mathematics

MDPI

Article

Comparison Among Modified Continual Reassessment Methods with Different Dose Allocation Methods for Phase I Clinical Trials

Jiacheng Xiao †, Weijia Zhang †, Rong Li * and Conghua Wen *

Department of Financial and Actuarial Mathematics, Xi'an Jiaotong-Liverpool University, Suzhou 215123, China; jiacheng.xiao17@student.xjtlu.edu.cn (J.X.); weijia.zhang@xjtlu.edu.cn (W.Z.)
* Correspondence: rong.li@xjtlu.edu.cn (R.L.); conghua.wen@xjtlu.edu.cn (C.W.)
† These authors contributed equally to this work.

Abstract: The continual reassessment method (CRM) has been an essential Bayesian finding design in phase I clinical trials. It utilizes all the information in observed data which contributes to its essential operational characteristics. However, the CRM has been criticized for its aggressive dose escalation. Model-assisted methods including BOIN, Keyboard, and mTPI improved the safety while retaining relative efficiency. In this paper, we propose four models combining the structure of the CRM and model-assisted methods. We show that these models could operate with comparable CRM performance through simulations. The results suggest that two of the proposed methods outperformed the traditional methods with a higher percentage of correct selection of true maximum tolerated dose. In addition, the interval-based approaches offered by the new models with greater flexibility regarding target toxicity achieved an improvement in the adaptability of the dose-finding process in clinical trials.

Keywords: clinical trial; biostatistics; CRM; Bayesian adaptive design

MSC: 82-10

Received: 2 February 2025
Revised: 27 February 2025
Accepted: 4 March 2025
Published: 5 March 2025

Citation: Xiao, J.; Zhang, W.; Li, R.; Wen, C. Comparison Among Modified Continual Reassessment Methods with Different Dose Allocation Methods for Phase I Clinical Trials. *Mathematics* **2025**, *13*, 863. https://doi.org/10.3390/math13050863

1. Introduction

Clinical trial designs assist clinicians in conducting clinical experiments. They are proposed to evaluate the toxicity and efficiency of new drugs on the human body in the short and long term. Clinical trials are categorized into four phases based on different research objects and administrative requirements. Phase I clinical trial designs aim at detecting the maximum tolerated dose which is generally abbreviated to MTD. It is defined as the dose level that possesses the toxicity probability nearest to the acceptable target toxicity rate. Phase II trials focus on the therapeutic effects of the novel agent. The trade-off between toxicity and efficacy has incurred a trend in conducting phase I/II clinical trials to achieve the highest efficacy with a limited toxic response. Phase III trials are devoted to comparing the agent with existing drugs to examine the long-term effectiveness. The side effects are monitored through phase IV trials as well as further investigations after marketing.

Phase I clinical trial designs could be broadly classified into three categories including model-based, model-assisted, and algorithm-based methods. For the algorithm-based methods, toxicity probabilities monotonically increase in dose levels without the presumption of statistical dose–toxicity models. For this reason, they are called non-parametric

models as well. Dose allocation rules under different phenomenons could be specified in advance without subsequent calculations. This is one of the major incentives for clinicians to favor algorithm-based designs. The most widely acknowledged 3+3 method has played an essential role in this field owing to its simplicity and efficiency. It was the most frequently used method in practice. Storer [1] summarized the procedure of this method with mathematical modeling. A+B and A+B+C methods [2], which are variations of the 3+3 method, were considered as well and the statistical inference of this category of methods was discussed. However, the actual implementation of algorithm-based methods did not perform as expected due to their conservativeness. They tended to select the dose levels below the true MTD. The time-consuming procedure was another critical defect that has been argued frequently.

In order to defeat these deficiencies, a number of modifications, for example, Rolling 6 Design [3] and Accelerated Titration Design [4], were introduced. Practical applications indicate the improvements of the modified algorithm-based methods. Taking advantage of these designs, the i3+3 method [5] was conducted by integrating the features of 3+3 and interval-based designs. It has been gradually adopted in modern phase I clinical trials.

From the perspective of statistics, it is feasible to construct models by assuming an explicit dose–toxicity relationship in advance. As a result, a statistical inference could be involved in the dose-finding process. Model-based designs were originated from this concept. They are also known as parametric designs because of the incorporation of presumed parametric models. O'Quigley et al. [6] proposed the continual reassessment method (CRM), which is the representative of the model-based method. It was initially created to incorporate the prior information and the acquired experimental outcomes with the application of Bayesian knowledge. A sophisticated design implementing statistical concepts allows the parametric methods to pool the toxicity response data and the decision rule shall utilize all information acquired in the experiment. Later, practical modifications in the CRM were conducted by Goodman et al. [7] including termination rules, dose escalation control, and initiating from the lowest. O'Quigley and Shen [8] applied a likelihood approach based on the original CRM. Daimon et al. [9] further investigated the Bayesian applications in clinical trial designs focusing on the issue of posterior probabilities. Garrett-Mayer [10] and Wheeler et al. [11] summarized the specific procedure of the CRM. In contrast, the performance could be highly influenced by the assumptions of the model, including the dose–toxicity model, prior distribution, and skeleton, because of a small sample size of trials.

To overcome these flaws, Gasparini and Eisele [12] proposed a curve-free method to introduce a more flexible model. BMA-CRM [13] conquered the issue of skeleton assignment. They averaged the results of the model with certain weights assigned according to the performance of skeletons. Pan and Yuan [14] further discussed the issue of specifying skeletons in BMA-CRM. Zhang et al. [15] examined the performance of the CRM with a novel dose–toxicity curve. DA-CRM [16] and EM-CRM [17] were designed for specific situations encountered in clinical trials. In addition, Abbas et al. [18] discussed the situation where the prior assumptions were violated. In addition, the time delay factor was introduced by Cheung and Chappell [19].

The latest designs tend to take advantage of both algorithm-based and model-based methods. As a result, model-assisted methods were proposed and they abandoned the usage of additional parameters required in the CRM but merely considered the relationship between toxicity probabilities and dose levels. The posterior probabilities were incorporated differently. These works deduced several different dose allocation rules and final recommendation schemes based on their characteristics.

Considering the complexities of model-based methods, Ji et al. [20] first proposed the TPI method to design a model that would be easily understood and applied by the non-statisticians. Modifications of these methods, mTPI [21] and mTPI-2 [22], were proposed subsequently. Later, Keyboard [23] was introduced to simplify the interval selection problems in mTPI-2. The idea of optimization was introduced as well and Bayesian optimal interval design (BOIN) [24] was designed by reducing the probabilities of wrong selection of dose level. Keyboard and BOIN could be operated through the application of EXCEL which is greatly convenient to apply. Yuan et al. [25] considered delayed toxicities in the application of BOIN. In 2022, Zhou et al. [26] further extended the TITE-BOIN in phase I/II.

In practice, the most frequently used methods would still be 3+3 and BOIN owing to the convenience of calculation and interpretation as well as the compliance with FDA requirements. Nevertheless, the model-based methods demonstrate remarkable performance. In this paper, our goal is to construct four trial designs integrating the benefits of model-based and model-assisted designs. The structure of the CRM would be adopted to capture all the information during the trials. This would remedy the model-assisted methods which make decisions only based on the information of the current dose. To enhance the safety of our design, the dose allocation would reference the methods that have been adopted in mTPI, Keyboard, and i3+3 as alternatives compared to the CRM. In addition, the usage of interval-based designs would be able to allow a more flexible DLT range instead of a single point. The objective of the proposed method would be to fully use the observed information while retaining the simplicity of the model-assisted designs and introduce the acceptable interval to enhance the flexibility of the designs. This would result in higher or comparable accuracy of finding the true MTD and considerable overdose control could be confirmed.

The rest of the article is organized as follows. In Section 2, the basic procedure and structure of the model are deduced, respectively. A simulation could be conducted in Section 3 and we shall discuss the results with certain indexes in Section 4. The conclusion will be demonstrated in the form of a discussion in Section 5.

2. Methods

2.1. Traditional CRM

The CRM introduced by O'Quigley et al. [6] has been widely adopted in phase I clinical trials as a result of its accuracy and efficiency. Compared with algorithm-based methods, the CRM combines prior knowledge with statistical models that adopt Bayesian statistics as the core methodology. A set of toxicity probabilities of p_i, $i = 1, \cdots k$ with progressively increasing relationship $p_1 < \cdots < p_k$ is pre-specified where k levels of dose are given. With prior knowledge acquired from experience, these probabilities are predetermined by statisticians and clinicians and are generally known as the skeleton of the CRM.

The dose–toxicity model is proposed with one or more unknown parameters that need to be estimated. This method aims to model the real toxicity curve based on consecutive observations to find the MTD, which is the dose level with the toxicity probability nearest to the target toxicity probability ϕ_T. Let target toxicity probability $\phi_T = 0.3$.

Denote a as the parameter and the prior distribution of $f(a)$ is specified in advance. It also represents the prior beliefs of clinicians before the trials begin. In addition, the dose–toxicity models are proposed. Power, logistics, and hyperbolic tangent models are frequently used in the literature. Yin [27] summarized the common models used in designs with predefined parameters in Table 1. Shen and O'Quigley [28] discussed that necessity of the one-parameter model. It was indicated that one parameter would be sufficient to model

the dose–toxicity curve and introduction of another parameter may lead to instability of the estimates. The one-parameter model was also applied in [13,14,29]. The toxicity probabilities would be defined as

$$\mathbb{P}(toxicity\ probability\ at\ dose\ level\ i) = \pi_i(a) = f(a, p_i),\ i = 1, \cdots, k.$$

The experiment will generally begin with the lowest dose level concerning safety once the model and assumptions are set. During the trials, assuming that y_i toxicities have been observed from n_i patients treated at dose level i, the likelihood function is formulated as

$$L(\mathcal{D}|\mathrm{a}) \propto \prod_{\mathrm{i}=1}^{\mathrm{k}} \{\mathrm{f}(\mathrm{a}, \mathrm{p_i})\}^{\mathrm{y_i}} \{1 - \mathrm{f}(\mathrm{a}, \mathrm{p_i})\}^{(\mathrm{n_i} - \mathrm{y_i})}.$$

According to the Bayesian theorem [30], the posterior distribution would be derived and the posterior mean toxicity probability $\hat{\pi}_i$ at dose level i would be calculated as

$$f(a|\mathcal{D}) = \frac{L(\mathcal{D}|\mathrm{a})\mathrm{f}(\mathrm{a})}{\int L(\mathcal{D}|a) f(a) da},$$

$$\hat{\pi}_i = \int f(a, p_i) f(a|\mathcal{D}) da = \int f(a, p_i) \frac{L(\mathcal{D}|\mathrm{a})\mathrm{f}(\mathrm{a})}{\int L(\mathcal{D}|a) f(a) da} da.$$

With the estimation obtained through the observations, the updated recommendation of dose level with posterior mean nearest to the MTD would be assigned to the next cohort of patients. The trial would be terminated till the maximum sample size is reached or the early termination is adopted.

Table 1. Parameter space.

Model Type	Toxicity Probabilities	Parameters
Power Model	$p_j^{exp(a)}$	$ln(\frac{ln(p_j)}{ln(\phi_T)})$
Logistics Model	$\frac{exp(-3+ap_j)}{1+exp(-3+ap_j)}$	$\frac{1}{p_j} ln(\frac{\phi_T}{1-\phi_T})$
Hyperbolic Tangent Model	$(\frac{tanh(p_j)+1}{2})^a$	$\frac{ln(\phi_T)}{ln(tanh(p_j)+1)}$

2.2. Proposed Methods

Based on the fundamental structure of the CRM, we propose to apply the dose allocation rule in model-assisted methods. The power model will be selected as in simulations. Specifically, dose selection rules in mTPI, i3+3, and Keyboard will be integrated into the model-based Bayesian framework. Posterior probabilities $\pi_i(a|\mathcal{D})$ will be calculated with the observed data $\mathcal{D}$. The information on dose response would be fully utilized in successively updating posterior distribution. The trial will sequentially recommend dose levels based on the procedures in the following subsections. The Bayesian formula could allow us to fully use the information observed throughout trials. Therefore, once the trial is terminated, the final recommended dose level would be the one chosen based on all the data from patients across the trial.

Taking the power model as an illustration for our idea, the toxicity model would be

$$\mathbb{P}(toxicity\ probability\ at\ dose\ level\ i) = \pi_i(a) = p_i^{exp(a)},$$

where a is the unknown parameter which would be re-estimated through the trials. In addition, the prior distribution $f(a)$ would be determined in advance, summarizing the

pre-knowledge of the agent. Assume that we could observe y_i toxicities occurring in n_i patients who are treated at dose level i; then, the likelihood function can be formulated as

$$L(\mathcal{D}|a) \propto \prod_{i=1}^{k} \{p_i^{exp(a)}\}^{y_i} \{1 - p_i^{exp(a)}\}^{(n_i - y_i)}.$$

The posterior distribution $f(a|\mathcal{D})$ would be derived with the observations of (n_i, y_i) by the formula

$$\begin{aligned} f(a|\mathcal{D}) &= \frac{L(\mathcal{D}|a)f(a)}{\int L(\mathcal{D}|a)f(a)da} \\ &= \frac{\prod_{i=1}^{k} \{p_i^{exp(a)}\}^{y_i} \{1 - p_i^{exp(a)}\}^{(n_i - y_i)} f(a)}{\int \prod_{i=1}^{k} \{p_i^{exp(a)}\}^{y_i} \{1 - p_i^{exp(a)}\}^{(n_i - y_i)} f(a) f(a) da}. \end{aligned}$$

The posterior mean would be calculated as the toxicity probabilities estimated at dose level i. Thus, the estimated toxicities $\hat{\pi}_i$ would be

$$\hat{\pi}_i = \int p_i^{exp(a)} f(a|\mathcal{D}) da.$$

Termination rules for safety control are adopted as those used in Ji et al. [21]. In addition, a rule of escalation control [7] is introduced for the reason that the CRM has been criticized for its aggressive escalation. We control the de-escalation procedure since ineffective treatment at lower levels could be harmful as well. Escalation with overdose control (EWOC) is employed to solve the overdose issue [31].

- Early Termination
  Suppose that dose level 1 has been recommended for patients. The trial would be terminated early if the observed data indicate the unacceptable toxicity probability at the lowest dose level. That is

  $$\mathbb{P}(\pi_1(a) > \phi_T \,|\mathcal{D}) > \xi,$$

  where the ξ is of a value that is close to 1. Be clear that, once the early termination is determined, no dose would be recommended because of the intolerable toxicities.
- EWOC
  For dose level i, the next dose level at $i + 1$ will not be recommended if

  $$\mathbb{P}(\pi_{i+1}(a) > \phi_T \,|\mathcal{D}) > \delta,$$

  where the δ should be large enough, usually say 0.65, to reduce the chance of overdose. The following cohorts of patients will be suggested to stay at the same level until there is enough evidence to indicate that the chance of over-toxicity in the next dose level is small enough for dose escalation.
- Escalation Control
  In the trial, the transition will be restricted within one dose range. In other words, when escalation/de-escalation is chosen for patient $j + 1$, we could select at most one dose level higher/lower than the level at which patient j has been treated.

2.2.1. iCRM

Instead of only considering the posterior mean toxicity probability, we focus on the toxicity interval with the updated posterior distribution of a. Starting at the lowest dose

level, we could acquire the toxicity response sequentially. Define the probability that the posterior probability at each dose level falls within the tolerable interval (0.25, 0.35) as

$$T_i = \mathbb{P}\{\pi_i(a) \in [0.25, 0.35] \mid \mathcal{D}\},\ for\ i = 1, \cdots, k.$$

The dose level would be selected according to the schemes that

$$J = \underset{i \in \{1, \cdots, k\}}{\arg\max}\{T_i\}.$$

That is, the dose level with the highest posterior probability of allocating in the target interval would be selected as the next recommendation.

Based on the power model, at a given dose level i, we have

$$\begin{aligned} T_i &= \mathbb{P}\{0.25 \leq \pi_i(a) \leq 0.35 \mid \mathcal{D}\} \\ &= \mathbb{P}\{0.25 \leq p_i^{exp(a)} \leq 0.35 \mid \mathcal{D}\} \\ &= \mathbb{P}\{ln(\frac{ln(0.35)}{ln(p_i)}) \leq a \leq ln(\frac{ln(0.25)}{ln(p_i)}) \mid \mathcal{D}\} \\ &= \int_{ln(\frac{ln(0.35)}{ln(p_i)})}^{ln(\frac{ln(0.25)}{ln(p_i)})} f(a|\mathcal{D})\ \mathrm{da}. \end{aligned}$$

The dose level with the largest T_i would be selected as a recommendation for the next cohort of patients. The final recommendation would follow the same idea with all patients' data collected.

2.2.2. mCRM

Motivated by mTPI [21], defining EI as $[\phi_T - \epsilon_1, \phi_T + \epsilon_2]$, we could be able to derive the dose allocation rule. Suppose we have collected the dose toxicity responses and the observed data $\mathcal{D} = \{(n_i, y_i),\ for\ i = 1, \cdots, k\}$. The probability of the posterior toxicity probabilities that belong to the interval $[L, U]$ could be defined as

$$\hat{p_i}(L, U) = \mathbb{P}(\pi_i(a) \in [L, U] \mid \mathcal{D}).$$

With the obtained responses of patients, the dose should be escalated, stayed, or de-escalated if $(0, \phi_T - \epsilon_1)$, $[\phi_T - \epsilon_1, \phi_T + \epsilon_2]$, $(\phi_T + \epsilon_2, 1]$ has the largest posterior probability mass. That is

$$j = \underset{j \in \{E, D, S\}}{\arg\max}\{U_j\},$$

where, for each interval, we have

$$U_E = \hat{p_i}(0, \phi_T - \epsilon_1) = \int_{\frac{ln(\phi_T - \epsilon_1)}{ln(p_i)}}^{\infty} f(a|\mathcal{D})\ \mathrm{da};$$

$$U_S = \hat{p_i}(\phi_T - \epsilon_1, \phi_T + \epsilon_2) = \int_{\frac{ln(\phi_T + \epsilon_2)}{ln(p_i)}}^{\frac{ln(\phi_T - \epsilon_1)}{ln(p_i)}} f(a|\mathcal{D})\ \mathrm{da};$$

$$U_D = \hat{p_i}(\phi_T + \epsilon_2, 1) = \int_{0}^{\frac{ln(\phi_T + \epsilon_2)}{ln(p_i)}} f(a|\mathcal{D})\ \mathrm{da}.$$

Generally, we choose $\epsilon_1 = \epsilon_2 = 0.05$ and the trial would be conducted until early termination or the sample size has been exhausted. The final recommendation would be derived from all the information collected throughout the trial.

2.2.3. bCRM

In this model, we propose to calculate the mean posterior toxicity probability at dose level i and choose the dose for the next patient based on the following rule

$$\begin{cases} \text{Escalation} & if\ \hat{\pi}_i \in (0, 0.25) \\ \text{Stay} & if\ \hat{\pi}_i \in [0.25, 0.35] \\ \text{De-escalation} & if\ \hat{\pi}_i \in (0.35, 1] \end{cases}$$

These intervals are similar to the one applied in i3+3 [5]. We named it the boundary CRM (bCRM). If the posterior mean toxicity probability of the current dose level i allocates in (0.35, 1], the dose level can be considered too toxic and de-escalation would be adopted. On the opposite, if $\hat{\pi}_i \in (0, 0.25)$, we would recommend one dose level higher because of the inefficient treatment. We would assign the same dose level i once $\hat{\pi}_i \in [0.25, 0.35]$ is calculated. Compared to the traditional CRM, this would allow for a tolerance interval for toxicity probabilities.

2.2.4. kCRM

Different from the mCRM, ten intervals would be divided and referred to as keys [11]. Firstly, we select our target key as (0.25, 0.35) which is the same interval we are aiming at in the mCRM and iCRM. With a width of 0.1, other intervals would be derived as (0.05, 0.15), (0.15, 0.25), (0.35, 0.45), (0.45, 0.55), (0.55, 0.65), (0.65, 0.75), (0.75, 0.85), and (0.85, 0.95). The recommended selection of boundaries of j-th key in Yan et al. [23] could also be written as

$$V_j = 0.1j - 0.05,\ for\ j = 1, \cdots 9.$$

Posterior probability in each interval for dose level i could be calculated and we define

$$M_j = \hat{p}_i(V_j, V_{j+1}) = \mathbb{P}\{\pi_i(a\ |\mathcal{D}) \in [V_j, V_{j+1}]\} = \int_{\frac{ln(V_{j+1})}{ln(p_i)}}^{\frac{ln(V_j)}{ln(p_i)}} f(a|\mathcal{D})\ \text{da}, \text{for j} = 1, \cdots, 9.$$

Then, for the current dose level i, the strongest key denoted as K would be the one with the highest posterior probability mass which is

$$K = \underset{j \in \{1, \cdots, 9\}}{\arg\max}\{M_j\}.$$

Consider the situation where the strongest key is located after a probability of 0.35, that is $K > 3$. It indicates that the posterior toxicity probability at dose level i would have a higher chance to locate above the upper bound of the tolerable interval. As a result, dose de-escalation would be adopted and the next lower dose would be recommended. If $K < 3$ is obtained, the current dose could be regarded as underestimated and the next higher dose will be suggested to the next patient. Only when the strongest key is allocated in the target key will we choose to stay at the current dose. The selection schemes are summarized as

$$\begin{cases} \text{Escalation} & if\ K < 3 \\ \text{Stay} & if\ K = 3 \\ \text{De-escalation} & if\ K > 3 \end{cases}$$

Once the patients in the trial are exhausted, the dose selection deduced with all observed data would be our final recommendation.

2.3. Numerical Integrations

The novel designs rely on a Bayesian statistical model to estimate the dose–toxicity relationship, and the application often involves solving integrals or performing simulations to update the model parameters based on observed data. Closed-form solutions are rarely available because the likelihood function can be complicated owing to the sequential enrollment of patients at different dose level. In addition, the prior and posterior distribution may not be conjugate, making it challenging to obtain the closed-form result. Due to the application of the one-parameter model, numerical integration methods could be adopted to calculate the posterior distribution. The integrate function in R has been used in our simulations. Techniques like Monte Carlo simulations could be used in models with more parameters.

3. Simulation

The operational characteristics would be investigated through simulation conducted through R (latest v. R4.4.3). For the experimental settings, we propose to examine six dose levels with the existence of a basic presumption of monotonicity in the dose–toxicity relationship. There would be at most 36 patients enrolled in each trial and three patients would be introduced for each dose assignment in a cohort. Ten scenarios are listed and demonstrated in Figure 1. The scenarios adopted vary in the location of the true MTD and the shape of the corresponding toxicity curve. The Yin et al. [29] has adopted similar scenarios for simulations. The extreme situations mentioned in [11] where no MTD would be found since the lowest dose level possesses the toxicity probability above target or the highest dose level possesses the toxicity probability below target are considered as well.

$$\begin{cases} (0.1, 0.12, 0.30, 0.50, 0.60, 0.65), \text{scenario1} \\ (0.06, 0.08, 0.10, 0.15, 0.30, 0.45), \text{scenario2} \\ (0.08, 0.12, 0.20, 0.30, 0.40, 0.50), \text{scenario3} \\ (0.02, 0.05, 0.08, 0.10, 0.14, 0.30), \text{scenario4} \\ (0.05, 0.14, 0.18, 0.20, 0.23, 0.30), \text{scenario5} \\ (0.05, 0.10, 0.30, 0.50, 0.65, 0.75), \text{scenario6} \\ (0.01, 0.02, 0.04, 0.06, 0.30, 0.50), \text{scenario7} \\ (0.20, 0.30, 0.45, 0.55, 0.60, 0.70), \text{scenario8} \\ (0.30, 0.45, 0.50, 0.60, 0.70, 0.80), \text{scenario9} \\ (0.50, 0.60, 0.70, 0.80, 0.85, 0.90), \text{scenario10} \end{cases}$$

A general pre-specified probability skeleton is applied, that is

$$(p_1, \cdots, p_6) = (0.1, 0.2, 0.3, 0.4, 0.5, 0.6).$$

The parameters $\epsilon_1 = \epsilon_2 = 0.05$ are determined in the mCRM. For safety control, $\xi = 0.98$ and $\delta = 0.65$ are proposed to reduce the chance of assigning doses with a high risk of toxicities. In total, 10,000 simulations were carried out for a target of $\phi_T = 0.3$. Normal prior distribution of a is assumed to be N(0, 2). For comparison, we compare the results of our model with the traditional CRM [6], BOIN [24], and Keyboard [23]. Basic settings for proposed methods are applied in the CRM as well.

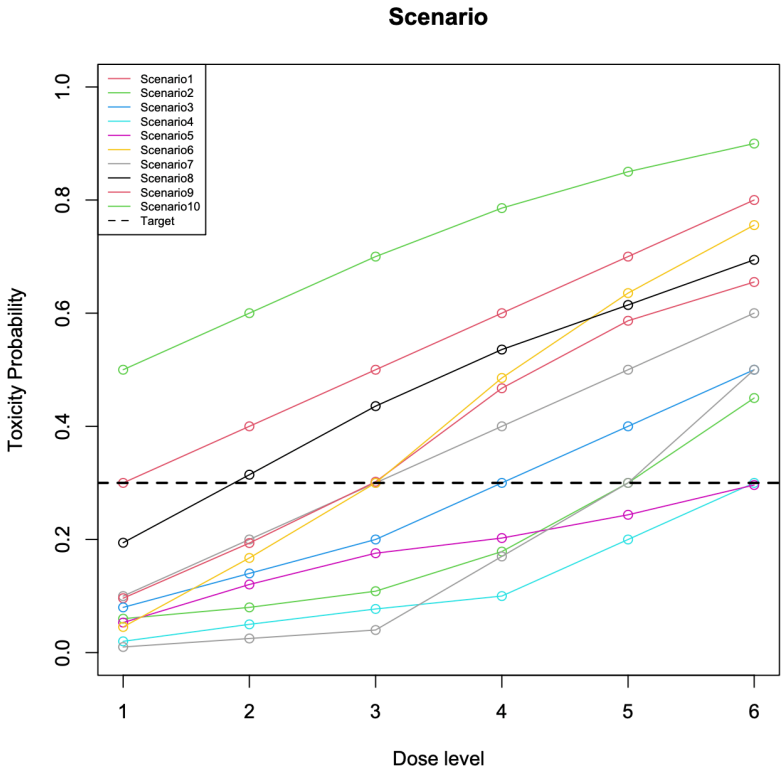

Figure 1. True toxicity probabilities under different scenarios.

4. Results and Discussion

4.1. Discussion of the Simulation Results

In Tables 2–5, we have the simulation results conducted with different methods. The first row under each scenarios is the true toxicity probabilities of each dose level. The following row compares the results of the traditional CRM, iCRM, kCRM, bCRM, and mCRM under Normal(0, 2) assumptions while the fourth and the fifth rows depict the performance of Keyboard and BOIN.

The percentage of correct selection, i.e., the percentage of dose recommendations at the true MTD, would be used as a primary measure to evaluate the performance of the operating models [9]. The inconclusive trials which are caused by early termination of an over-toxic initial dose would be denoted by "NONE". Three additional measures would be adopted as supplements for the evaluation.

- Allocation percentage at each dose is calculated as percentage of patients treated at each dose within each trial.
- Average number of DLTs that occurred in trials across the simulations measures the safety of the design.
- The rate of overdose is the percentage of trials that recommended the dose level above the true MTD.

As suggested in [32], full optimization of the average number of DLTs and rate of overdose may not be desirable. Allocating more patients at lower doses would achieve better performance of these criteria but with an overly conservative escalation schedule.

The results are summarized in Figures 2–4. In general, most of our proposed methods performed well compared to the traditional CRM, Keyboard, and BOIN. To be precise, the mCRM achieved the highest percentage of selecting the correct MTD in scenarios 1, 3, 4, 5, and 6 and the iCRM achieved the highest percentage of selecting the correct MTD in scenarios 2 and 8. The kCRM outperformed other methods in scenario 9. In scenario 10, no dose shall be selected as the MTD since the probability of toxicity at the lowest is

unacceptable already. BOIN achieved the highest rate of early termination which shall be considered as safe.

Table 2. Simulation results of final recommendation percentage at each dose level and average number of toxicities [scenarios 1–5].

	Final Recommendation Percentage at Dose Level							**Ave # of Toxicities**	**Overdose**
Designs	**1**	**2**	**3**	**4**	**5**	**6**	**None**		
Scenario1	0.1	0.12	0.3	0.5	0.6	0.65			
kCRM	0.009	0.226	0.653	0.107	0.003	0.000	0.001	9.981	0.110
iCRM	0.002	0.126	0.678	0.187	0.005	0.000	0.001	9.957	0.192
bCRM	0.021	0.180	0.602	0.176	0.020	0.001	0.001	9.840	0.196
mCRM	0.041	0.021	0.694	0.238	0.004	0.000	0.001	11.311	0.243
CRM	0.003	0.147	0.683	0.161	0.004	0.000	0.001	9.831	0.165
Keyboard	0.006	0.189	0.669	0.123	0.010	0.001	0.003	9.340	0.133
BOIN	0.007	0.196	0.651	0.134	0.009	0.001	0.003	8.341	0.144
Scenario2	0.06	0.08	0.1	0.15	0.3	0.45			
kCRM	0.001	0.007	0.033	0.267	0.551	0.141	0.000	7.838	0.141
iCRM	0.000	0.004	0.020	0.182	0.584	0.209	0.000	7.830	0.209
bCRM	0.000	0.004	0.044	0.216	0.503	0.232	0.000	7.720	0.232
mCRM	0.006	0.005	0.012	0.166	0.544	0.268	0.000	7.597	0.268
CRM	0.000	0.004	0.024	0.200	0.573	0.198	0.000	7.722	0.198
Keyboard	0.001	0.005	0.022	0.231	0.543	0.199	0.000	7.413	0.199
BOIN	0.001	0.007	0.035	0.279	0.490	0.189	0.000	6.265	0.189
Scenario3	0.08	0.12	0.2	0.3	0.4	0.5			
kCRM	0.003	0.049	0.295	0.451	0.180	0.022	0.001	8.830	0.202
iCRM	0.001	0.027	0.239	0.471	0.228	0.033	0.001	8.809	0.261
bCRM	0.003	0.049	0.266	0.426	0.212	0.043	0.001	8.710	0.255
mCRM	0.022	0.016	0.151	0.531	0.237	0.043	0.001	8.978	0.280
CRM	0.001	0.034	0.259	0.463	0.212	0.030	0.001	8.717	0.242
Keyboard	0.005	0.053	0.277	0.418	0.199	0.046	0.001	8.364	0.245
BOIN	0.006	0.067	0.313	0.410	0.166	0.038	0.001	7.336	0.203
Scenario4	0.02	0.05	0.08	0.1	0.14	0.3			
kCRM	0.000	0.000	0.004	0.032	0.274	0.690	0.000	6.250	0.000
iCRM	0.000	0.000	0.002	0.018	0.180	0.801	0.000	6.257	0.000
bCRM	0.000	0.000	0.003	0.037	0.176	0.785	0.000	6.201	0.000
mCRM	0.000	0.000	0.001	0.017	0.157	0.825	0.000	6.104	0.000
CRM	0.000	0.000	0.003	0.021	0.190	0.787	0.000	6.209	0.000
Keyboard	0.000	0.001	0.004	0.025	0.237	0.732	0.000	5.883	0.000
BOIN	0.000	0.001	0.009	0.058	0.248	0.684	0.000	4.998	0.000
Scenario5	0.05	0.14	0.18	0.2	0.23	0.3			
kCRM	0.001	0.034	0.116	0.234	0.323	0.293	0.000	7.008	0.000
iCRM	0.000	0.021	0.093	0.212	0.310	0.363	0.000	7.010	0.000
bCRM	0.002	0.024	0.110	0.209	0.287	0.368	0.000	6.9657	0.000
mCRM	0.008	0.009	0.065	0.229	0.300	0.389	0.000	6.933	0.000
CRM	0.000	0.023	0.105	0.216	0.310	0.346	0.000	6.972	0.000
Keyboard	0.009	0.051	0.107	0.181	0.271	0.382	0.000	6.804	0.000
BOIN	0.010	0.068	0.146	0.253	0.235	0.288	0.000	6.120	0.000

Table 3. Simulation results of final recommendation percentage at each dose level and average number of toxicities [scenarios 6–10].

	Final Recommendation Percentage at Dose Level							**Ave # of Toxicities**	**Overdose**
Designs	**1**	**2**	**3**	**4**	**5**	**6**	**None**		
Scenario6	0.05	0.1	0.3	0.5	0.65	0.75			
kCRM	0.001	0.188	0.698	0.112	0.002	0.000	0.000	10.358	0.113
iCRM	0.000	0.090	0.705	0.201	0.003	0.000	0.000	10.331	0.204
bCRM	0.018	0.146	0.627	0.187	0.021	0.000	0.000	10.205	0.208
mCRM	0.004	0.004	0.734	0.254	0.004	0.000	0.000	11.921	0.257
CRM	0.000	0.106	0.715	0.176	0.003	0.000	0.000	10.205	0.178
Keyboard	0.003	0.175	0.695	0.121	0.005	0.000	0.000	9.300	0.127
BOIN	0.003	0.177	0.681	0.132	0.007	0.000	0.000	8.159	0.139
Scenario7	0.01	0.02	0.04	0.06	0.3	0.5			
kCRM	0.000	0.000	0.000	0.131	0.700	0.168	0.000	8.809	0.168
iCRM	0.000	0.000	0.000	0.061	0.680	0.258	0.000	8.808	0.258
bCRM	0.000	0.000	0.015	0.136	0.570	0.280	0.000	8.641	0.280
mCRM	0.000	0.000	0.000	0.059	0.582	0.359	0.000	8.407	0.359
CRM	0.000	0.000	0.000	0.070	0.683	0.248	0.000	8.638	0.248
Keyboard	0.000	0.000	0.000	0.151	0.704	0.145	0.000	7.347	0.145
BOIN	0.000	0.000	0.001	0.147	0.678	0.175	0.000	6.108	0.175
Scenario8	0.2	0.3	0.45	0.55	0.6	0.7			
kCRM	0.378	0.486	0.115	0.009	0.001	0.000	0.012	10.518	0.124
iCRM	0.234	0.565	0.176	0.012	0.001	0.000	0.012	10.561	0.189
bCRM	0.293	0.504	0.167	0.023	0.002	0.000	0.012	10.413	0.191
mCRM	0.403	0.029	0.528	0.027	0.001	0.000	0.012	11.625	0.556
CRM	0.261	0.557	0.160	0.009	0.001	0.000	0.012	10.429	0.170
Keyboard	0.257	0.537	0.150	0.019	0.001	0.000	0.037	10.339	0.170
BOIN	0.262	0.530	0.151	0.019	0.002	0.000	0.037	9.863	0.172
Scenario9	0.3	0.45	0.5	0.6	0.7	0.8			
kCRM	0.795	0.104	0.015	0.001	0.000	0.000	0.085	11.634	0.120
iCRM	0.700	0.195	0.019	0.001	0.000	0.000	0.085	11.686	0.215
bCRM	0.717	0.168	0.029	0.002	0.000	0.000	0.085	11.570	0.198
mCRM	0.669	0.005	0.238	0.003	0.000	0.000	0.084	11.957	0.247
CRM	0.720	0.179	0.016	0.001	0.000	0.000	0.085	11.577	0.196
Keyboard	0.630	0.154	0.023	0.002	0.000	0.000	0.192	10.937	0.179
BOIN	0.633	0.152	0.020	0.002	0.000	0.000	0.193	10.748	0.174
Scenario10	0.5	0.6	0.7	0.8	0.85	0.9			
kCRM	0.264	0.001	0.000	0.000	0.000	0.000	0.735	10.468	0.001
iCRM	0.262	0.003	0.000	0.000	0.000	0.000	0.735	10.471	0.003
bCRM	0.261	0.002	0.000	0.000	0.000	0.000	0.737	10.462	0.002
mCRM	0.227	0.000	0.041	0.000	0.000	0.000	0.732	10.677	0.041
CRM	0.261	0.002	0.000	0.000	0.000	0.000	0.737	10.449	0.002
Keyboard	0.129	0.002	0.000	0.000	0.000	0.000	0.868	7.898	0.002
BOIN	0.128	0.002	0.000	0.000	0.000	0.000	0.870	7.834	0.002

In scenario 1, the MTD is at the third dose level. Though the traditional CRM achieved 68.34% of patients assigned to the MTD, the mCRM outperformed it with a higher risk of allocating patients at a higher dose level. The iCRM achieved the third-best results. In addition, results of Keyboard and BOIN indicated well-performed implementations of these methods and a lower rate of average number of toxicities occurred. In scenario 6, the mCRM achieved similar results. It achieved the highest percentage of correct selection but with a higher percentage of patients allocated in dose levels considered to be unacceptable. As a result, a higher rate of toxicities occurred in the experiment. The fourth dose level

was chosen as the MTD in scenario 2 as the highest dose level that is tolerable enough for clinicians; the iCRM defeated other methods with the highest proportion of right selection which was at 58.42%. The CRM performed slightly better than the other proposed methods at 57.31%. The mCRM achieved the lowest number of average toxicities in the CRM-based methods. Still, the safest method is BOIN, but it achieved the worst experimental results as well. In scenario 3, compared with other methods, the mCRM achieved a remarkable selection percentage at 53.07% which is around 8% higher than others. It also allocated the highest percentage of patients at the true MTD with a slightly higher number of toxicities. Similar results could be observed in scenario 4. With the highest dose level within the tolerable interval, the mCRM achieved 82.46% correct selection rate which is about 5% higher than others on average. The average number of toxicities of the mCRM was the lowest compared with the traditional CRM, kCRM, iCRM, and bCRM. Scenario 5 was quite similar to scenario 4 and the same results were drawn for the data.

Table 4. Simulation results of allocation percentage at each dose level [scenarios 1–5].

	Allocation Percentage at Dose Level					
Designs	**1**	**2**	**3**	**4**	**5**	**6**
Scenario1						
kCRM	0.137	0.220	0.437	0.175	0.030	0.003
iCRM	0.136	0.220	0.439	0.176	0.027	0.003
bCRM	0.140	0.225	0.438	0.171	0.025	0.002
mCRM	0.173	0.097	0.406	0.298	0.023	0.002
CRM	0.144	0.222	0.434	0.173	0.025	0.002
Keyboard	0.133	0.270	0.429	0.147	0.019	0.001
BOIN	0.173	0.344	0.347	0.123	0.013	0.001
Scenario2						
kCRM	0.106	0.100	0.120	0.218	0.309	0.147
iCRM	0.106	0.100	0.120	0.221	0.304	0.149
bCRM	0.106	0.100	0.126	0.221	0.309	0.138
mCRM	0.118	0.090	0.109	0.247	0.312	0.124
CRM	0.109	0.097	0.123	0.226	0.306	0.139
Keyboard	0.106	0.118	0.139	0.231	0.275	0.132
BOIN	0.129	0.148	0.179	0.260	0.202	0.082
Scenario3						
kCRM	0.121	0.138	0.252	0.296	0.154	0.038
iCRM	0.121	0.139	0.254	0.299	0.149	0.039
bCRM	0.123	0.141	0.261	0.296	0.145	0.035
mCRM	0.149	0.096	0.196	0.390	0.139	0.030
CRM	0.126	0.138	0.257	0.301	0.143	0.035
Keyboard	0.121	0.179	0.278	0.264	0.124	0.035
BOIN	0.161	0.239	0.285	0.220	0.077	0.018
Scenario4						
kCRM	0.089	0.086	0.096	0.123	0.212	0.393
iCRM	0.089	0.086	0.096	0.124	0.210	0.395
bCRM	0.089	0.086	0.099	0.123	0.215	0.386
mCRM	0.091	0.085	0.093	0.132	0.230	0.369
CRM	0.090	0.086	0.097	0.126	0.214	0.387
Keyboard	0.090	0.101	0.118	0.133	0.209	0.349
BOIN	0.104	0.126	0.146	0.170	0.218	0.236
Scenario5						
kCRM	0.106	0.127	0.179	0.215	0.210	0.163
iCRM	0.106	0.128	0.179	0.217	0.206	0.164
bCRM	0.107	0.129	0.191	0.210	0.204	0.158
mCRM	0.123	0.097	0.164	0.262	0.209	0.146
CRM	0.108	0.127	0.184	0.219	0.203	0.158
Keyboard	0.114	0.182	0.198	0.186	0.160	0.160
BOIN	0.164	0.229	0.219	0.184	0.117	0.085

Table 5. Simulation results of allocation percentage at each dose level [scenarios 6–10].

	Allocation Percentage at Dose Level					
Designs	**1**	**2**	**3**	**4**	**5**	**6**
Scenario6						
kCRM	0.104	0.194	0.467	0.202	0.032	0.002
iCRM	0.103	0.193	0.469	0.203	0.029	0.002
bCRM	0.104	0.200	0.468	0.198	0.028	0.002
mCRM	0.113	0.091	0.424	0.344	0.026	0.002
CRM	0.106	0.200	0.466	0.200	0.027	0.002
Keyboard	0.105	0.266	0.454	0.156	0.019	0.001
BOIN	0.138	0.350	0.367	0.131	0.013	0.001
Scenario7						
kCRM	0.086	0.084	0.085	0.134	0.371	0.241
iCRM	0.086	0.084	0.085	0.135	0.368	0.242
bCRM	0.086	0.084	0.086	0.137	0.385	0.222
mCRM	0.086	0.084	0.085	0.146	0.400	0.199
CRM	0.086	0.084	0.086	0.139	0.382	0.224
Keyboard	0.086	0.089	0.096	0.211	0.371	0.147
BOIN	0.091	0.100	0.115	0.286	0.296	0.112
Scenario8						
kCRM	0.415	0.368	0.174	0.037	0.005	0.001
iCRM	0.403	0.379	0.176	0.037	0.005	0.001
bCRM	0.429	0.362	0.170	0.033	0.005	0.000
mCRM	0.471	0.096	0.354	0.075	0.004	0.000
CRM	0.427	0.364	0.170	0.035	0.004	0.000
Keyboard	0.379	0.418	0.170	0.030	0.003	0.000
BOIN	0.454	0.378	0.143	0.023	0.002	0.000
Scenario9						
kCRM	0.748	0.183	0.058	0.010	0.001	0.000
iCRM	0.738	0.193	0.058	0.011	0.001	0.000
bCRM	0.758	0.175	0.056	0.009	0.001	0.000
mCRM	0.736	0.059	0.180	0.023	0.001	0.000
CRM	0.758	0.175	0.056	0.010	0.001	0.000
Keyboard	0.700	0.244	0.049	0.007	0.000	0.000
BOIN	0.729	0.227	0.039	0.005	0.000	0.000
Scenario10						
kCRM	0.940	0.048	0.011	0.001	0.000	0.000
iCRM	0.939	0.049	0.011	0.001	0.000	0.000
bCRM	0.943	0.046	0.010	0.001	0.000	0.000
mCRM	0.907	0.026	0.065	0.001	0.000	0.000
CRM	0.945	0.044	0.010	0.001	0.000	0.000
Keyboard	0.912	0.083	0.005	0.000	0.000	0.000
BOIN	0.920	0.076	0.004	0.000	0.000	0.000

In scenario 7, the risk of overdose in the mCRM reduced with the cost of low accuracy. Keyboard beat the other methods in recommending the correct MTD and the kCRM was the second-best method, accordingly. The iCRM outperformed others in scenario 8 with a correct selection rate at 56.48%. With a remarkable correct selection rate of 79.49%, the kCRM achieved the best performance among others. For scenario 10, the model-assisted methods generally beat the model-based methods in safety. Specifically, BOIN mostly adopted early termination in simulations, indicating its safety.

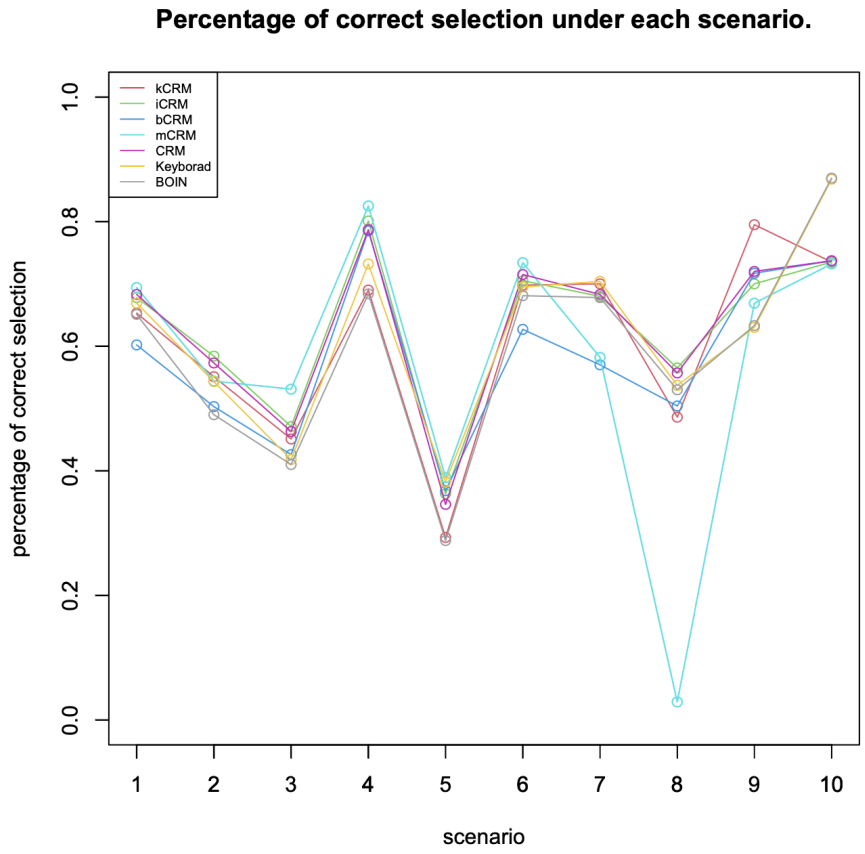

Figure 2. Percentage of correct selection of each scenario.

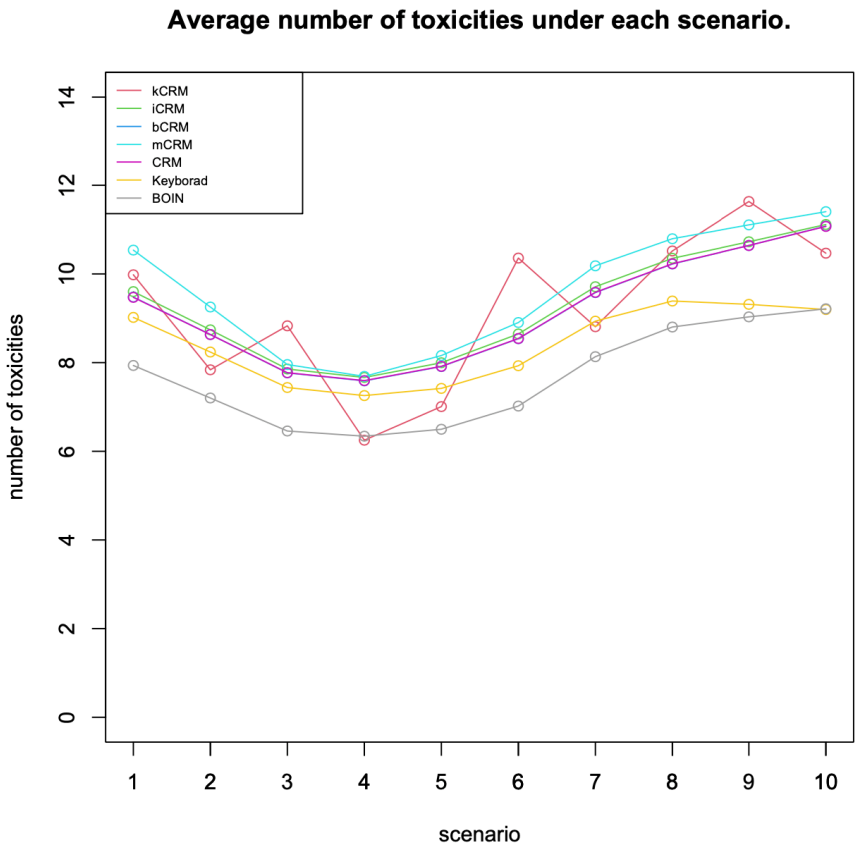

Figure 3. Average number of toxicities.

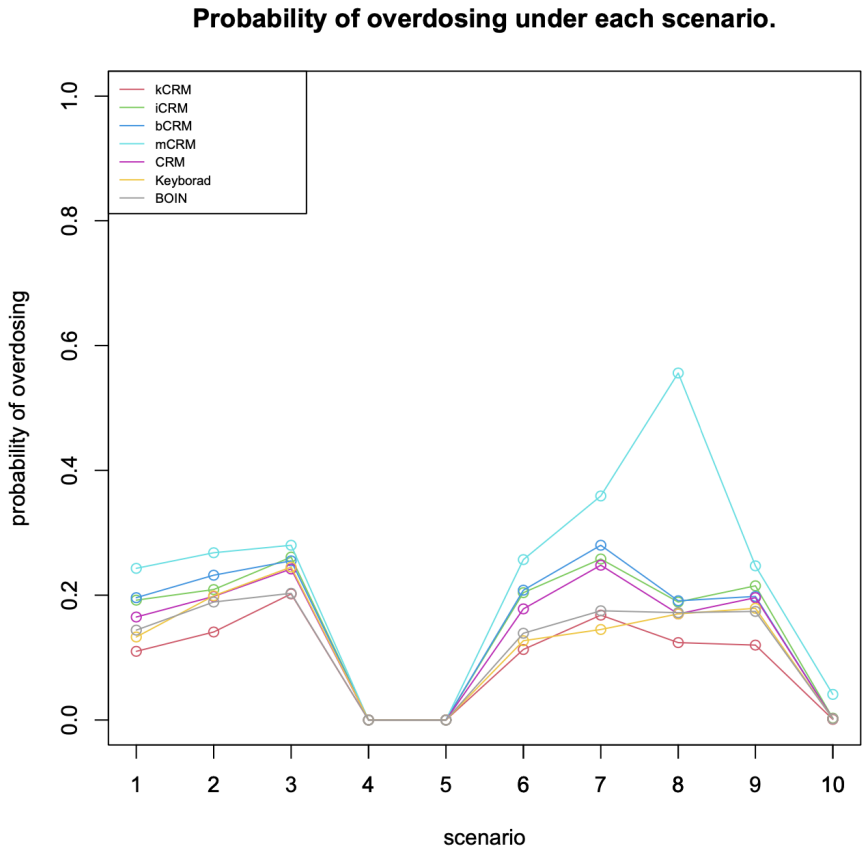

Figure 4. Probability of overdosing.

In summary, our proposed CRM-based methods achieved the highest performance in most of the scenarios. Though they did not completely beat other methods universally, comparable results could be observed in every scenario. The mCRM outperformed in half the number of the scenarios proposed but with slightly higher probabilities of overdose which could be considered acceptable. For certain situations like scenario 8 where the skeleton is wrongfully supposed compared with true toxicity probabilities and the distance between each probability is relatively small, the mCRM could fail to identify the true MTD and poor performance would be obtained. It seems reasonable to consider that the mCRM may be troubled with the mismatch of model and true toxicity probabilities, and a prior distribution of higher variance could mitigate the issue by allowing the model to fluctuate based on the latest observations with a larger range. According to empirical knowledge, we found that the mCRM is most suitable to be applied in situations where the toxicity probabilities of adjacent dose levels are relatively large and the true MTD is located at a higher dose level.

The other three methods are more stable compared to the mCRM, and iCRM performed better than others except the mCRM in most scenarios. In particular, it beat the traditional CRM frequently while retaining a similar level of safety. It could be considered a more robust and reliable method for finding a dose with an acceptable toxicity probability interval.

4.2. Sensitivity Test

In order to examine the performance of proposed designs, simulations with a maximum number of 54 patients have been conducted under each scenario with other settings unchanged [33]. Table 6 demonstrated the simulation results. The percentage of dose recommendations at true MTD, the rate of overdose, and the average number of DLTs are

used as criteria to inspect the operating characteristics. Owing to the Bayesian statistical framework and additional information acquired from extra patients, all the models have been improved significantly in terms of accuracy and safety in most of the situations. Specifically, the bCRM achieved the highest rate of correct selection in scenarios 1, 3, 4, and 6 but failed to selected the true MTD in scenario 8. The escalation schedule of the bCRM may improve the accuracy with more information at higher doses. It would also lead to a higher overdose rate and failure when the model was misspecified. The kCRM beat other models in scenarios 7 and 9 and the iCRM outperformed others in scenario 2. The application of the kCRM, iCRM, and mCRM would lead to a less average number of DLTs compared to the CRM except in scenario 10, which may be resulted from the more prudent dose allocation rules we proposed.

Table 6. Simulation results of percentage of dose recommendations at true MTD, rate of overdose, and average number of DLTs.

Percentage of Dose Recommendations at True MTD							
Scenarios	**kCRM**	**iCRM**	**bCRM**	**mCRM**	**CRM**	**Keyboard**	**BOIN**
Scenario1	0.770	0.793	0.837	0.777	0.798	0.776	0.759
Scenario2	0.680	0.711	0.665	0.693	0.696	0.668	0.644
Scenario3	0.540	0.574	0.592	0.538	0.563	0.511	0.492
Scenario4	0.765	0.871	0.908	0.784	0.854	0.822	0.792
Scenario5	0.324	0.383	0.434	0.259	0.400	0.500	0.425
Scenario6	0.806	0.809	0.869	0.811	0.819	0.799	0.775
Scenario7	0.838	0.806	0.719	0.823	0.815	0.809	0.798
Scenario8	0.559	0.646	0.025	0.597	0.654	0.619	0.623
Scenario9	0.826	0.764	0.649	0.810	0.762	0.652	0.732
Scenario10	0.881	0.864	0.881	0.881	0.882	0.942	0.892
Rate of Overdose							
Scenarios	**kCRM**	**iCRM**	**bCRM**	**mCRM**	**CRM**	**Keyboard**	**BOIN**
Scenario1	0.062	0.139	0.123	0.060	0.101	0.081	0.099
Scenario2	0.088	0.154	0.215	0.078	0.140	0.149	0.154
Scenario3	0.160	0.210	0.233	0.158	0.204	0.229	0.217
Scenario4	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Scenario5	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Scenario6	0.065	0.147	0.129	0.063	0.107	0.075	0.096
Scenario7	0.071	0.162	0.255	0.097	0.132	0.075	0.096
Scenario8	0.079	0.147	0.565	0.080	0.124	0.136	0.145
Scenario9	0.075	0.159	0.243	0.090	0.139	0.134	0.142
Scenario10	0.000	0.000	0.029	0.000	0.000	0.000	0.000
Average Number of DLTs							
Scenarios	**kCRM**	**iCRM**	**bCRM**	**mCRM**	**CRM**	**Keyboard**	**BOIN**
Scenario1	15.207	15.213	17.110	15.191	15.462	14.810	13.572
Scenario2	12.916	12.777	12.689	12.381	13.180	12.794	11.274
Scenario3	14.028	13.933	14.319	13.913	14.165	13.796	12.446
Scenario4	10.946	10.872	10.825	8.933	11.176	10.567	9.268
Scenario5	11.317	11.277	11.264	10.890	11.446	11.314	10.321
Scenario6	15.690	15.668	17.904	15.674	15.936	14.721	13.317
Scenario7	14.338	14.166	13.974	13.467	14.591	12.568	11.017
Scenario8	15.558	15.855	17.783	15.664	15.932	15.755	15.412
Scenario9	16.824	17.340	17.622	16.922	17.050	15.850	16.944
Scenario10	11.982	13.022	11.890	11.983	11.956	8.766	11.799

5. Conclusions

Phase I clinical trials aim at detecting the dose level with tolerable toxicity probability. Considering the preciseness and scarcity of patients enrolled, the information collected from each patient should be highly valued and fully utilized. Model-based and model-assisted methods have become more practical for the industry because the application of statistical knowledge offers a chance to take the best advantage of the clinical data. In addition, targets of a single toxicity probability may not always be obtainable considering the conservativeness of clinical trials. A trial design with a tolerable toxicity interval and fully utilized observations from patients are desired for industrial practice.

In this research, we have proposed a series of CRM-based methods to study the dose-finding research once a tolerable toxicity probability interval is given. By alternating the dose allocation methods and escalation rule, we would be able to construct novel models called the iCRM, kCRM, mCRM, and bCRM. Simulations were conducted to compare the operational performance of these methods and standard methods of model-based and model-assisted methods. Though with the same CRM structure, different allocation rules would result in different operating characteristics. In terms of dose-finding accuracy, the higher percentage of correct selection indicates the superiorities of our methods while maintaining certain performance in overdose control which can be measured by the overdose rate and number of toxicities occurred. The iCRM and mCRM have been recommended for application with corresponding situations. In addition, the simulation results provided insights about the operating characteristics. For instance, the mCRM tends to outperform others when the true toxicity probabilities of each dose spread widely. A sensitivity test considering a larger sample size has been performed and the results indicate that the proposed methods possess remarkable properties from the perspective of accuracy and safety.

Author Contributions: Conceptualization, J.X. and W.Z.; methodology, J.X. and W.Z.; software, J.X.; validation, J.X., W.Z., C.W. and R.L.; formal analysis, J.X. and W.Z.; investigation, J.X., W.Z., C.W. and R.L.; writing—original draft preparation, J.X. and W.Z.; writing—review and editing, J.X., W.Z., C.W. and R.L.; visualization, J.X.; supervision, W.Z., C.W. and R.L.; funding acquisition, W.Z. All authors have read and agreed to the published version of the manuscript.

Funding: The study is supported by the Science and Technology Project of Chongqing Education Commission [No. KJQN202100804], Research Development Fund (RDF) [No. RDF-24-01-069] and School Research Grant (SGR), Xi'an Jiaotong-Liverpool University.

Data Availability Statement: The data were acquired through simulations.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Storer, B.E. Design and analysis of phase I clinical trials. *Biometrics* **1989**, *45*, 925–937. [CrossRef] [PubMed]
2. Lin, Y.; Shin, W. Statistical properties of the traditional algorithm-based designs for phase I cancer clinical trials. *Biostatistics* **2011**, *2*, 203–215. [CrossRef] [PubMed]
3. Skolnik, J.M.; Barrett, J.S.; Jayaraman, B.; Patel, D.; Adamson, P.C. Shortening the timeline of pediatric phase I trials: The rolling six design. *J. Clin. Oncol.* **2008**, *26*, 190–195. [CrossRef] [PubMed]
4. Simon, R.; Freidlin, B.; Rubinstein, L.; Arbuck, S.G. Accelerated titration designs for phase I clinical trials in oncology. *J. Natl. Cancer Inst.* **1997**, *89*, 1138–1147. [CrossRef]
5. Liu, M.; Wang, S.J.; Ji, Y. The i3+3 design for phase I clinical trials. *J. Biopharm. Stat.* **2019**, *30*, 294–304. [CrossRef]
6. O'Quigley, J.; Pepe, M.; Fisher, L. Continual reassessment method: A practical design for phase I clinical trials in cancer. *Biometrics* **1990**, *46*, 33–48. [CrossRef]
7. Goodman, S.N.; Zahurak, M.Y.; Piantadosi, S. Some practical improvements in the continual reassessment method for phase I studies. *Stat. Med.* **1995**, *14*, 1149–1161. [CrossRef]
8. O'Quigley, J.; Shen, L. Continual reassessment method: A likelihood approach. *Biometrics* **1996**, *52*, 673–684. [CrossRef]

9. Daimon, T.; Zohar, S.; O'Quigley, J. Posterior maximization and averaging for Bayesian working model choice in the continual reassessment method. *Stat. Med.* **2011**, *30*, 1563–1573. [CrossRef]
10. Garrett-Mayer, E. The continual reassessment method for dose-finding studies:a tutorial. *Clin. Trials* **2006**, *3*, 57–71. [CrossRef]
11. Wheeler, G.M.; Mander, A.P.; Bedding, A.; Brock, K.; Cornelius, V.; Grieve, A.P.; Jaki, T.; Love, S.B.; Odondi, L.; Weir, C.J.; et al. How to design a dose-finding study using the continual reassessment method. *BMC Med. Res. Methodol.* **2019**, *19*, 18. [CrossRef] [PubMed]
12. Gasparini, M.; Eisele, J. A curve-Free method for phase I clinical trials. *Biometrics* **2000**, *56*, 609–615. [CrossRef] [PubMed]
13. Yin, G.; Yuan, Y. Bayesian model averaging continual reassessment method in phase I clinical trials. *J. Am. Stat. Assoc.* **2009**, *104*, 954–968. [CrossRef]
14. Pan, H.; Yuan, Y. A default method to specify skeletons for Bayesian model averaging continual reassessment method for phase I clinical trials. *Stat. Med.* **2017**, *36*, 266–279. [CrossRef] [PubMed]
15. Zhang, W.; Wang, X.; Yang, P. A new design of the continual reassessment method. *Commun. Stat.-Simul. Comput.* **2019**, *50*, 2015–2024. [CrossRef]
16. Liu, S.; Yin, G.; Yuan, Y. Bayesian data augmentation dose finding with continual reassessment method and delayed toxicity. *Ann. Appl. Stat.* **2013**, *7*, 2138–2156. [CrossRef]
17. Yuan, Y.; Yin, G. Robust EM continual reassessment method in oncology dose finding. *J. Am. Stat. Assoc.* **2011**, *106*, 818–831. [CrossRef]
18. Abbas, R.; Rossoni, C.; Jaki, T.; Paoletti, X.; Mozgunov, P. A comparison of phase I dose-finding designs in clinical trials with monotonicity assumption violation. *Clin. Trials* **2020**, *17*, 522–534. [CrossRef]
19. Cheung, Y.K.; Chappell, R. Sequential designs for phase I clinical trials with late-onset toxicities. *Biometrics* **2000**, *56*, 1177–1182. [CrossRef]
20. Ji, Y.; Li, Y.; Bekele, N. Dose-finding in phase I clinical trials based on toxicity probability intervals. *Clin. Trials* **2017**, *4*, 235–244. [CrossRef]
21. Ji, Y.; Liu, P.; Li, Y.; Bekele, N. A modified toxicity probability interval method for dose-finding trials. *Clin. Trials* **2010**, *7*, 653–663. [PubMed]
22. Guo, W.; Wang, S.J.; Yang, S.; Lynn, H.; Ji, Y. A Bayesian interval dose-finding design addressing Ockham's razor: MTPI-2. *Contemp. Clin. Trials* **2017**, *58*, 23–33. [CrossRef]
23. Yan, F.; Mandrekar, S.J.; Yuan, Y. Keyboard: A novel Bayesian toxicity probability interval design for phase I clinical trials. *Clin. Cancer Res.* **2017**, *23*, 3994–4003. [CrossRef] [PubMed]
24. Liu, S.; Yuan, Y. Bayesian optimal interval designs for phase I clinical trials. *J. R. Stat. Soc. Ser. C* **2014**, *64*, 507–523. [CrossRef]
25. Yuan, Y.; Lin, R.; Li, D.; Nie, L.; Warren, K.E. Time-to-Event Bayesian optimal interval design to accelerate phase I trials. *Clin. Cancer Res.* **2018**, *24*, 4921–4930. [CrossRef]
26. Zhou, Y.; Lin, R.; Lee, J.J.; Li, D.; Wang, L.; Li, R.; Yuan, Y. TITE-BOIN12: A Bayesian phase I/II trial design to find the optimal biological dose with late-onset toxicity and efficacy. *Stat. Med.* **2022**, *41*, 1918–1931. [CrossRef]
27. Yin, G. *Clinical Trial Design: Bayesian and Frequentist Adaptive Methods*; Wiley: Hoboken, NJ, USA, 2012.
28. Shen, L.; O'Quigley, J. Consistency of continual reassessment method under model misspecification. *Biometrics* **1996**, *83*, 395–405. [CrossRef]
29. Yin, G.; Zheng, S.; Xu, J. Fractional dose-finding methods with late-Onset toxicity in phase I clinical trials. *J. Biopharm. Stat.* **2014**, *23*, 856–870. [CrossRef]
30. Casella, G.; Berger, R.L. *Statistical Inference*, 2nd ed.; Duxbury Press: New York, NY, USA, 2001.
31. Babb, J.; Rogatko, A.; Zacks, S. Cancer phase I clinical trials: Efficient dose escalation with overdose control. *Stat. Med.* **1998**, *17*, 1103–1120. [CrossRef]
32. Clertant, M.; O'Quigley, J. Semiparametric dose finding methods: Special cases. *J. R. Stat. Soc. Ser. C Appl. Stat.* **2019**, *68*, 271–288. [CrossRef]
33. Mozgunov, P.; Jaki, T.; Gounaris, I.; Goddemeier, T.; Victor, A.; Grinberg, M. Practical implementation of the partial ordering continual reassessment method in a Phase I combination-schedule dose-finding trial. *Stat. Med.* **2022**, *41*, 5789–5809. [CrossRef]