Prior and Posterior Distributions

Tighiouart et al. [51] assumed that ρ_{01}, ρ_{10}, and η are independent a priori with $\rho_{01} \sim beta(a_1, b_1)$, $\rho_{10} \sim beta(a_2, b_2)$, and conditional on (ρ_{01}, ρ_{10}), $\frac{\rho_{00}}{min(\rho_{01}, \rho_{10})} \sim beta(a_3, b_3)$. A gamma distribution with mean $E(\eta) = a/b$ and variance $Var(\eta) = a/b^2$ is placed on the interaction term η. Vague $beta$ priors are achieved by taking $a_j = b_j = 1$, for $j = 0, 1, 2, 3$, while a vague gamma prior is chosen with a mean of 21 and a variance of 540.

Let $D_n = \{x_i, y_i\}, i = 1, \ldots, n$, be the data collected after enrolling n patients in the trial. The likelihood function for the model parameters is

$$\begin{aligned} \mathscr{L}(\rho_{00}, \rho_{10}, \rho_{01}, \eta) = \prod_{i=1}^{n} &(H(\rho_{00}, \rho_{10}, \rho_{01}, \eta; x_i, y_i))^{z_i} \\ &\times (1 - H(\rho_{00}, \rho_{10}, \rho_{01}, \eta; x_i, y_i))^{1-z_i}, \end{aligned} \tag{8}$$

where, using Eq. (6),

$$H(\rho_{00}, \rho_{10}, \rho_{01}, \eta; x_i, y_i) = F\left(F^{-1}(\rho_{00}) + (F^{-1}(\rho_{10}) - F^{-1}(\rho_{00}))x_i + (F^{-1}(\rho_{01}) - F^{-1}(\rho_{00}))y_i + \eta x_i y_i\right). \tag{9}$$

Therefore, using Bayes rule, the posterior distribution of the model parameters ρ_{00}, ρ_{01}, ρ_{10}, and η is proportional to the product of the likelihood and the prior distribution

$$\pi(\rho_{00}, \rho_{01}, \rho_{10}, \eta|D_n) \propto \pi(\rho_{00}|\rho_{01}, \rho_{10}) \times \pi(\rho_{01}) \times \pi(\rho_{10}) \times \pi(\eta) \times \mathscr{L}(\rho_{00}, \rho_{10}, \rho_{01}, \eta), \tag{10}$$

which is analytically intractable. Therefore, Monte Carlo Markov Chain (MCMC) methods are employed such as R [37] and JAGS [35] to estimate the features of the posterior distribution of the model parameters.

Trial Design

Tighiouart et al. [51] use a dose escalation/de-escalation algorithm that treats cohorts of two patients simultaneously based on the EWOC criterion, where at each stage of the trial, one subject receives a new dose of agent A for a given dose of agent B that was previously assigned and the other patient receives a new dose of agent B for a given dose of agent A that was previously assigned. Diniz et al. [12] extended this algorithm to the CRM criterion. The dose escalation algorithm is illustrated in Fig. 1.

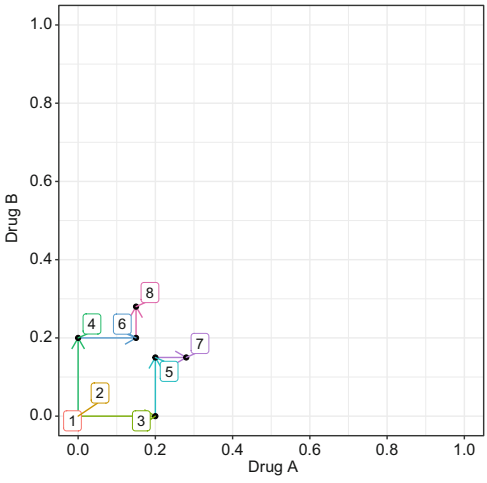

Fig. 1 Illustration of the dose escalation algorithm for the first 8 patients

Specifically, the design proceeds as follows:

1. Each patient in the first cohort of 2 patients receives the same dose combination $(x_i, y_i) = (0, 0)$ for $i = 1, 2$.
2. In the i-th cohort of 2 patients, for $i \geq 2$,
   (a) If i is even, then patient $2i - 1$ receives dose (x_{2i-1}, y_{2i-3}) and patient $2i$ receives dose (x_{2i-2}, y_{2i}), where

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

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

   for EWOC criterion.

$$x_{2i-1} = \underset{x}{\operatorname{argmin}}|H(\hat{\rho}_{00}, \hat{\rho}_{01}, \hat{\rho}_{10}, \hat{\eta}; x, y_{2i-3}) - \theta|$$

$$y_{2i} = \underset{y}{\operatorname{argmin}}|H(\hat{\rho}_{00}, \hat{\rho}_{01}, \hat{\rho}_{10}, \hat{\eta}; x_{2i-2}, y) - \theta|$$

   for CRM criterion.
   (b) If i is odd, then patient $2i - 1$ receives dose (x_{2i-3}, y_{2i-1}) and patient $2i$ receives dose (x_{2i}, y_{2i-2}), where

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

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

   for EWOC criterion.

$$x_{2i} = \underset{x}{\operatorname{argmin}}|H(\hat{\rho}_{00}, \hat{\rho}_{01}, \hat{\rho}_{10}, \hat{\eta}; x, y_{2i-2}) - \theta|$$

$$y_{2i-1} = \underset{y}{\operatorname{argmin}}|H(\hat{\rho}_{00}, \hat{\rho}_{01}, \hat{\rho}_{10}, \hat{\eta}; x_{2i-3}, y) - \theta|$$

   for CRM criterion.
3. Repeat step 2 until n patients are enrolled in the trial subject to a safety stopping rule in which the trial is stopped if the estimated probability of DLT at the lowest dose combination is higher than a pre-specified threshold.

$\Pi^{-1}_{\Gamma_{A|B=y}}(\cdot|D)$ denotes the inverse c.d.f. of the posterior distribution of the MTD of drug A given the level of drug $B = y$, and for the CRM method, $\hat{\rho}_q, \hat{\eta}, q \in \{00, 01, 10\}$ are the posterior medians.

The EWOC criterion consists of finding a dose $x^\star$ such that the posterior probability that the MTD exceeds this dose is bounded by a feasibility bound α. For example, in step 2 of the above algorithm, the dose of drug A is the maximum dose level of A such that the posterior probability that the MTD of A given that the level of drug B is y_{2i-3} exceeds $x^\star$ is bounded by α, i.e., $x^\star = x_{2i-1} = \Pi^{-1}_{\Gamma_{A|B=y_{2i-3}}}(\alpha|D_{i-1})$. Babb et al. [1] suggested a fixed feasibility boundary α equal to 0.25. Babb and Rogatko [2] introduced an increasing feasibility boundary until 0.5 with initial α equal 0.25, while Wheeler et al. [63] suggested a similar strategy, but conditional on the previous patient having no DLT.

The CRM criterion consists of finding a dose $x^\star$ such that it minimizes the absolute value difference between the estimated probabilities of DLT for the target toxicity rate θ. For example, in step 2 of the above algorithm, the dose of drug A is the dose $x^\star$ that minimizes $|H(\hat{\rho}_{00}, \hat{\rho}_{01}, \hat{\rho}_{10}, \hat{\eta}; x^\star, y_{2i-3}) - \theta|$.

At the end of the trial, the MTD (7) is estimated as

$$\widehat{C} = \left\{ (x^*, y^*) \in S : y^* = \frac{F^{-1}(\theta) - F^{-1}(\widehat{\rho}_{00}) - (F^{-1}(\widehat{\rho}_{10}) - F^{-1}(\widehat{\rho}_{00}))x^*}{F^{-1}(\widehat{\rho}_{01}) - F^{-1}(\widehat{\rho}_{00}) + \widehat{\eta}_3 x^*} \right\}. \tag{11}$$

The discussed approach can be easily extended to a discrete grid of doses, i.e., $(x_1, \cdots, x_r)$ and $(y_1, \cdots, y_r)$ be the doses of agents A and B, respectively. Trial design proceeds using the algorithm described in Sect. 2.1 with the continuous doses recommended in step 2 being rounded to the nearest discrete dose level.

At the end of the trial, a discrete set Γ of dose combinations satisfying (i) and (ii) below is selected as MTDs. Let C_i be the estimated MTD curve at the end of the trial and denote by $d((x_j, y_k), C_i)$ the Euclidian distance between the dose combination (x_j, y_k) and C_i as in (14).

(i) Let $\Gamma_A = \bigcup_{t=1}^{r} \left\{ (x, y_t) : x = \underset{x_j}{\text{argmin}}\, d((x_j, y_t), C_i) \right\}$,

$\Gamma_B = \bigcup_{t=1}^{r} \left\{ (x_t, y) : y = \underset{y_j}{\text{argmin}}\, d((x_t, y_j), C_i) \right\}$, and $\Gamma_0 = \Gamma_B \cap \Gamma_A$.

(ii) Let $\Gamma = \Gamma_0 \backslash \{(x*, y*) : P(|P(Z = 1|(x*, y*)) - \theta| > \delta_1 | D_n) > \delta_2\}$.

In (i), dose combinations closest to the MTD are selected by first minimizing the distances across the levels of drug A and then across the levels of drug B. In (ii), we exclude MTDs from (i) that likely to be either too toxic or too low. The design parameter δ_1 is selected after consultation with a clinician, and the parameter δ_2 is selected after exploring a large number of scenarios for a given prospective trial. Following [51], $\delta_1 = 0.1, \delta_2 = 0.3$.

Design Operating Characteristics

The performance of trial designs with finite sample size is assessed based on operating characteristics calculated from a Monte Carlo simulation study with m replicates, often with $m = 1000$.

In single agents, there are several operating characteristics such as bias, mean-squared error, average DLT rate, percentage of trials in which DLT rate is within an optimal toxicity interval, the percentage of trials with the estimated MTD within an optimal MTD interval, and the percentage of patients receiving optimal doses defined by those optimal intervals among others [13]. These operating characteristics can be divided in two classes measuring the safety of the trial design and the efficiency to estimate the MTD curve.

However, not all of them can be easily extended when estimating the MTD as a curve instead of a point in the dose space. Tighiouart et al. [53] presented some of these extensions.

Safety

The average percentage of DLT and the percentage of trials that have a DLT rate exceeding $\theta + \delta$ are, respectively, given by

$$\bar{\theta} = \frac{1}{m} \sum_{i=1}^{m} \hat{\theta}_i \tag{12}$$

$$\bar{\theta}_\delta = \frac{1}{m} \sum_{i=1}^{m} \mathbb{1}(\hat{\theta}_i > \theta + \delta) \tag{13}$$

where $\hat{\theta}_i$ is the estimated DLT rate for ith replicate for $i = 1, \ldots, m$. It is expected that $\bar{\theta}$ is close to θ, and the threshold $\delta = 0.1$ is considered to be an indication of an excessive DLT rate.

Efficiency

The pointwise average relative minimum distance from the true MTD curve to the estimated MTD curve can be interpreted as the pointwise average bias when estimating the MTD.

Let C_i be the estimated MTD curve for the ith Monte Carlo replicate and C_{true} be the true MTD curve. For every point $(x, y) \in C_{true}$, the minimum relative distance of the point (x, y) on the true MTD curve to the estimated MTD curve C_i can be calculated as follows:

$$d^{(i)}_{(x,y)} = sign(y' - y) \times min_{\{(x^*,y^*):(x^*,y^*)\in C_i\}}((x - x^*)^2 + (y - y^*)^2)^{1/2}, \tag{14}$$

where y' is such that $(x, y') \in C_i$ for $i = 1, \ldots, m$. If the point (x, y) is below C_i, then $d^{(i)}_{(x,y)}$ is positive. Otherwise, it is negative.

Then, the pointwise average bias is defined as follows:

$$d_{(x,y)} = \frac{1}{m} \sum_{i=1}^{m} d^{(i)}_{(x,y)}. \tag{15}$$

As the magnitude of bias is relative to the true MTD, it is also important to quantify the percentage of trials satisfying a given condition relative to the true MTD value. Let $\Delta(x, y)$ be the Euclidean distance between the minimum dose combination $(0, 0)$ and the point (x, y) on the true MTD curve, such that the minimum distance of the point (x, y) from the true MTD curve to the estimated MTD curve C_i is no more than $(100 \times p)\%$ of the distance of the true MTD from the minimum dose,

$$R_{(x,y)} = \frac{1}{m} \sum_{i=1}^{m} I\left(|d^{(i)}_{(x,y)}| \leq p\Delta(x, y)\right), \tag{16}$$

where $0 < p < 1$.

One can interpret (16) as drawing a circle with center (x, y) on the true MTD curve and radius $p\Delta(x, y)$, and then the percent of trials with the MTD curve estimate C_i within this circle is given by $R_{(x,y)}$. Therefore, the statistic (16) measures the percentage of correct recommendations.

Results

The methodology for a phase I trial proposed by Tighiouart et al. [51, 53] and Diniz et al. [12] has also been used by Diniz et al. [10], Tighiouart [47], and Jiménez et al. [21]. Therefore, there are several scenarios available in the literature with different values for ρ_{00}, ρ_{01}, ρ_{10}, and η. For illustration purposes, we present the operating characteristics of one of these multiple scenarios based on 2000 simulated trials. Dose escalation proceeds following EWOC and CRM criteria with the target toxicity rate $\theta = 0.33$. For EWOC, the feasibility boundary α starts at 0.25 with an increment of 0.05 for each new cohort of patients up to 0.5. Cohorts of two patients were accrued with the total sample size of 40 patients.

Assuming $(\rho_{00}, \rho_{10}, \rho_{01}, \eta) = (1, 0.01, 0.6, 10)$ from [51], Table 1 shows safety operating characteristics indicating that the proposed designs rarely surpass the toxicity rate given that one drug has low toxicity. Figure 2A shows the estimated MTD, with Fig. 2B indicating increasing bias at the edges of the MTD curves, varying from -0.06 to 0.06 for EWOC and -0.045 to 0.045 for CRM. The percentage of correct recommendation in Fig. 2C displays high values for both tolerances $p = 0.1, 0.2$ reaching the minimum values on the far left edge of MTD curve, with 63% for EWOC and 74% for CRM. Therefore, the CRM criterion

Table 1 Safety results from the simulated scenario presented in Fig. 2 from [51]

Design	Average % of toxicities	% of trials with DLT rate $> \theta + 0.05$	% of trials with DLT rate $> \theta + 0.10$
CRM	27.21	0.20	0.00
EWOC	25.25	0.05	0.00

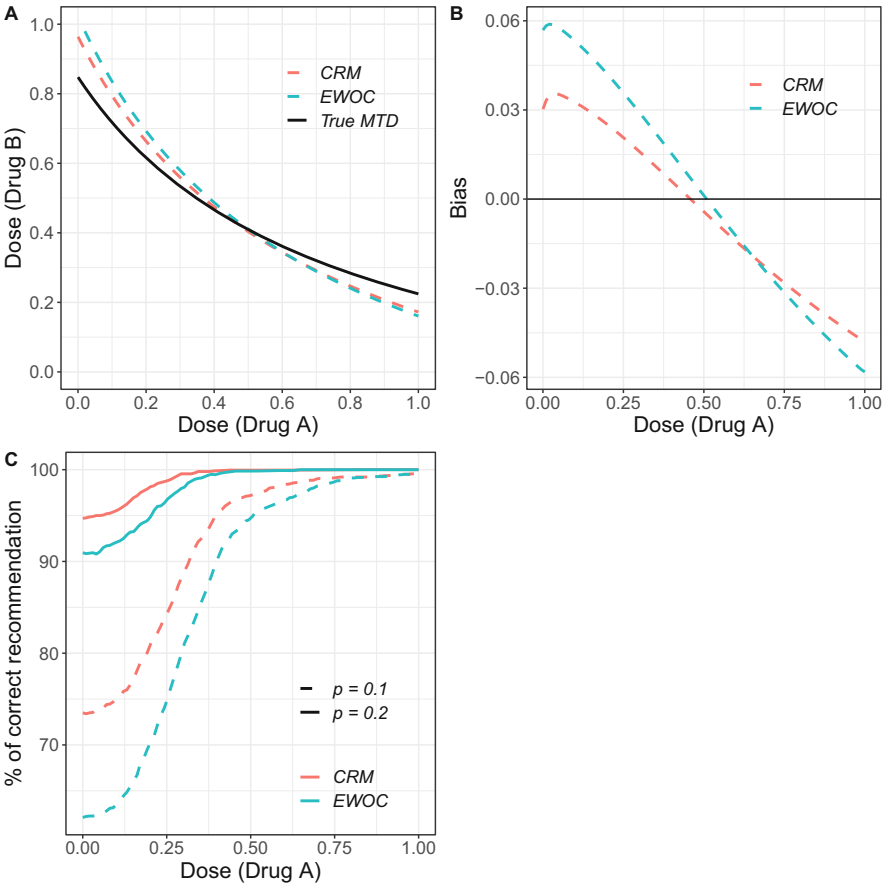

Fig. 2 Simulated scenario taken from [51], $(\rho_{00}, \rho_{10}, \rho_{01}, \eta) = (1, 0.01, 0.6, 10)$. In (A), we show the true and estimated MTD curves as defined in Eq. (7) as well as each final recommended dose combination after simulating 1000 trials. In (B) and (C), we observe the bias and the percentage of correct recommendation, respectively, for each value of dose for drug A contained in the MTD curve

presents superior operating characteristics for this scenario. Overall, both designs have good operating characteristics and are able to estimate the MTD curve while keeping the proportion of DLTs within reasonable boundaries.

2.2 Attributing Dose-Limiting Toxicities

Model

Following [22] and using the same notation as the one defined in Sect. 2.1, the doses of drugs A and B are standardized to be in a desired interval. The marginal probability of DLT of each compound is defined in terms of the power model (i.e., $P(Z = 1|x) = x^{\alpha}$ and $P(Z = 1|y) = y^{\beta}$), and we specify the joint probability of DLT using the Gumbel copula model (see [31]) as

$$\begin{aligned} &P(\delta_A, \delta_B|x, y) \\ &\quad = (x^{\alpha})^{\delta_A} \left(1 - x^{\alpha}\right)^{1-\delta_A} (y^{\beta})^{\delta_B} \left(1 - y^{\beta}\right)^{1-\delta_B} + (-1)^{(\delta_A+\delta_B)} \gamma(x, y), \end{aligned} \tag{17}$$

where $\gamma(x, y) = x^{\alpha} (1 - x^{\alpha}) y^{\beta} \left(1 - y^{\beta}\right) \frac{e^{-\gamma}-1}{e^{-\gamma}+1}$, δ_A is the binary indicator of DLT attributed to drug A, δ_B is the binary indicator of DLT attributed to drug B, and γ is the interaction parameter. A sufficient condition for the monotonicity assumption to hold is to assume that x^{α} and y^{β} are the increasing functions (i.e., $\alpha > 0$ and $\beta > 0$). Using (17), if the DLT is attributed exclusively to drug A, then

$$P(\delta_A = 1, \delta_B = 0|x, y) = \pi^A = x^{\alpha}(1 - y^{\beta}) - \gamma(x, y). \tag{18}$$

If the DLT is attributed exclusively to drug B, then

$$P(\delta_A = 0, \delta_B = 1|x, y) = \pi^B = y^{\beta}(1 - x^{\alpha}) - \gamma(x, y). \tag{19}$$

If the DLT is attributed to both A and B, then

$$P(\delta_A = 1, \delta_B = 1|x, y) = \pi^{AB} = x^{\alpha} y^{\beta} + \gamma(x, y). \tag{20}$$

Equation (18) represents the probability that a DLT is caused only by drug A. This can happen, for example, when a type of DLT of taxotere (A), such as grade 4 neutropenia, is observed. However, this type of DLT can never be observed with metformin (B). This can also happen when the clinician attributes a grade 4 diarrhea to taxotere (A) but not to metformin (B) in the case of a low-dose level of this later even though both drugs have this common type of side effect. The fact that dose level y is present in Eq. (18) is a result of the joint modeling of the two marginals and accounts for the probability that drug B does not cause a DLT. This later case is, of course, based on the clinician's judgment. Equations (19) and (20) can be interpreted similarly.

The overall probability of DLT is calculated following [65] as the sum of (18), (19), and (20), which translates into

$$P(\text{DLT}|x, y) = \pi = x^{\alpha} + y^{\beta} - x^{\alpha} y^{\beta} - \gamma(x, y). \tag{21}$$

To calculate the MTD, re-write Eq. (1) as a second-degree polynomial in y^{β} and solve for the solutions. This allows us to define the MTD set C as

$$C = \left\{ (x_*, y_*) : y_* = \left[\frac{-(1 - x_*^{\alpha} - \kappa) \pm \sqrt{(1 - x_*^{\alpha} - \kappa)^2 - 4\kappa(x_*^{\alpha} - \theta)}}{2\kappa} \right]^{\frac{1}{\beta}} \right\}, \tag{22}$$

where

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

Among patients treated with dose combination (x, y) who exhibit DLT, suppose that an unknown fraction η of these patients have a DLT with known attribution, i.e., the clinician knows if the DLT is caused by drug A only, or drug B only, or both drugs A and B. Let $\mathscr{A}$ be the indicator of DLT attribution when $Z = 1$. It follows that for each patient treated with dose combination (x, y), there are five possible toxicity outcomes. This is illustrated in the chance tree diagram in Fig. 3. Using Eqs. (18), (19), (20), (21), and Fig. 3, we can define the contributions to the likelihood from each of the five observable outcomes as defined in Table 2.

The likelihood function is defined as

$$\mathscr{L}(\alpha, \beta, \gamma, \eta) = \prod_{i=1}^{n} \left[\left(\eta \pi_i^{(\delta_{A_i}, \delta_{B_i})} \right)^{\mathscr{A}_i} (\pi_i \, (1 - \eta))^{1 - \mathscr{A}_i} \right]^{Z_i} (1 - \pi_i)^{1 - Z_i}, \tag{23}$$

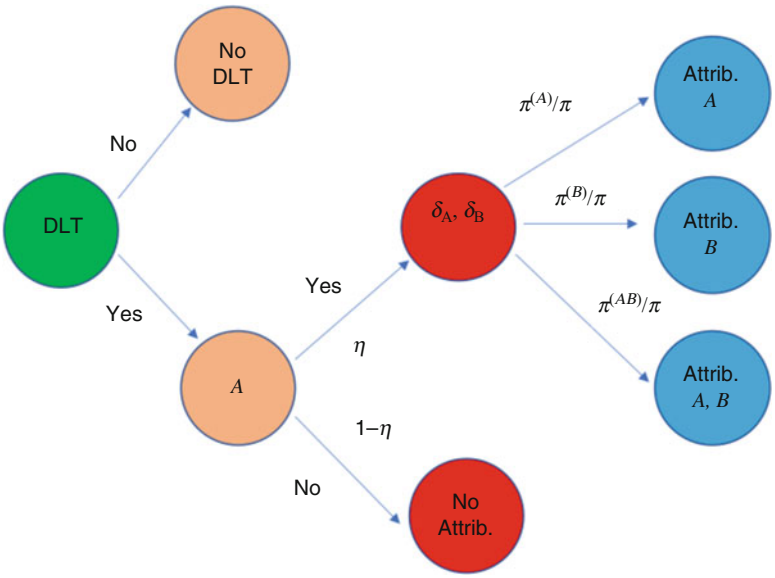

Fig. 3 A chance tree illustrating the 5 possible outcomes we can find in a trial

Table 2 Contributions to the likelihood function based on the observed outcomes: toxicity, attribution, attribution to drug A (δ_A), and attribution to drug B (δ_B) for each patient

Toxicity (Z)	Attribution ($\mathscr{A}$)	δ_A	δ_B	Likelihood
0	–	–	–	$1-\pi = 1-\left[x^{\alpha}+y^{\beta}-x^{\alpha}\times y^{\beta}-\gamma(x,y)\right]$
1	0	–	–	$\pi\times(1-\eta) = \left[x^{\alpha}+y^{\beta}-x^{\alpha}\times y^{\beta}-\gamma(x,y)\right]\times(1-\eta)$
1	1	1	0	$\pi\times\eta\times\frac{\pi^{(1,0)}}{\pi} = \eta\times\left[x^{\alpha}(1-y^{\beta})-\gamma(x,y)\right]$
1	1	0	1	$\pi\times\eta\times\frac{\pi^{(0,1)}}{\pi} = \eta\times\left[y^{\beta}(1-x^{\alpha})-\gamma(x,y)\right]$
1	1	1	1	$\pi\times\eta\times\frac{\pi^{(1,1)}}{\pi} = \eta\times\left[x^{\alpha}y^{\beta}+\gamma(x,y)\right]$

and the joint posterior probability distribution of the model parameters as

$$P(\alpha,\beta,\gamma,\eta|x,y,\delta_A,\delta_B) \propto P(\alpha)P(\beta)P(\gamma)P(\eta)\times\mathscr{L}(\alpha,\beta,\gamma,\eta). \tag{24}$$

Using Eq. (24), we can easily sample and obtain MCMC estimates of α, β, γ, and η.

Trial Design

Dose escalation/de-escalation proceeds using the following modified univariate continual reassessment method (CRM) [32] described in Sect. 2.1:

1. Each patient in the first cohort of 2 patients receives the same dose combination $(x_i, y_i) = (0, 0)$ for $i = 1, 2$.
2. In the i-th cohort of 2 patients, for $i \geq 2$,
   (a) If i is even, then patient $2i-1$ receives dose (x_{2i-1}, y_{2i-3}) and patient $2i$ receives dose (x_{2i-2}, y_{2i}), where

$$x_{2i-1} = \underset{x}{\text{argmin}}\left|\widehat{\text{Prob}}(Z=1|x, y_{2i-3})-\theta\right|,$$

$$y_{2i} = \underset{y}{\text{argmin}}\left|\widehat{\text{Prob}}(Z=1|x_{2i-2}, y)-\theta\right|.$$

   If a DLT was observed in the previous cohort of two patients and was attributable to drug A, then x_{2i-1} is further restricted to be no more than x_{2i-3}. On the other hand, if a DLT was observed in the previous cohort of two patients and was attributable to drug B, then y_{2i} is further restricted to be no more than y_{2i-2}.
   (b) If i is odd, then patient $2i-1$ receives dose (x_{2i-3}, y_{2i-1}) and patient $2i$ receives dose (x_{2i}, y_{2i-2}), where

$$x_{2i} = \underset{x}{\text{argmin}}\left|\widehat{\text{Prob}}(Z=1|x, y_{2i-2})-\theta\right|$$

$$y_{2i-1} = \underset{y}{\text{argmin}}\left|\widehat{\text{Prob}}(Z=1|x_{2i-3}, y)-\theta\right|.$$

If a DLT was observed in the previous cohort of two patients and was attributable to drug A, then x_{2i} is further restricted to be no more than x_{2i-2}. On the other hand, if a DLT was observed in the previous cohort of two patients and was attributable to drug B, then y_{2i-1} is further restricted to be no more than y_{2i-3}.

3. Repeat step 2 until n patients are enrolled in the trial subject to a safety stopping rule in which the trial is stopped if the estimated probability of DLT at the lowest dose combination is higher than a pre-specified threshold.

Results

An extensive simulation study is performed by Jimenez et al. [22]. For illustration purposes, in Fig. 4, we present the results of one scenario that illustrates the main conclusion of this chapter.

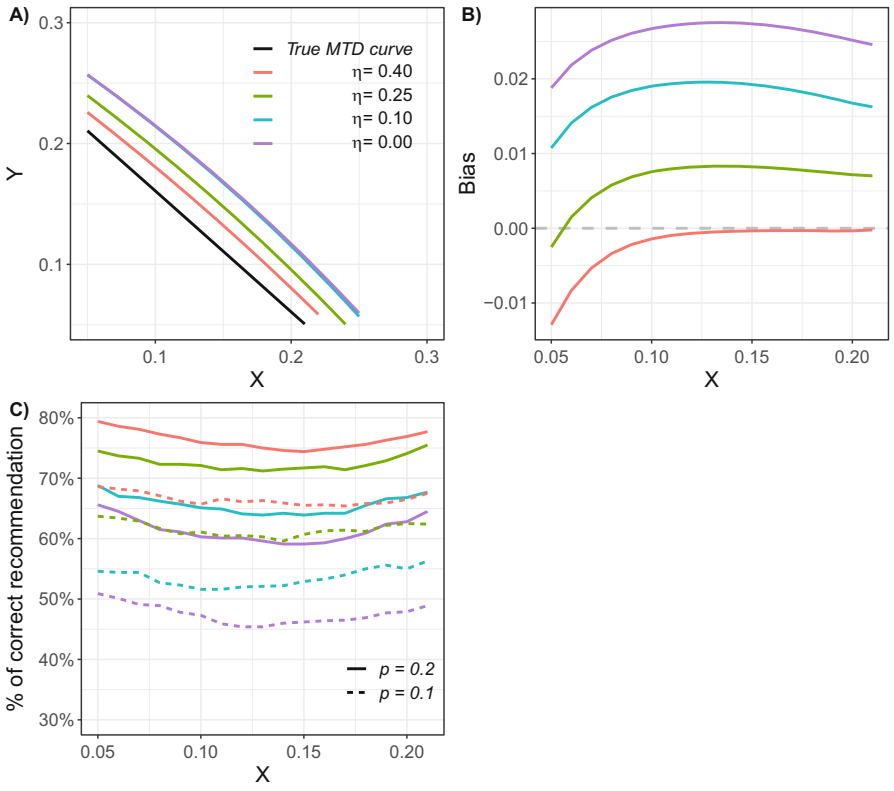

Fig. 4 Simulated scenario taken from [22]. In (A), we show the true and estimated MTD curves as defined in Eq. (22) as well as each final recommended dose combination after simulating 1000 trials for different levels of toxicity attribution. In (B) and (C), we observe the bias and the percentage of correct recommendation, respectively, for each value of X contained in the MTD curve

Table 3 Safety results from the simulated scenario presented in Fig. 4 from [22]

η	Average % of toxicities	% of trials with DLT rate $> \theta + 0.05$	% of trials with DLT rate $> \theta + 0.10$
0.00	33.62	25.90	4.10
0.10	32.67	22.60	4.80
0.25	31.55	17.60	2.70
0.40	30.70	13.30	2.00

Jimenez et al. [22] evaluate the effect of toxicity attribution in several scenarios assuming proportions of attributed DLTs of 0%, 10%, 25% and 40% (i.e., $\eta = \{0, 0.1, 0.25, 0.4\}$). These values are reasonable because higher values of η in practice are very rare. In general, increasing the value of η increases the pointwise percent of MTD recommendation and reduces bias. The approach of partial toxicity attribution generates safe trial designs, as presented in Table 3, and efficient estimation of the MTD. Further details about the approach and computer codes can be found in [22].

2.3 Adding a Baseline Covariate

Although chemotherapy and radiotherapy are still the main cancer treatments for tumors after surgical excision, these conventional therapies may be combined with targeted agents to enhance treatment efficacy. Traditional drug combination designs as presented in the previous section assume that the patient population is homogeneous of treatment tolerance. Therefore, a design that specifies the dose-toxicity relationship given a baseline covariate that indicates when a patient is more susceptible to a given targeted agent is desirable for drug combinations.

Model

Diniz et al. [10] proposed a parametric model to identify tolerable dose combinations of two synergistic drugs A and B given a patient with a binary baseline covariate with value w. Assuming the same notation used along this chapter, the proposed model is defined as

$$P(Z = 1|x, y, w) = F(\beta_0 + \beta_1 x + \beta_2 y + \beta_3 xy + \beta_4 w). \tag{25}$$

The MTD for a patient with covariate value w is defined as the set of combinations (x^*, y^*) such that

$$C = \left\{ (x^*, y^*) \in S : y^* = \frac{F^{-1}(\theta) - \beta_0 - \beta_1 x^* - \beta_4 w}{\beta_2 + \beta_3 x^*} \right\}. \tag{26}$$

The model (25) is reparameterized to allow a more meaningful prior elicitation defining ρ_{000} as the probability of DLT when the level of drugs A and B is minimum, and $w = 0$; ρ_{100} as the probability of DLT when the level of drug A is maximum, the level of drug Y is minimum and $w = 0$; ρ_{101} as the probability of DLT when the level of drug X is maximum, the level of drug B is minimum and $w = 1$; ρ_{010} as the probability of DLT when the level of drugs A is minimum, the level of drug B is maximum, and $w = 0$. Then, it follows that

$$\begin{aligned}
\beta_0 &= F^{-1}(\rho_{000}) \\
\beta_1 &= F^{-1}(\rho_{100}) - F^{-1}(\rho_{000}) \\
\beta_2 &= F^{-1}(\rho_{010}) - F^{-1}(\rho_{000}) \\
\beta_3 &= \eta \\
\beta_4 &= F^{-1}(\rho_{101}) - F^{-1}(\rho_{100}).
\end{aligned} \tag{27}$$

The MTD set defined in (26) can be written as

$$C = \left\{ (x^*, y^*) \in S : y^* = \frac{G(\theta, \rho_{000}) - (G(\rho_{100}, \rho_{000}))x^* - (G(\rho_{101}, \rho_{100}))w}{G(\rho_{010}, \rho_{000}) + \eta x^*} \right\}, \tag{28}$$

where $G(a, b) = F^{-1}(a) - F^{-1}(b)$.

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

$$\begin{aligned}
\mathscr{L}(\rho_{000}, \rho_{100}, \rho_{101}, \rho_{010}, \eta | D_n) = \prod_{i=1}^{n} &(H(\rho_{000}, \rho_{100}, \rho_{101}, \rho_{010}, \eta; x_i, y_i, w_i))^{\delta_i} \\
&\times (1 - H(\rho_{000}, \rho_{100}, \rho_{101}, \rho_{010}, \eta; x_i, y_i, z_i))^{1-\delta_i},
\end{aligned} \tag{29}$$

where

$$\begin{aligned}
&H(\rho_{000}, \rho_{100}, \rho_{101}, \rho_{010}, \eta; x_i, y_i, z_i) \\
&\quad = F(F^{-1}(\rho_{000}) + (F^{-1}(\rho_{100}) - F^{-1}(\rho_{000}))x_i + (F^{-1}(\rho_{010}) - F^{-1}(\rho_{000}))y_i \\
&\qquad + (F^{-1}(\rho_{101}) - F^{-1}(\rho_{100}))w_i + \beta_3 x_i y_i).
\end{aligned} \tag{30}$$

Prior and Posterior Distributions

Diniz et al. [10] consider the priors $\rho_{100} \sim \text{beta}(a_1, b_1)$, $\rho_{010} \sim \text{beta}(a_3, b_3)$, $\rho_{101} \sim \text{beta}(a_2, b_2)$, $\rho_{000}/min(\rho_{100}, \rho_{010}) \sim \text{beta}(a_0, b_0)$, and $\eta \sim \text{gamma}(a, b)$

with mean $E(\eta) = a/b$ and variance $Var(\eta) = a/b^2$. See Sect. 2.1 for the definition of the hyperparameter values of each distribution. The posterior distribution is given by

$$\begin{aligned} P(\rho_{000}, \rho_{100}, \rho_{101}, \rho_{010}, \eta | D_n) \propto & \prod_{i=1}^{n} (H(\rho_{000}, \rho_{100}, \rho_{101}, \rho_{010}, \eta; x_i, y_i, w_i))^{Z_i} \\ & \times (1 - H(\rho_{000}, \rho_{100}, \rho_{101}, \rho_{010}, \eta; x_i, y_i, w_i))^{1-Z_i} \\ & \times P(\rho_{000} | \rho_{100}, \rho_{010}) P(\rho_{100}) P(\rho_{101}) P(\rho_{010}) P(\eta). \end{aligned} \tag{31}$$

Trial Design

The algorithm for dose escalation/de-escalation is similar to the one discussed in Sect. 2.1 with the additional binary covariate information. It uses the EWOC principle [1] where at each stage of the trial, we seek a dose of one agent using the current posterior distribution of the MTD of the agent given the current dose of the other agent and the next patient's baseline covariate value. Specifically, for the i-th cohort of two patients, the design proceeds as follows:

1. If i is even, patient $(2i - 1)$ receives dose (x_{2i-3}, y_{2i-1}) and patient $2i$ receives dose (x_{2i}, y_{2i-2}), where $y_{2i-1} = \Pi^{-1}_{\Gamma_{B|A=x_{2i-3}, W=w_{2i-1}}}(\alpha | D_{2i-2})$ and $x_{2i} = \Pi^{-1}_{\Gamma_{A|B=y_{2i-2}, Z=z_{2i}}}(\alpha | D_{2i-2})$. Here, $\Pi^{-1}_{\Gamma_{A|B=y, W=w}}(\alpha | D)$ is the inverse cumulative distribution function of the posterior distribution, $\pi(\Gamma_{A|B=y, Z=z} | D)$.
2. Similarly, if i is odd, patient $(2i - 1)$receives dose (x_{2i-1}, y_{2i-3}) and patient $2i$ receives dose (x_{2i-2}, y_{2i}), where $x_{2i-1} = \Pi^{-1}_{\Gamma_{A|B=y_{2i-3}, W=w_{2i-1}}}(\alpha | D_{2i-2})$ and $y_{2i} = \Pi^{-1}_{\Gamma_{B|A=x_{2i-2}, W=w_{2i}}}(\alpha | D_{2i-2})$.

As described in Sect. 2.1, dose escalation is further restricted to be no more than a pre-specified fraction of the dose range of the corresponding agent as well as stopping rules.

At the completion of the trial, an estimate of the MTD curve for $w = 0, 1$ is obtained using Eq. (28) as

$$\hat{C} = \left\{ (x^*, y^*) \in S : y^* = \frac{G(\theta, \hat{\rho}_{000}) - (G(\hat{\rho}_{100}, \hat{\rho}_{000}))x^* - (G(\hat{\rho}_{101}, \hat{\rho}_{100}))w}{G(\hat{\rho}_{010}, \hat{\rho}_{000}) + \hat{\beta}_3 x^*} \right\}, \tag{32}$$

where $G(a, b) = F^{-1}(a) - F^{-1}(b)$, and $\hat{\rho}_{000}$, $\hat{\rho}_{100}$, $\hat{\rho}_{101}$, $\hat{\rho}_{010}$, and $\hat{\beta}_3$ are the posterior medians given the data D_n.

Table 4 Safety results from the simulated scenario presented in Fig. 5 from [10]

Covariate (W)	Average % of toxicities	% of trials with DLT rate $> \theta + 0.05$	% of trials with DLT rate $> \theta + 0.10$
Overall	30.70	4.80	0.40
0	22.10	0.60	0.10
1	39.40	58.80	32.80

Results

In [10] several scenarios we evaluated, including a comparison between including and not including a baseline covariate in parallel trials. We illustrate the design for drug combination with a baseline covariate using a simulation study with 1000 trials. Dose escalation proceeds following EWOC criterion with the target toxicity rate $\theta = 0.33$, and the feasibility boundary α starts at 0.25 with an increment of 0.05 for each new cohort of patients up to 0.5. Cohorts of two patients were accrued with the total sample size of 40 patients such that two sub-groups of 20 patients randomly accrued over each trial.

Table 4 shows safety operating characteristics indicating that the proposed design is able to control the overall average DLT, with higher overdose for patients with $W = 1$ because they are more susceptible, i.e., their MTD curve is closer to the minimum dose. Figure 5A shows the estimated MTD for both sub-groups, with Fig. 5B indicating increasing bias at the edges of the MTD curves, but still with negligible absolute values. The percentage of correct recommendation in Fig. 5C displays high values for both tolerances $p = 0.1, 0.2$ when $W = 0$ and only $p = 0.2$ when $W = 1$.

3 Designs for Phase I–II Clinical Trials

3.1 Binary Endpoint

Let $\widehat{C}$ be the estimated MTD curve obtained in Eq. (11) and suppose it is defined for $(x, y) \in [X_1, X_2] \times [Y_1, Y_2] \subset [X_{\min}, X_{\max}] \times [Y_{\min}, Y_{\max}]$. Let E be the indicator of treatment response, $E = 1$ if we have a positive response, and $E = 0$ otherwise. Let p_0 be the probability of efficacy of the standard-of-care treatment. The goal of the stage II trial is to identify dose combinations $(x, y) \in \widehat{C}$ such that $P(E = 1|(x, y)) > p_0$.

Model

Tighiouart [47] models the probability of response by first mapping dose combinations on $\widehat{C}$ to [0, 1] as follows. For $(x, y) \in \widehat{C}$, let x be the vertical projection of

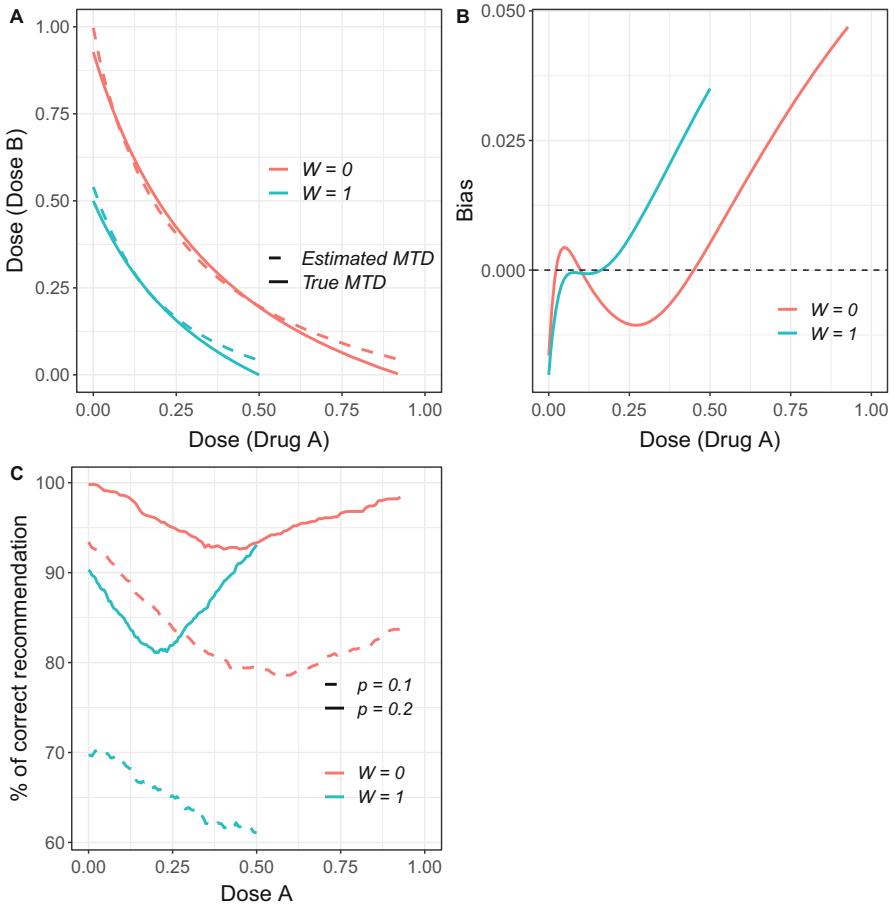

Fig. 5 Simulated scenario taken from [10], $(\rho_{000}, \rho_{100}, \rho_{010}, \rho_{101}, \eta) = (0.01, 0.40, 0.40, 0.80, 10)$. In (A), we show the true and estimated MTD curves as defined in Eq. (28) as well as each final recommended dose combination after simulating 1000 trials. In (B) and (C), we observe the bias and the percentage of correct recommendation, respectively, for each value of drug A contained in the MTD curve

(x, y) on the interval $[X, Y]$ and $z = h(x) = (x - X)/(Y - X)$. z can be considered as a dose combination since there is a one-to-one transformation mapping $z \in [0, 1]$ to $(x, y) \in \widehat{C}$. Let

$$P(E = 1|z, \boldsymbol{\psi}) = F(f(z; \boldsymbol{\psi})) \tag{33}$$

be the probability of efficacy, where F is a known link function, $f(z; \boldsymbol{\psi})$ is unknown, and $\boldsymbol{\psi}$ is an unknown parameter. A flexible way to model the probability of efficacy along the MTD curve is the cubic spline function

$$f(z;\boldsymbol{\psi}) = \beta_0 + \beta_1 z + \beta_2 z^2 + \sum_{j=3}^{k} \beta_j (z-\kappa_j)_+^3, \tag{34}$$

where $\boldsymbol{\psi} = (\boldsymbol{\beta}, \boldsymbol{\kappa})$, $\boldsymbol{\beta} = (\beta_0, \ldots, \beta_k)$, $\boldsymbol{\kappa} = (\kappa_3, \ldots, \kappa_k)$ with $\kappa_3 = 0$. Let $D_m = \{(z_i, E_i), i = 1, \ldots, m\}$ be the data after enrolling m patients in the trial, where E_i is the response of the i-th patient treated with dose combination z_i and let $\pi(\boldsymbol{\psi})$ be a prior density on the parameter $\boldsymbol{\psi}$. The posterior distribution is

$$\pi(\boldsymbol{\psi}|D_m) \propto \prod_{i=1}^{m} [F(f(z_i;\boldsymbol{\psi}))]^{E_i} [1 - F(f(z_i;\boldsymbol{\psi}))]^{1-E_i} \pi(\boldsymbol{\psi}). \tag{35}$$

Trial Design

This stage of the trial makes use of response-adaptive randomization to allocate patients to dose combinations that are likely to have high probability of efficacy. Let p_z be the probability of efficacy at dose combination z and p_0 be the probability of a treatment not worthy of further investigation. To test the hypothesis

$$\mathrm{H}_0 : p_z \leq p_0 \text{ for all } z \text{ versus } \mathrm{H}_1 : p_z > p_0 \text{ for some } z,$$

we enroll n patients in the trial according to the following design:

1. The first n_1 patients are randomly assigned to dose combinations $z_1, \ldots, z_{n_1}$ equally spaced along the MTD curve C_{est}.
2. Update the posterior in (35) and obtain a Bayes estimate $\hat{\boldsymbol{\psi}}$.
3. Generate n_2 dose combinations from the standardized density $F(f(z;\hat{\boldsymbol{\psi}}))$ and assign them to the next cohort of n_2 patients.
4. Repeat steps (2) and (3) until n patients have been enrolled subject to pre-specified stopping rules.

This algorithm can be viewed as an extension of a Bayesian adaptive design to select a superior arm among a finite number of arms [3] to selecting a superior arm from an infinite number of arms.

Decision Rule At the end of the trial, accept H_1 if

$$\mathrm{Max}_z[P(F(f(z;\boldsymbol{\psi})) > p_0|D_n)] > \delta_u, \tag{36}$$

where δ_u is a design parameter.

Stopping Rules For ethical considerations and to avoid exposing patients to sub-therapeutic doses, the trial may be stopped for futility after j patients are evaluable for efficacy if there is strong evidence that none of the dose combinations are promising, i.e., $\mathrm{Max}_z[P(F(f(z;\boldsymbol{\psi})) > p_0|D_j)] < \delta_0$, where δ_0 is a small pre-

specified threshold. In cases where the investigator is interested in stopping the trial early for superiority, the trial may be terminated after j patients are evaluable for efficacy if $\text{Max}_z[P(F(f(z;\boldsymbol{\psi})) > p_0|D_j)] > \delta_1$, where $\delta_1 \geq \delta_u$ is a pre-specified threshold and the corresponding dose combination $z^* = \text{argmax}_u\{P(F(f(u;\boldsymbol{\psi})) > p_0|D_j)\}$ is selected for future randomized phase II or III studies.

Results

Performance of this design depends on a number of parameters including the sample size n, the probability of a poor treatment efficacy p_0, design parameter δ_u, and desired effect size. Using extensive simulations, [47] showed that this phase 2 response-adaptive design has good operating characteristics using sample sizes and effect sizes comparable to single-arm phase 2 trials with one dose level. For illustration purpose, we provide in Fig. 6 the 6 scenarios presented in [47], where scenarios A–C favor the null hypothesis and scenarios D–F favor the alternative hypothesis (see Sect. 3.1). These were used to derive the operating characteristics of a combination trial cisplatin–cabazitaxel in advanced prostate cancer patients with clinical benefit as the treatment response. The probability of a poor treatment response is $p_0 = 0.15$ and the effect size is 0.25. Thirty patients were enrolled in stage 2 following 30 patients in stage 1. Scenarios A–C have estimated powers of 0.896, 0.921, and 0.81, respectively. Scenarios D–F have estimated type-I error probabilities of 0.1, 0.19, and 0.143, respectively. Additional results such as average bias and percentage of correct recommendation, and safety for stage 1 are presented in [47] as well as in its supplementary material. These results allow to conclude that

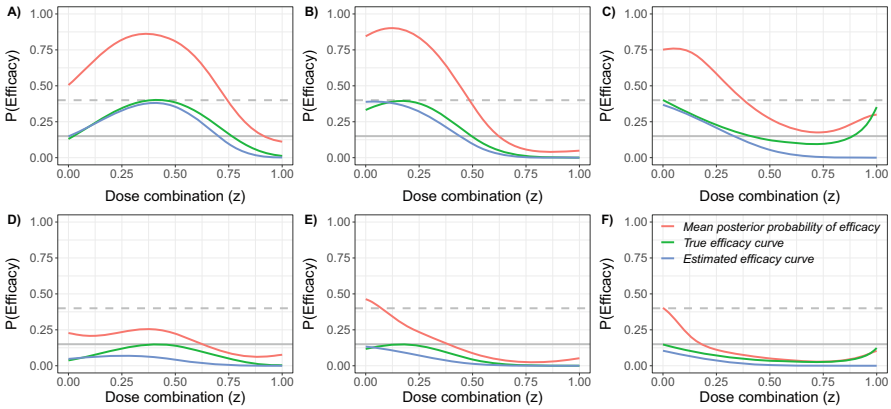

Fig. 6 True efficacy curves, mean posterior probability of efficacy curves, and estimated efficacy curves for different dose combinations (z) in 6 scenarios under H_0 and H_1. The grey-solid lines represent the null probability of efficacy p_0 (i.e., the probability of a poor treatment efficacy) and the grey-dashed lines represent the target probability of efficacy (i.e., the effect size)

this design has, in general, good operating characteristics allowing to identify the dose combination that maximizes the efficacy.

3.2 Survival Endpoint

Introduction

In this section, we review the work of Jiménez et al. [21] that extends the methodology in [47] from binary efficacy endpoint to time-to-event endpoint.

Model

Jiménez et al. [21] model the time to progression as a Weibull distribution with probability density function

$$f(t; z) = \frac{k}{\lambda(z; \boldsymbol{\psi})} \left(\frac{t}{\lambda(z; \boldsymbol{\psi})} \right)^{k-1} \exp\left(-\frac{t}{\lambda(z; \boldsymbol{\psi})} \right)^{k}, \tag{37}$$

where $\lambda > 0$ is the shape parameter and $k > 0$ is the scale parameter.

The median TTP is

$$\text{Med}(z) = \lambda(z; \boldsymbol{\psi})(\log 2)^{\frac{1}{k}}. \tag{38}$$

A flexible way of modeling the median TTP along the MTD curve is through the use of the cubic spline function

$$\lambda(z; \boldsymbol{\psi}) = \exp\left(\beta_0 + \beta_1 z + \beta_2 z^2 + \sum_{j=3}^{5} \beta_j (z - \phi_j)_+^3 \right), \tag{39}$$

where $\boldsymbol{\psi} = (\boldsymbol{\beta}, \boldsymbol{\phi})$, with $\boldsymbol{\beta} = (\beta_0, \ldots, \beta_5)$ and $\boldsymbol{\phi} = (\phi_3, \ldots, \phi_5)$, being $\phi_3 = 0$. Let $D_n = \{(z_i, t_i, \delta_i), i = 1 \ldots, n\}$ be the data after enrolling n patients in the trial where t represents the TTP or last follow-up, and δ the censoring status, and let $\pi(\boldsymbol{\psi}, k)$ be the joint prior density on the parameter vectors $\boldsymbol{\psi}$ and k. The posterior distribution is

$$\pi(\boldsymbol{\psi}, k | D_m) \propto \pi(\boldsymbol{\psi}, k) \prod_{i=1}^{n} \left[\frac{k}{\lambda(z_i; \boldsymbol{\psi})} \left(\frac{t_i}{\lambda(z_i; \boldsymbol{\psi})} \right)^{k-1} \right]^{\delta_i} \times \exp\left(-\frac{t_i}{\lambda(z_i; \boldsymbol{\psi})} \right)^{k}. \tag{40}$$

Let Med_z be the median TTP at dose combination z, and let Med_0 be the median TTP of the standard-of-care treatment. We propose an adaptive design in order to test the hypothesis

$$\begin{aligned} &\text{H}_0 : \text{Med}_z \leq \text{Med}_0 \text{ for all } z \quad \text{vs.} \\ &\text{H}_1 : \text{Med}_z > \text{Med}_0 \text{ for some } z. \end{aligned} \tag{41}$$

It is important to keep in mind that the reason why [21] propose a model with a fairly large number of parameters is because they work in a continuous dose space. In a discrete dose space, it is not common to test so many dose combinations. Also, a model with a large number of parameters would most likely be non-identifiable, even with large sample sizes. The use of continuous dose combinations is not uncommon in dose-finding studies since the drugs are administered intravenously and this allows to administer any drug concentration we desire.

Trial Design

This stage of the trial makes use of response-adaptive randomization to decide in which dose combinations cohorts of patients are allocated. The algorithm is similar to the one discussed in Sect. 3.1 with the difference that in this one [21] work with time-to-event data:

1. We first treat n_1 patients at dose combinations $x_1, \ldots, x_{n_1}$, which are equally spaced along the estimated maximum tolerated dose combination curve C_{est}.
2. Obtain posterior distribution of $\boldsymbol{\psi}$ and k given the data D_{n_1} using Eq. (40). Note that prior to obtaining the posterior distribution of the model parameters, patients who have not progressed are right censored.
3. Generate n_2 dose combinations from the standardized density $\text{Med}(z) = \lambda(z; \boldsymbol{\psi})(\log 2)^{\frac{1}{k}}$, and assign them to the next n_2 patients.
4. Repeat steps 2 and 3 until a total of n patients have been enrolled in the trial subject to pre-specified stopping rules.

Decision Rule: At the end of the trial, we reject the null hypothesis if $\text{Max}_z\{P(\text{Med}(z; \boldsymbol{\psi}_i) > \text{Med}_0 | D_{n,i})\} > \delta_u$, where δ_u is a design parameter.

Stopping Rule (Safety): For a prospective trial, a separate stopping rule for safety using, for example, a Bayesian continuous monitoring for toxicity (see e.g., [66]) should be implemented as discussed in [47].

Stopping Rule (Futility): For ethical reasons and to avoid treating patients at sub-therapeutic dose levels, we will stop the trial for futility if there is strong evidence that none of the dose combinations are promising, i.e., $\text{Max}_z\{P(\text{Med}(z; \boldsymbol{\psi}_i) > \text{Med}_0 | D_{n,i})\} < \delta_0$, where δ_0 is a design parameter.

Stopping Rule (Efficacy): For ethical reasons, if the investigator considers there is enough evidence in favor of one or more dose combinations being

tested, and no further patients need to be enrolled, the trial can be terminated if $\text{Max}_z\{P(\text{Med}(z;\boldsymbol{\psi}_i) > \text{Med}_0|D_{n,i})\} > \delta_1$, where $\delta_1 \geq \delta_u$ is a study parameter and the dose combination $z^{\text{opt}} = \arg\max_v\{P(\text{Med}(v;\boldsymbol{\psi}_i) > \text{Med}_0|D_{n,i})\}$ is selected for further randomized phase II or phase III clinical trials.

The rational for this approach is based on the rejection-sampling principle, which can be used to generate observations from a target distribution (in our case (38)). Hence, if we generate data from (38), we will be allocating patients to dose combinations that are more likely to have higher TTP according to the current estimation of (38) (i.e., the shape of (38) will be updated as patients enroll).

Results

An extensive simulation with several scenarios was performed by Jiménez et al. [21]. For illustration purposes, in Fig. 7, we show one scenario that summarizes the main conclusions of this chapter.

In Fig. 7A, we show the dose–efficacy relationship within the MTD curves selected in stage 1. For this particular case, this curve represents a scenario where high levels of drug Y and low levels of X are more efficacious. In Fig. 7B, we observe how the proposed design identifies lower levels of z, which represents high levels of drug Y and low levels of X as the more efficacious dose combinations.

In Table 5, we observe the corresponding Bayesian power, type-I error probability, and type-I + type-II error probability with effect sizes of 1.5 and 2 months and accrual rates of 1 and 2 patients per month. Additional results are presented in the supplementary material of [21] such as average bias and percentage of correct recommendation. These results allow to conclude that this design has, in general, good operating characteristics allowing to identify the dose combination that maximizes the efficacy.

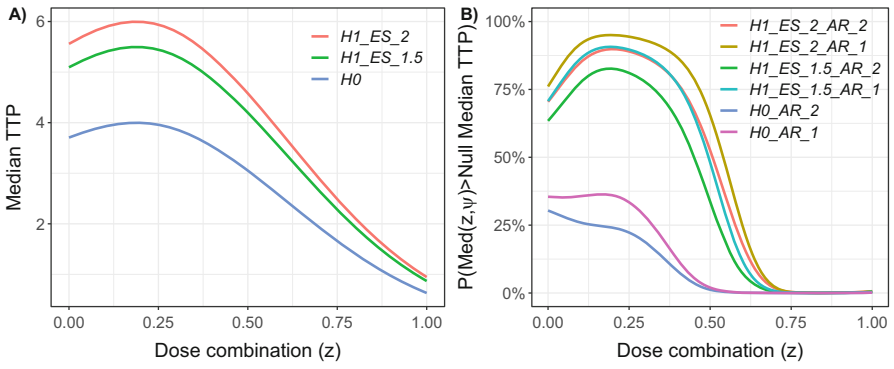

Fig. 7 Median TTP (A) and posterior probability of having Med(z, ψ) >null Median TTP (B) for different dose combinations (z) under H_0 and H_1, with effect sizes (ES) of 1.5 and 2 months and accrual rates (AR) of 1 and 2 patients per month

Table 5 Bayesian power, type-I error probability, and type-I + type-II error probability with effect sizes of 1.5 and 2 months and accrual rates of 1 and 2 patients per month

	Power (effect size of 1.5 months)		Power (effect size of 2 months)		Probability of type-I error		Probability of type-I + type-II errors (effect size of 1.5 months)		Probability of type-I + type-II errors (effect size of 2 months)	
	δ_u		δ_u		δ_u		δ_u		δ_u	
Accrual rate	0.8	0.9	0.8	0.9	0.8	0.9	0.8	0.9	0.8	0.9
1	0.924	0.844	0.971	0.927	0.227	0.121	0.303	0.277	0.256	0.194
2	0.824	0.674	0.920	0.829	0.107	0.048	0.283	0.374	0.187	0.219

4 Discussion

The use of drug combinations in early phase cancer clinical trials has been extensively studied over the last decade. The overall goal of early phase clinical trials in oncology is to find a set of one or more safe dose combinations that maximize efficacy. To achieve this goal, we propose that, in the first step, a phase I trial is designed to identify one or more maximum tolerated doses (MTDs). Following this step, a phase II trial is designed to search for a combination that maximizes efficacy within the set of MTDs. It is worth noting that the main objective of the majority of phase I designs is to identify a single MTD. We recommend the use of designs that select more than one MTD for efficacy trials as this may result in less failed phase II trials.

In this chapter, we focused on dose-finding methods tailored for drug combinations with continuous dose levels. The use of continuous dose levels is very common in clinical oncology research, especially in early phase trials where the existing or first-in-human drugs are delivered as infusions intravenously. In addition, discretizing the dose levels may lead to a recommended phase 2 dose that has either a small or high probability of DLT relative to the target risk of DLT, if the true MTD is not part of the discrete set of doses. As showed by Diniz et al. [13], continuous dose schemes generally have equal or better safety and efficiency results than the discrete dose schemes, although alternative approaches to improve efficiency of discrete dose schemes have been recently published where new doses are added during the trial into the original discrete set [18]. In cases where there is no information about the location of the MTD, a continuous dose scheme would certainly be much more appealing. Furthermore, although the seminal CRM design and several related dose-finding methods are based on regression models, their dose escalation algorithms are based on pre-specified skeletons to incorporate prior information, which cannot be adapted to continuous doses.

In phase I trial designs, consecutive cohorts of two patients were treated simultaneously with different dose combinations to better explore the space of doses. The method was studied extensively by Tighiouart et al. [50, 51, 53] under the EWOC criterion and by Diniz et al. [12] and Jimenez et al. [22] using the

CRM principle via simulations. Comparisons of EWOC and CRM in the settings of dichotomous DLTs and ordinal toxicity grades can be found in [11].

Most drug combination designs assume that the binary DLT is attributable to either one of the drugs or both. This is a reasonable assumption because of the rarity of cancer drugs with non-overlapping toxicities of any grade. However, certain combinations may lead to some non-overlapping toxicities. For instance, when combining taxotere with metformin, an occurrence of a grade 4 neutropenia can only be attributed to taxotere and not to metformin. This event will guide the clinician to hold the current dose of metformin and decrease the taxotere dose for the next cohort of patients, even if the statistical algorithm recommends a dose decrease for both agents. We described the work developed by Jimenez et al. [22], where a clinician can attribute the DLT to one or more drugs in an unknown fraction of attributable DLTs by extending the previous statistical models. This is useful in a situation where the two drugs do not have many overlapping toxicities (see, e.g. [28]). However, it is also important to note that this method relies on clinical judgment regarding DLT attribution.

Another approach reviewed in this chapter is the inclusion of a baseline covariate to estimate patient-specific MTD curves [10]. We found that in the presence of a clinically significant baseline covariate, the design with a covariate had a smaller pointwise average bias and a higher percent of MTD recommendation relative to a design that ignores the covariate. Moreover, we stand to lose little in terms of safety of the trial and efficiency of the estimated MTD curve, if we include a practically not important covariate in the model.

In the second part of this chapter, we described two-stage designs developed by Tighiouart [47] and Jiménez et al. [21] where the estimated MTD curve from a phase I trial is used as input to a phase II efficacy trial using Bayesian adaptive randomization. Two-stage designs are required when it takes several cycles of therapy to resolve treatment efficacy or patient characteristics in phases I and II are clinically different. For instance, efficacy in the cisplatin–cabazitaxel trial that was described in [47] is resolved after three cycles of treatment, and patients in stage I must have metastatic, castration resistant prostate cancer, whereas patients in stage II must have visceral metastasis. As mentioned in these articles, these designs can be viewed as an extension of the Bayesian adaptive design comparing a finite number of arms [3] to that with an infinite number of arms. In particular, when the dose levels of the two agents are discrete, methods such as the ones described in [45, 59, 62] can be used to identify a set of MTDs in stage I, and the trial in stage II can select the most efficacious dose by adaptive randomization. One limitation of these two-stage approaches is that uncertainty of the estimated MTD curve in stage I is not taken into account in stage II of the design, which implies that the MTD curve is not updated as a result of observing DLTs in stage II. However, this problem is also inherent to single-agent two-stage designs where the MTD from the phase I trial is used in phase II studies. In both cases, safety is monitored continuously during the second stage of the design. A potential alternative design would account for first-, second-, and third-cycle DLT in addition to efficacy outcome at each cycle. In addition, the nature of DLT (reversible vs. non-reversible) should be taken into

account since patients with a reversible DLT are usually treated for that side effect and kept in the trial with dose reduction in subsequent cycles. These topics are the subject of future research.

Successful implementation of these designs requires active involvement and collaboration between the clinicians and the biostatisticians in many situations. This includes the design stage, prior distribution calibration, specification of scenarios with various locations of the true MTD set of doses or safe and efficacious doses, and computations of sequential posterior probabilities for dose allocation. This process may be challenging since it requires special expertise of biostatisticians who can program MCMC algorithms, adapt the existing computer codes to their trial, and modify them as needed since every trial is unique. The process is also time-consuming at the design stage to derive the operating characteristics. An R package `EWOC2` for designing the trials in [51] can be found in [25], and R codes for deriving the operating characteristics of the trials in [21, 22, 47] can be found in the supplementary material of the corresponding journal web site. An application of the phase I–II design in Sect. 3.1 is described in [47] where the clinician Dr. Posadas worked with Dr. Tighiouart in calibrating the prior distributions of the model parameters of the phase I part using preliminary data from a similar phase I trial, using the same combination of cabazitaxel and cisplatin. Operating characteristics were derived based on scenarios elicited by the clinician regarding the location of the true MTD curve and expected clinical benefit rate. Other recent applications of these methods for single-agent trials were designed by Drs. Tighiouart and Rogatko and published in [36, 57].

References

1. Babb, J., Rogatko, A., Zacks, S.: Cancer phase I clinical trials: efficient dose escalation with overdose control. Stat. Med. **17**(10), 1103–1120 (1998)
2. Babb, J.S., Rogatko, A.: Patient specific dosing in a cancer phase I clinical trial. Stat. Med. **20**(14), 2079–2090 (2001)
3. Berry, S.M., Carlin, B.P., Lee, J.J., Muller, P.: Bayesian Adaptive Methods for Clinical Trials. CRC Press, New York (2010)
4. Braun, T.M.: The bivariate continual reassessment method: extending the CRM to phase I trials of two competing outcomes. Contemp. Clin. Trials **23**(3), 240–256 (2002)
5. Braun, T.M., Wang, S.: A hierarchical Bayesian design for phase I trials of novel combinations of cancer therapeutic agents. Biometrics **66**(3), 805–812 (2010)
6. Cai, C., Yuan, Y., Ji, Y.: A Bayesian dose finding design for oncology clinical trials of combinational biological agents. Appl. Stat. **63**, 159–173 (2014)
7. Chen, Z., Yuan, Y., Li, Z., Kutner, M., Owonikoko, T., Curran, W.J., Khuri, F., Kowalski, J.: Dose escalation with over-dose and under-dose controls in phase I/II clinical trials. Contemp. Clin. Trials **43**, 133–141 (2015)
8. Chen, Z., Zhao, Y., Cui, Y., Kowalski, J.: Methodology and application of adaptive and sequential approaches in contemporary clinical trials. J. Probab. Stat. **2012**, 20 (2012). https://doi.org/10.1155/2012/527351
9. Cheung, Y.K., Chappell, R.: Sequential designs for phase I clinical trials with late-onset toxicities. Biometrics **56**(4), 1177–1182 (2000)

10. Diniz, M.A., Kim, S., Tighiouart, M.: A Bayesian adaptive design in cancer phase I trials using dose combinations in the presence of a baseline covariate. J. Probab. Stat. **2018**, 11 (2018). https://doi.org/10.1155/2018/8654173
11. Diniz, M.A., Kim, S., Tighiouart, M.: A Bayesian adaptive design in cancer phase I trials using dose combinations with ordinal toxicity grades. Stats **3**, 221–238 (2020)
12. Diniz, M.A., Li, Q., Tighiouart, M.: Dose finding for drug combination in early cancer phase I trials using conditional continual reassessment method. J. Biom. Biostat. **8**, 6 (2017)
13. Diniz, M.A., Tighiouart, M., Rogatko, A.: Comparison between continuous and discrete doses for model based designs in cancer dose finding. PloS One **14**, 1 (2019)
14. Faries, D.: Practical modifications of the continual reassessment method for phase I cancer clinical trials. J. Biopharm. Stat. **4**(147), 164 (1994)
15. Gasparini, M., Eisele, J.: A curve-free method for phase I clinical trials. Biometrics **56**(2), 609–615 (2000)
16. Gatsonis, C., Greenhouse, J.B.: Bayesian methods for phase I clinical trials. Stat. Med. **11**(10), 1377–1389 (1992)
17. Goodman, S., Zahurak, M., Piantadosi, S.: Some practical improvements in the continual reassessment method for phase I studies. Stat. Med. **14**, 1149–1161 (1995)
18. Hu, B., Bekele, B.N., Ji, Y.: Adaptive dose insertion in early phase clinical trials. Clinical Trials **10**, 216–224 (2013)
19. Huang, X., Biswas, S., Oki, Y., Issa, J.-P., Berry, D.A.: A parallel phase I/II clinical trial design for combination therapies. Biometrics **63**(2), 429–436 (2007)
20. Ivanova, A.: A new dose-finding design for bivariate outcomes. Biometrics **59**(4), 1001–1007 (2003)
21. Jiménez, J.L., Kim, S., Tighiouart, M.: A Bayesian seamless phase I–II trial design with two stages for cancer clinical trials with drug combinations. Biom. J. **62**(5), 1300–1314 (2020)
22. Jimenez, J.L., Tighiouart, M., Gasparini, M.: Cancer phase I trial design using drug combinations when a fraction of dose limiting toxicities is attributable to one or more agents. Biom. J. **61**(2), 319–332 (2019)
23. Lam, C.K., Lin, R., Yin, G.: Non-parametric overdose control for dose finding in drug combination trials. J. R. Stat. Soc.: Ser. C: Appl. Stat. **68**(4), 1111–1130 (2019)
24. Le Tourneau, C., Lee, J.J., Siu, L.L.: Dose escalation methods in phase I cancer clinical trials. JNCI J. Natl. Cancer Inst. **101**(10), 708–720 (2009)
25. Li, Q., Tighiouart, M.: EWOC2: Escalation with Overdose Control using 2 Drug Combinations (2019). R package version 1.0
26. Lyu, J., Ji, Y., Zhao, N., Catenacci, D.: AAA: Triple-adaptive Bayesian designs for the identification of optimal dose combinations in dual-agent dose-finding trials. J. R. Stat. Soc.: Ser. C: Appl. Stat. **68**(2), 385–410 (2019)
27. Mander, A.P., Sweeting, M.J.: A product of independent beta probabilities dose escalation design for dual-agent phase I trials. Stat. Med. **34**(8), 1261–1276 (2015)
28. Miles, D., von Minckwitz, G., Seidman, A.D., et al.: Combination versus sequential single-agent therapy in metastatic breast cancer. ONCOLOGIST-MIAMISBURG **7**, 13–19 (2002)
29. Moller, S.: 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. Stat. Med. **14**, 911–922 (1995)
30. Mozgunov, P., Gasparini, M., Jaki, T.: A surface-free design for phase I dual-agent combination trials. Stat. Methods Med. Res. **29**(10), 3093–3109 (2020). 0962280220919450
31. Murtaugh, P.A., Fisher, L.D.: Bivariate binary models of efficacy and toxicity in dose-ranging trials. Commun. Stat. Theory Methods **19**(6), 2003–2020 (1990)
32. O'Quigley, J., Pepe, M., Fisher, L.: Continual reassessment method: a practical design for phase 1 clinical trials in cancer. Biometrics, 33–48 (1990)
33. O'Quigley, J., Shen, L.Z.: Continual reassessment method: a likelihood approach. Biometrics **52**, 673–684 (1996)
34. Piantadosi, S., Fisher, J.D., Grossman, S.: Practical implementation of a modified continual reassessment method for dose-finding trials. Cancer Chemother. Pharmacol. **41**, 429–436 (1998)

35. Plummer, M., et al.: JAGS: A program for analysis of Bayesian graphical models using Gibbs sampling. In: *Proceedings of the 3rd International Workshop on Distributed Statistical Computing*, vol. 124, pp. 1–10. Vienna, Austria (2003)
36. Qayed, M., Cash, T., Tighiouart, M., MacDonald, T., Goldsmith, K., Tanos, R., Kean, L., Watkins, B.C., Wetmore, C., Katzenstein, H.M.: A phase I study of sirolimus in combination with metronomic therapy (CHOAnome) in children with recurrent or refractory solid and brain tumors. Pediatr. Blood Cancer **67**(4), e28134 (2019)
37. R Core Team. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria (2018)
38. Riviere, M., Yuan, Y., Dubois, F., Zohar, S.: A Bayesian dose-finding design for drug combination clinical trials based on the logistic model. Pharm. Stat. **13**, 247–257 (2014)
39. Riviere, M., Yuan, Y., Dubois, F., Zohar, S.: A Bayesian dose-finding design for clinical trials combining a cytotoxic agent with a molecularly targeted agent. J. R. Stat. Soc. Ser. C **64**, 215–229 (2015)
40. Rogatko, A., Ghosh, P., Vidakovic, B., Tighiouart, M.: Patient-specific dose adjustment in the cancer clinical trial setting. Pharm. Medicine **22**(6), 345–350 (2008)
41. Sato, H., Hirakawa, A., Hamada, C.: An adaptive dose-finding method using a change-point model for molecularly targeted agents in phase I trials. Stat. Med. **35**(23), 4093–4109 (2016)
42. Shi, Y., Yin, G.: Escalation with overdose control for phase I drug-combination trials. Stat. Med. **32**(25), 4400–4412 (2013)
43. Shimamura, F., Hamada, C., Matsui, S., Hirakawa, A.: Two-stage approach based on zone and dose findings for two-agent combination phase I/II trials. J. Biopharm. Stat., **28**(6), 1025–1037 (2018)
44. Thall, P.F., Cook, J.D.: Dose-finding based on efficacy-toxicity trade-offs. Biometrics **60**(3), 684–693 (2004)
45. Thall, P.F., Millikan, R.E., Mueller, P., Lee, S.-J.: Dose-finding with two agents in phase I oncology trials. Biometrics **59**(3), 487–496 (2003)
46. Thall, P.F., Russell, K.E.: A strategy for dose-finding and safety monitoring based on efficacy and adverse outcomes in phase I/II clinical trials. Biometrics, 251–264 (1998)
47. Tighiouart, M.: Two-stage design for phase I–II cancer clinical trials using continuous dose combinations of cytotoxic agents. J. R. Stat. Soc.: Ser. C: Appl. Stat. **68**(1), 235–250 (2019)
48. Tighiouart, M., Cook-Wiens, G., Rogatko, A.: Incorporating a patient dichotomous characteristic in cancer phase I clinical trials using escalation with overdose control. Journal of Probab. Stat. **2012**, 10 (2012)
49. Tighiouart, M., Cook-Wiens, G., Rogatko, A.: A Bayesian adaptive design for cancer phase I trials using a flexible range of doses. J. Biopharm. Stat. **28**(3), 562–574 (2018)
50. Tighiouart, M., Li, Q., Piantadosi, S., Rogatko, A.: A Bayesian adaptive design for combination of three drugs in cancer phase I clinical trials. Am. J. Biostat. **6**(1), 1 (2016)
51. Tighiouart, M., Li, Q., Rogatko, A.: A Bayesian adaptive design for estimating the maximum tolerated dose curve using drug combinations in cancer phase I clinical trials. Stat. Med. **36**(2), 280–290 (2017)
52. Tighiouart, M., Liu, Y., Rogatko, A.: Escalation with overdose control using time to toxicity for cancer phase I clinical trials. Plos One **9**, 3 (2014)
53. Tighiouart, M., Piantadosi, S., Rogatko, A.: Dose finding with drug combinations in cancer phase I clinical trials using conditional escalation with overdose control. Stat. Med. **33**(22), 3815–3829 (2014)
54. Tighiouart, M., Rogatko, A.: Dose finding with escalation with overdose control (EWOC) in cancer clinical trials. Stat. Sci. **25**(2), 217–226 (2010)
55. Tighiouart, M., Rogatko, A.: Number of patients per cohort and sample size considerations using dose escalation with overdose control. J. Probab. Stat. **2012**, 16 (2012). https://doi.org/10.1155/2012/692725
56. Tighiouart, M., Rogatko, A., Babb, J.S.: Flexible Bayesian methods for cancer phase I clinical trials. dose escalation with overdose control. Stat. Med. **24**(14), 2183–2196 (2005)

57. Tuli, R., Shiao, S.L., Nissen, N., Tighiouart, M., Kim, S., Osipov, A., Bryant, M., Ristow, L., Placencio-Hickok, V.R., Hoffman, D., Rokhsar, S., Scher, K., Klempner, S.J., Noe, P., Davis, M.J., A, W., Lo, S., Jamil, L., Sandler, H., Piantadosi, S., Hendifar, A.: A phase 1 study of veliparib, a PARP-1/2 inhibitor, with gemcitabine and radiotherapy in locally advanced pancreatic cancer. EBioMedicine **40**, 374–381 (2019)
58. US Department of Health and Human Services. National Cancer Institute: Common Terminology Criteria for Adverse Events (CTCAE) Version 4.0. [internet] National Cancer Institute, Bethesda (2009). [cited 2015 sep 22]
59. Wages, N.A.: Identifying a maximum tolerated contour in two-dimensional dose finding. Stat. Med. **36**(2), 242–253 (2017)
60. Wages, N.A., Conaway, M.R.: Phase I/II adaptive design for drug combination oncology trials. Stat. Med. **33**(12), 1990–2003 (2014)
61. Wages, N.A., Conaway, M.R., O'Quigley, J.: Continual reassessment method for partial ordering. Biometrics **67**(4), 1555–1563 (2011)
62. Wang, K., Ivanova, A.: Two-dimensional dose finding in discrete dose space. Biometrics **61**(1), 217–222 (2005)
63. Wheeler, G.M., Sweeting, M.J., Mander, A.P.: Toxicity-dependent feasibility bounds for the escalation with overdose control approach in phase I cancer trials. Stat. Med. **36**(16), 2499–2513 (2017)
64. Whitehead, J., Thygesen, H., Whitehead, A.: A Bayesian dose-finding procedure for phase I clinical trials based only on the assumption of monotonicity. Stat. Med. **29**(17), 1808–1824 (2010)
65. Yin, G., Yuan, Y.: A latent contingency table approach to dose finding for combinations of two agents. Biometrics **65**(3), 866–875 (2009)
66. Yu, J., Hutson, A.D., Siddiqui, A.H., Kedron, M.A.: Group sequential control of overalltoxicity incidents in clinical trials–non-bayesian and bayesian approaches. Stat. Methods Med. Res. **25**(1), 64–80 (2016)
67. Yuan, Y., Yin, G.: Bayesian dose finding by jointly modelling toxicity and efficacy as time-to-event outcomes. J. R. Stat. Soc.: Ser. C: Appl. Stat. **58**(5), 719–736 (2009)
68. Yuan, Y., Yin, G.: Bayesian phase I/II adaptively randomized oncology trials with combined drugs. Ann. Appl. Stat. **5**(2A), 924 (2011)
69. Zacks, S., Rogatko, A., Babb, J.: Optimal Bayesian-feasible dose escalation for cancer phase I trials. Statist. Probab. Lett. **38**(3), 215–220 (1998)
70. Zhang, L., Yuan, Y.: A practical Bayesian design to identify the maximum tolerated dose contour for drug combination trials. Stat. Med. **35**(27), 4924–4936 (2016)
71. Zhengjia, C., Mourad, T., Jeanne, K.: Dose escalation with overdose control using a quasi-continuous toxicity score in cancer phase I clinical trials. Contemp. Clin. Trials **33**(5), 949–958 (2012)