Kuwait J. Sci. **47 (1)** pp. 22-32, 2020

Investigating dose finding methods and different priors in Bayesian continual reassessment method

Tuba Koc[1,*], Mehmet A. Cengiz[2]

[1]*Dept. of Statistics, Cankiri Karatekin, University Cankiri, Turkey*

[2]*Dept. of Statistics, Ondokuz Mayis, University Samsun, Turkey*

[*]*Corresponding author: tubakoc@karatekin.edu.tr*

Abstract

A well-designed and properly analyzed clinical trial is a powerful tool for the development of new drugs. Clinical trials are studies that explore whether a treatment, drug or device is safe and effective for humans. These studies can show which drug development method works best for certain diseases. The first step in drug discovery (phase I) is very important to determine maximum tolerated dose (MTD).

In the first part of the study, the classical 3+3 design, Continual Reassessment Method (CRM), and Bayesian Continual Reassessment Method (B-CRM) are compared in terms of selection probability of MTD and the number of treated patients. Among these designs, the B-CRM produced better results than the 3+3 and the CRM. In the second part of the study, we considered different model structures and priors in the B-CRM design. We considered three model structures; power model, hyperbolic tangent model, logit model and three different prior distributions; gamma, uniform and lognormal prior respectively. It was found that if power or hyperbolic tangent model structure and uniform prior were selected, the MTD selection rates were the highest in B-CRM.

Keywords: Bayesian inference; clinical trial; maximum tolerated dose; prior selection.

1. Introduction

Clinical trials are studies that are applied directly on humans as test subjects under convenient circumstances. They mainly focus on experimental medicine, new treatment methods, medical devices and clinical processes. Clinical trials play an important role in developing new treatment methods, helping researchers to find best medicine dosage levels. They also contribute to finding potential side effects of medicines, which safeguards patients.

Phase I trials focus on four major rules: the ethics of the trial, the initial selection of the dose, the speed of efficiency of the dose, the possibility of target toxicity, and the sufficiency of the trial design. Phase I trials aim to determine the maximum tolerated dose (MTD) of a suggested dose amount for Phase II studies and to calculate its safety and tolerability. The MTD is the amount that is closest to the toxicity levels which is pre-determined by researchers. The recommended dose should be at the MTD level or below it. The dose-limiting toxicity (DLT) is the toxicity that blocks the further administration of the drug at current dose level.

There are many statistical algorithms and model-based methods for dose finding studies in Phase I. One is the standard 3+3 design, an algorithm-based approach, which finds the highest dose of MTD with its toxicity probability less than 33% (Storer, 1989). With its simple and easy applicability, the 3+3 design is the most commonly used rule-based design. Le Tourneau *et al*. (2009) explained drawbacks of the various dose escalation methods and recommended model-based designs rather than the 3+3 design. Hansen *et al*. (2014) how that the 3+3 design identified the correct dose with an acceptable level of precision, but no single escalation method was proven superior in all conditions. Nie *et al*. (2016) showed that in the 3+3 design, the important dose-ranging clinical trials have not been routinely performed. Ulas & Karaman (2018) found that the 3+3 design produced the worst performance compared to model-based designs based on different simulation studies.

The most commonly used model-based design is continual reassessment methods (CRM) (O'Quigley *et al*., 1990). As more toxicity data is collected during the trial, CRM updates the estimation of toxicity probabilities for all doses. Each new cohort of patients is sequentially

assigned to the most appropriate dose based on the updated toxicity probabilities. The MTD is determined when the total sample size is exhausted. Babb *et al*. (1998) developed an increasing dose scheme which controls overdose so as to protect patients from excessive toxicity. Heyd & Carlin (1999) developed an advanced application for CRM in order to stop an experiment earlier when the probability interval is narrow enough for MTD. Ishizuka & Ohashi (2001) proposed a method that reduces the number of patients that are treated at the next dose levels of the MTD by using Bayesian computation for the CRM method during phase I clinical trials. Leung & Wang (2002) applied a decision theory to optimize the number of patients allocated for the highest dose using tolerated toxicity. Garret-Mayer (2006) discussed qualities that define a CRM design and showed examples of CRMs and standard designs. Cheung (2011) explained widely the CRM and its extended versions for dose finding studying in his book.

In the Bayesian paradigm, Yin (2012) investigated clinical trial designs with Bayesian and frequents methodologies. Ji *et al*. (2012) consider the Bayesian continual reassessment method for patients with malignancies and when sample sizes were small. The authors considered three different scenarios: Early Stopping (ES), Fast Escalation (FE), and Bracketed MTD (BR). They found that the 3+3 design could not increase quickly in an FE scenario. However, model-based designs were more effective. Thus, model-based designs can be recommended for trials in which the number of patients is small. Sweeting *et al*. (2013) demonstrated, through example of the BCRM package, how a variety of possible designs can be easily implemented within R statistical software. In addition, the authors showed how properties of the design can be communicated to trial investigators using simple textual and graphical output obtained from the software.

Section 2 includes a review of a one-parameter CRM design, Bayesian Continual Reassessment Method (B-CRM) and stopping rules. The main focus of Section 3 is on how different priors for different model types affect the selection probability of the MTD. Simulation studies were carried out in order to compare CRM, B-CRM and 3+3 designs. A selection of different prior and model structures for B-CRM is considered by using two different simulation studies in order to examine true MTD rates. Finally, a brief discussion is given in Section 4.

2. Methods

2.1. One parameter CRM design

CRM links to the true toxicity probability in each dose with pre-determined toxicity probability, using a single-parameter model. As toxicity data is collected during the trial, CRM updates the estimation of toxicity probability of all doses constantly. Commonly, we expect toxicity to increase monotonically with respect to the dose. Let $p_1 < p_2 < \cdots < p_d$ be the predetermined toxicity probabilities of a set of d doses for the drug under consideration, which is often known as the skeleton of the CRM. Let ϕ_T be the target toxicity probability specified by the researcher and Y_j denote the binary toxicity outcome observed in the jth patient. The jth patient recruited to the trial, $y_j = 1$ denotes a DLT, for $j = 1,2, \dots, d$

CRM power model is as follows:

$$P(y_j = 1) = \pi_j(\alpha) = {p_j}^{\exp(\alpha)} \tag{1}$$

where a is the only unknown parameter (O'Quigley & Shen, 1996).

One parameter CRM hyperbolic tangent dose-response model as follows

$$P(y_j = 1) = \pi_j(\alpha) = \left[(\tanh(d_j) + 1)/2\right]^{\alpha} \tag{2}$$

where $\alpha \epsilon \mathbb{R}^+$ and d_j is the standardized dose at dose level j.

One parameter logistic CRM model which has a fixed intercept as follows

$$\pi_j(\alpha) = \frac{\exp(c+\alpha d_j)}{1+\exp(c+\alpha d_j)} \tag{3}$$

The fixed intercept c is set to 3 for ease of calculation as it mentioned in (O'Quigley & Chevret, 1991).

In order to increase the robust of design, multiple CRM models can be used after creating a different skeleton (Yin & Yuan 2009). Different skeletons estimate the toxicity profile of a medicine with a different prior knowledge. So that it guarantees the researcher to estimate the toxicity probability closest to the best estimations among all the candidate models.

2.2. Bayesian framework

Estimation of α parameter is desired by using statistical model defined with probability density function $p(y|\alpha)$

depending on $y = (y_1, \dots, y_d)$ data. According to the Bayesian approach, when information on distribution and probability is insufficient, α cannot be exactly determined. The fundamental of the Bayesian inference is given below:

1. For α, a probability distribution, represented as $f(\alpha)$ is formulized. This distribution is called a prior distribution.
2. For observed y dataset and when α is given, a likelihood function $L(\alpha; y)$ is determined to define the distribution of y given α

$$L(\alpha; y_1, \dots, y_d) = \prod_{j=1}^{d} \pi(d_j; \alpha)^{y_j} (1 - \pi(d_j; \alpha))^{1-y_j}$$

$f(\alpha|y)$ is the posterior distribution which is defined using the prior and the likelihood function as the following:

$$f(\alpha|y_1, \dots, y_d) = \frac{f(a)L(\alpha; y_1, \dots, y_d)}{\int f(\alpha)L(\alpha; y_1, \dots, y_d)d\alpha} \quad (4)$$

Then, all statistical inferences about α are obtained through the posterior distribution.

Above steps can be formulated for CRM model by using Bayesian approach.

When the posterior estimates of toxicity probabilities for all doses are updated, the recommended dose level for the next cohort of patients is the one that has a toxicity probability closest to the target ϕ_T. Thus, a new cohort of patients is assigned to dose level j^* such that

$$j^* = \underset{j \in (1,2,\dots,d)}{argmin} |f(\alpha|y) - \phi_T| \quad (5)$$

Denote the current dose level as j^{curr}.

- If $j^{curr} > j^*$, de escalate to the dose level $j^{curr} - 1$;
- If $j^{curr} < j^*$, escalate to the dose level $j^{curr} + 1$;
- Otherwise the dose stays at the same level for the next cohort of patient.

The trail is continued until the total sample size is reached, after which the dose with a posterior toxicity probability closest to ϕ_T is selected MTD (Yin, 2012).

Bayesian inference or modeling cannot be done without prior distribution. It is possible to take such prior distributions as Gamma, uniform and lognormal for the α parameter.

For instance, let α ~lognormal (μ, σ^2) prior and $f(\alpha)$ represent the prior distribution for CRM power model.

Then,

$$L(\alpha|y) \propto \prod_{j=1}^{d} \left(p_j{}^{\exp(\alpha)}\right)^{y_j} (1 - p_j{}^{\exp(\alpha)})^{1-y_j}$$

$$f(\alpha|y) = \int p_j{}^{\exp(\alpha)} \frac{\prod_{j=1}^{d} \left(p_j{}^{\exp(\alpha)}\right)^{y_j} (1 - p_j{}^{\exp(\alpha)})^{1-y_j} . f(\alpha)}{\int \prod_{j=1}^{d} \left(p_j{}^{\exp(\alpha)}\right)^{y_j} (1 - p_j{}^{\exp(\alpha)})^{1-y_j} . f(\alpha) d\alpha} \, d\alpha$$

2.3. Stopping criteria

In phase I trials, patients are continued to include in trial until the maximum sample size is reached or until it encounters one of the possible stopping rules. There are several stopping criteria in phase I dose finding studies. O'Quigley & Reiner (1998) proposed a stopping rule based on the probability that the remaining patients were included in the trial. According to this rule, together with the final recommended MTD, the same dose level is appointed and if, at any time, this predicted probability is high, then the trial is stopped. In this study, we considered two different criteria. First one is the O'Quigley & Reiner (1998) stopping criteria and the second one is the one of the commonly used practice stopping rule is as follows:

If $P(dose\ level\ in\ the\ level\ j > \phi_T) > 0.9$ (O'Quigley *et al.,* 1990) then the trial is terminated for safety concerns.

3. Simulation Studies

We compare the classical 3+3 design, CRM and B-CRM design using a pharmacokinetic phase I clinical trial that aim to determine the MTD for Antroquinonol ($C_{24}H_{38}O_{4}$) which is an active component of Anthodia camphorate. Anthodia camphorate is a unique mushroom of Taiwan, which has been used as a traditional medicine for protection of diverse health-related conditions. This clinical trial studied six doses of the drug. It was taken orally, daily within 15 minutes after a breakfast at the assigned dose level: 50, 100, 200, 300, 450, 600 mg/day for 4 weeks. A total of 21 assessable patients were used for dose escalation.

We took the MTD as the dose with a DLT rate of 30%, it can be seen that different values can be used for DLT in the literature, (Iasonos & O'Quigley, 2014) The values

can be prepared five sets of initial guesses of toxicity probabilities:

$$(p_1, p_2, p_3, p_4, p_5, p_6) = \begin{cases} 0.05, 0.10, 0.20, 0.30, 0.35, 0.40 \\ 0.03, 0.13, 0.30, 0.48, 0.64, 0.76 \\ 0.01, 0.03, 0.09, 0.18, 0.30, 0.42 \\ 0.10, 0.20, 0.30, 0.40, 0.50, 0.60 \\ 0.15, 0.30, 0.46, 0.61, 0.73, 0.81 \end{cases}$$

The probabilities of skeletons are simulated by using R software with "dfcrm" (dose- finding by the continual reassessment method) package.

These probabilities have effect on the true MTD when the range between target toxicity and pre-determined toxicity are higher. We generated several skeletons for all CRM models. And we use the skeleton for each CRM model based on the best performance produced. The skeleton which produced the best selection probability is considered in our study. We took the cohort size 3 and treated the first cohort of patients at the lowest dose level. We created five scenarios for five skeletons and for each scenario; we carried out 10000 simulated trials for each scenarios.

In this paper, the classical 3+3 design, CRM and Bayesian CRM methods are compared in order to select the true dose level as MTD and the numbers of patients who are treated at each dose level are highlighted. We used "East 6.3.1" program for MTD dose selection of probability in the 3+3 design and CRM. Furthermore, we investigated the different model structure and different priors of B-CRM method in order to determine the true dose level as MTD. In Table 1, we list the dose selection probability and the average number of patients treated at each dose using the 3+3, CRM and B-CRM design, respectively, for five scenarios at different dose levels.

Table1. Simulation results under different five scenarios for 3+3, CRM and B-CRM.

			Dose levels						Average patient
			50	**100**	**200**	**300**	**450**	**600**	
	Methods	**True toxicity rate**	**0.03**	**0.05**	**0.06**	**0.1**	**0.3**	**0.5**	
Scenario 1	3+3	Probability selection	1.74	13.15	15.56	40.62	26.2	2.76	
		number of patients treated	3.173	2.452	2.561	4.324	4.38	4.101	21
	CRM	Probability selection	3.3	5.1	20.1	35.1	36.4	0	
		number of patients treated	3.099	3.158	4.697	4.467	5.57	0	21
	B-CRM	Probability selection	0	0	0.04	18.8	73	8.23	
		number of patients treated	2.7	3.09	3.4	5.03	6.49	0.3	21
	Methods	**True toxicity rate**	**0.15**	**0.2**	**0.25**	**0.3**	**0.35**	**0.4**	
Scenario 2	3+3	Probability selection	24.35	29.37	22.4	16.86	6.88	0.24	
		number of patients treated	3.223	3.008	4.145	4.119	3.05	3.454	21
	CRM	Probability selection	0	16.5	58.9	24.6	0	0	
		number of patients treated	3.895	5.896	5.977	5.232	0	0	21
	B-CRM	Probability selection	0	0.4	5.2	94.4	0	0	
		number of patients treated	3.2	3.6	4.9	9.3	0	0	21
	Methods	**True toxicity rate**	**0.01**	**0.3**	**0.55**	**0.65**	**0.8**	**0.95**	
Scenario 3	3+3	Probability selection	31.8	57.36	10.33	0.51	0	0	
		number of patients treated	4.015	5.324	4.245	3.77	3.64	0	21
	CRM	Probability selection	30.07	44.7	21.01	4.02	0.2	0	
		number of patients treated	3.006	3.926	4.943	5.459	3.66	0	21
	B-CRM	Probability selection	47	52.8	0.05	0	0	0	
		number of patients treated	7.1	10.4	3.5	0	0	0	21

	Methods	True toxicity rate	0.05	0.09	0.16	0.21	0.23	0.24	
Scenario 4	3+3	Probability selection	4.5	13.28	20.99	19.3	8.09	1.35	
		number of patients treated	3.481	3.448	3.392	3.422	4.1	3.15	21
	CRM	Probability selection	0.04	0.02	4.5	3.7	20.9	70.6	
		number of patients treated	2.7	2.9	3.1	3.1	3.287	5.91	21
	B-CRM	Probability selection	0	0.12	1.4	3.2	19.7	74.6	
		number of patients treated	1.7	2.24	3.3	3.1	3.36	7.3	21
	Methods	**True toxicity rate**	**0.5**	**0.6**	**0.6**	**0.7**	**0.8**	**0.9**	
Scenario 5	3+3	Probability selection	27.4	2.17	0.25	0	0	0	
		number of patients treated	4.442	3.451	4.106	3	6	0	21
	CRM	Probability selection	26	23	7.9	0	0	0	
		number of patients treated	3.78	5.233	5.418	0	0	0	13
	B-CRM	Probability selection	19.3	0.01	0	0	0	0	
		number of patients treated	4.47	1.17	0	0	0	0	5

The true probabilities that given in Table 1 are set based on some notion of testing different "dose-response" scenarios. Ji *et al.,* (2012) have explained more clearly. Suppose that CRM model is one parameter power model and for all the methods except the 3+3 design, the trial is stopped when the maximum sample size is reached. We took the prior distributions of α as a Gamma for BCRM.

In the first scenario, simulations are carried out by using skeleton 1 which is obtained from CRM skeletons. While fifth dose is the maximum tolerated dose, selection probability of this dose as MTD differs for three different methods. The 3+3 design has the lowest probability to select the MTD with 26%. When dose selection rates are considered for the 3+3 design, it tends to choose the forth dose as MTD. Selection probability, where the dose is MTD, is 36% in CRM approach and selection probability, where the dose is MTD, is 73% for B-CRM approach. B-CRM approach gives the best performance in scenario 1. When treated patients for each dose are considered, generally all dose levels show similar results.

Simulations are carried out using skeleton 2 for scenario 2 and the fourth dose is found as the maximum tolerated dose. Exclusively, B-CRM selected dose 4 as true the MTD. Moreover, when the number of patients treated at each dose is examined, there is no toxic dose more than the MTD for CRM and B-CRM. In contrast, there are 6 patients who are treated at a dose which has a higher toxic level than the MTD in the 3+3 design.

In scenario 3, the second dose is found as maximum tolerated dose and the selection probability of this dose as MTD is similar for three different methods. When treated patients for each dose are considered, the 3+3 design has the highest probability to select MTD dose. However, when the number of patients treated at each dose is taken into account, patients are treated at higher toxic dose levels at B-CRM and CRM approach. Thus, although selection probability of MTD dose in B-CRM is lower than the highest selection probability in scenario, it is a better approach for treating fewer patients at extreme dose levels.

In scenario 4, the right dose toxicity probabilities determined for all levels of toxicity is lower than probability of the target toxicity. In such a situation, the trial is expected to be stopped with a rapid dose escalation. The B-CRM and CRM showed a good performance in this scenario because they constantly update existing toxicity possibilities for each dose level. The scenario 5, the right dose toxicity probabilities determined for all levels of toxicity is higher than probability of the target toxicity. So that, all doses are extremely high. In such a case, the trial is expected to immediately stop before selection of MTD.

So, the second stopping criteria, which is "if $P\left(dose\ level\ in\ the\ level\ j > \phi_T\right) > 0.9$", is used for this scenario. In this scenario, the best performance is showed by the B-CRM. In total, only 5 patients are treated at extreme highest dose and trial is stopped at an early stage.

In the most cases, the B-CRM method performed better than other trial designs. Therefore, different model structures and priors of B-CRM design are investigated and the selection probabilities of MTD in B-CRM are given in Table 2 and 3.

Table 2. Probability selection of MTD under the different three scenarios.

	Model	Prior distribution	Dose levels					
			50	100	200	300	450	600
Scenario 1	Power model	Gamma	0	0	0.01	17.1	74.8	0.8
		Uniform	0	0	0	17.5	82.5	0
		Log normal	0	0	0.03	20.7	46.2	32.8
	Hyperbolic tangent model	Gamma	0	0	0	21	68.8	10.2
		Uniform	0	0	0	16.9	83.1	0
		Log normal	0	0	0.01	20.8	47.9	31.2
	Logit model	Gamma	0	0	0.05	17.9	41	40.6
		Uniform	0	0	0.13	14.4	35.1	48.8
		Log normal	0	0.01	0.14	16.5	27.5	54.5
Scenario 2	Power model	Gamma	0	0	0.52	93.7	0.11	0
		Uniform	0	0	0.9	99.1	0	0
		Log normal	0	0	19.7	64.5	15.8	0
	Hyperbolic tangent model	Gamma	0	0	5.5	93.7	0.8	0
		Uniform	0	0	0.11	98.9	0	0
		Log normal	0	0	19.8	65.9	14.3	0
	Logit model	Gamma	0	0.05	25.6	54.3	18.8	0.08
		Uniform	0	0	28.2	46.4	25.4	0
		Log normal	0	0.33	30.2	38.8	22.2	0.4
Scenario 3	Power model	Gamma	48.8	50.1	0.07	0.04	0	0
		Uniform	0.05	99.5	0	0	0	0
		Log normal	16.6	68.5	14.9	0	0	0
	Hyperbolic tangent model	Gamma	48.9	50.5	0.4	0.02	0	0
		Uniform	0.08	99.2	0	0	0	0
		Log normal	16.2	72.5	11.1	0.02	0	0
	Logit model	Gamma	0.11	55.3	20	0.33	0.03	0
		Uniform	0.36	96.4	0	0	0	0
		Log normal	0.48	63.4	29.7	0.2	0.01	0

Fifth dose in the first scenario was the MTD. When the model structure was taken as a power-CRM model, selection probability of the MTD at the fifth dose level by using gamma prior was 75% and the probability of selection of MTD at the fifth dose level with uniform prior was 83% while probability of selection of the MTD at the fifth dose level with log normal prior was 46%. When the priors are taken as gamma and uniform distribution, B-CRM performed better results for power-CRM in the first scenario.

On the other hand, if the model structure was selected as hyperbolic tangent model, the probability selection of the MTD at the fifth dose by using gamma prior was 69% uniform prior was 83% and log normal prior was 48% respectively. The selection probabilities of hyperbolic tangent model structure with different priors produced similar results as in the power-CRM model. In both model structures, uniform prior selected the true dose level as MTD with the highest rates.

Additionally, when the model structure was taken as the logit model, probability selection of MTD at the fifth dose level with gamma prior was 41%, uniform prior was 35% and log-normal prior was 28%. The logit model structure for B-CRM produces low probability selection

of MTD for different priors when compared to power model and hyperbolic tangent model. In three different model structures and three different priors, the selection probability of the MTD at the fifth dose level was the highest, when the power or hyperbolic tangent model structure and uniform prior was used.

In scenario 2, the fourth dose was selected as MTD. The highest selection probability of the MTD at the fourth dose level was found, when the power or hyperbolic tangent model structure was taken with uniform prior. In scenario 3 similar results were obtained as in scenario 2.

Table 3. Probability selection of MTD under the different two scenarios.

	Model	Prior distribution	Dose levels					
			50	100	200	300	450	600
Scenario 4	Power model	Gamma	0	0.12	1.5	3.2	18.7	75.6
		Uniform	0	0	0	0	0.09	99.1
		Log normal	0	0.17	4.1	4.53	40.8	50.4
	Hyperbolic tangent model	Gamma	0	0	0	0.88	18.5	72.7
		Uniform	0	0	0	0	1	99
		Log normal	0	0	0	0	41.7	58.3
	Logit model	Gamma	0	0	0	0.68	38.6	54.5
		Uniform	0	0	0	0	3	97
		Log normal	0	0.1	0.47	15	33.3	51
Scenario 5	Power model	Gamma	18.9	0	0	0	0	0
		Uniform	11.5	0	0	0	0	0
		Log normal	21.2	0	0	0	0	0
	Hyperbolic tangent model	Gamma	19.7	0	0	0	0	0
		Uniform	13.4	0.02	0	0	0	0
		Log normal	25.2	0	0	0	0	0
	Logit model	Gamma	27	0.7	0.07	0	0	0
		Uniform	18.5	0.06	0.01	0	0	0
		Log normal	24.4	0.27	0.06	0	0	0

In scenario 4, the right toxicity probabilities determined for all dose levels are lower than the probability of selection of MTD dose. In such a case, the trial is expected to rapidly increase the dose. For scenario 4, when the three different models and three different priors are investigated, it seems that fastest dose increase happens when the proper prior is selected for three different structures.

In scenario 5, all doses are highly toxic. In such a case, the trial is expected to stop immediately without selecting the MTD.

In this scenario, we used the stopping rule "$P(dose\ level\ in\ the\ level\ j > \phi_T) > 0{,}9$" and that's why the MTD is not selected.

In the next simulation study, the pharmacokinetic phase I clinical trial that aim to determine the MTD for aldoxorubicin is applied. This clinical trial studied five doses of the drug. The assigned dose levels are: 130, 175, 180, 240, 320mg/day for 12 weeks. A total of 30 assessable patients were used for dose escalation. We took the MTD as the dose with a DLT rate of 25% and prepared five sets of initial guesses of toxicity probabilities:

$$(p_1, p_2, p_3, p_4, p_5) = \begin{cases} 0.01, 0.08, 0.25, 0.46, 0.64 \\ 0.08, 0.25, 0.46\ 0.64, 0.76 \\ 0.01, 0.05, 0.12, 0.25, 0.39 \\ 0.15, 0.25, 0.35, 0.46, 0.55 \\ 0.03, 0.06, 0.11, 0.17, 0.25 \end{cases}$$

and for each scenario we carried out 10000 simulated trials. It can be seen in Table 4.

Table 4. Probability selection of MTD under different four scenarios.

					Dose levels				
					130	**175**	**180**	**240**	**320**
				True toxicity rate	**0.03**	**0.05**	**0.10**	**0.25**	**0.30**
Scenario 1	**Model**	Power model	**Prior distribution**	Gamma	0.01	0.12	25	54.3	19.4
				Uniform	0	0.07	20.4	78.9	0
				Lognormal	0	0.08	15.2	53.3	30.7
		Hyperbolic tangent model		Gamma	0	0.15	23	54.1	21.4
				Uniform	0	0	10.1	89.9	0
				Lognormal	0	0.02	17.6	51.5	30.7
		Logit model		Gamma	0	0.06	13.7	50.7	35
				Uniform	0	0.17	15.4	61.1	21.8
				Lognormal	0.03	0.13	22.8	50.4	25.2
				True toxicity rate	**0.01**	**0.05**	**0.10**	**0.10**	**0.25**
Scenario 2	**Model**	Power model	**Prior distribution**	Gamma	0	0.01	0.09	0.95	89.5
				Uniform	0	0.01	0.05	6.1	93.3
				Lognormal	0	0	0.53	21.4	73.3
		Hyperbolic tangent model		Gamma	0	0.01	0.11	11.7	87.1
				Uniform	0	0.02	0.06	0.93	89.9
				Lognormal	0	0	0.18	14.8	83.4
		Logit model		Gamma	0	0	0.14	35.1	63.5
				Uniform	0	0.01	0.10	20.7	78.2
				Lognormal	0	0.02	0.14	25.3	73.1
				True toxicity rate	**0.01**	**0.25**	**0.45**	**0.65**	**0.80**
Scenario 3	**Model**	Power model	**Prior distribution**	Gamma	0.24	80.3	17.2	0.01	0
				Uniform	0.17	86.1	12.1	0.01	0
				Lognormal	0.20	81.9	16.1	0	0
		Hyperbolic tangent model		Gamma	0.20	80.3	17.7	0	0
				Uniform	0.18	83.3	14.7	0.02	0
				Lognormal	0.30	79.4	17.5	0.01	0
		Logit model		Gamma	0.32	69.8	26.7	0.03	0
				Uniform	0.16	78.9	19.2	0.03	0
				Lognormal	0.30	62.6	34.3	0.01	0
				True toxicity rate	**0.10**	**0.25**	**0.30**	**0.40**	**0.50**
Scenario 4	**Model**	Power model	**Prior distribution**	Gamma	0.94	55	30.1	0.55	0
				Uniform	0.76	65.4	19.5	0.75	0
				Lognormal	12.1	55.6	27.4	0.48	0.01
		Hyperbolic tangent model		Gamma	0.92	64	20.6	0.61	0.01
				Uniform	0.65	73.6	12.1	0.78	0
				Lognormal	12	58	25.3	0.44	0.03
		Logit model		Gamma	15.7	42.9	35.2	0.6	0.02
				Uniform	10.5	52	28.8	0.79	0.08
				Lognormal	15.6	44.3	33.1	0.70	0

In the first scenario, the fourth dose was the MTD. It is important that the model structure was critical on selecting the true toxicity probabilities. Results of the first scenario shows that the best performance was calculated by hyperbolic tangent model structure and uniform prior distribution. In contrast, logit model structure and log-normal prior was produced the worst performance. In the second and third scenario, the results of the simulation runs were similar. The best performance was examined when the power model structure and uniform prior distribution were selected. On the other hand, if the model structure was selected as logit model and prior distribution was selected as log normal, the probability selection of the true MTD was very low. Results of the last scenario were very similar to the first scenario. The best performance was calculated by hyperbolic tangent model structure and uniform prior distribution. Moreover, the worst performance was examined by logit model structure and gamma prior.

4. Conclusions

Determination of MTD in phase I trials plays an important role in developing a new drug. A wrong MTD dose selection can cause a trail failure. This results and a loss of time and investment. It can also delay critical medical treatment resulting in loss of life.

In this study, the classical 3+3 design, CRM and B-CRM designs were compared in order to select the true dose level as MTD. When these methods were compared, five different skeletons of CRM and five different scenarios were used. The results show that the B-CRM gives the best performance in scenario 1. For scenario 2, only the B-CRM selected the true dose as MTD. In the third scenario, the 3+3 design performed the best true toxicity probability. However, the 3+3 design were treated many patients with extremely toxic dose levels in this design. Although selection probability of the MTD dose in the B-CRM was lower than the highest selection probability, it is a better approach for treating patients at extreme dose levels. So, B-CRM appears to be a robust method. In scenario 4, all doses had lower toxicity levels than the target toxicity. Therefore, the last dose level is expected to be the MTD.B-CRM and CRM showed good performance in this scenario. In scenario 5, all doses were extremely toxic, so the trial was immediately stopped before the selection of the MTD. Generally, B-CRM showed a better performance than the 3+3 design and CRM.

We also considered power, hyperbolic tangent and logit models for gamma, uniform and lognormal priors via a simulation study in the B-CRM. In most cases, the better results were obtained when the uniform prior was taken with the power and hyperbolic model structures. The worst performances were produced by the logit model structure and log-normal prior in almost all scenarios. In conclusion, it is clear that B-CRM was found to be the most robust method among the dose finding methods in phase I trials. Lastly, it is important to note that if a power or hyperbolic tangent model structure and uniform prior are selected, the MTD selection rates appear to be at their highest rate in B-CRM.

References

Babb, J., Rogatko, A. & Zacks, S. (1998). Cancer phase I clinical trials: Efficient dose escalation with overdose control. Statistics in medicine, **17**(10): 1103-1120.

Cheung, Y.K. (2011). Dose finding by the continual reassessment method. Boca Raton, FL: Chapman & Hall. Pp.178.

Garrett-Mayer, E. (2006). The continual reassessment method for dose-finding studies: A tutorial. Clinical Trials, **3**(1): 57-71.

Hansen, A.R., Graham, D.M., Pond, G.R. & Siu, L.L. (2014). Phase 1 trial design: Is 3+ 3 the best? Cancer Control: **21**(3): 200-208.

Heyd, J.M. & Carlin, B.P. (1999). Adaptive design improvements in the continual reassessment method for phase I studies. Statistics in Medicine, **18**(11): 1307-1321.

Iasonos, A. & O'Quigley, J. (2014). Adaptive dose-finding studies: A review of model-guided phase I clinical trials. Journal of Clinical Oncology, **32**(23): 2505-2511.

Ishizuka, N. & Ohashi, Y. (2001). The continual reassessment method and its applications: A Bayesian methodology for phase I cancer clinical trials. Statistics in Medicine, **20**(17-18): 2661-2681.

Ji, Y., Feng, L., Liu, P., Shpall, E.J., Kebriaei, P., Champlin, R. & Cooper, L.J. (2012). Bayesian continual reassessment method for dose-finding trials infusing T-cells with limited sample size. Journal of Biopharmaceutical Statistics, **22**(6): 1206-1219.

Le Tourneau, C., Lee, J.J. & Siu, L.L. (2009). Dose escalation methods in phase I cancer clinical trials. JNCI: Journal of the National Cancer Institute, **101**(10): 708-720.

Leung, D.H.Y. & Wang, Y.G. (2002). An extension of the continual reassessment method using decision theory. Statistics in Medicine, **21**(1):51-63.

Nie, L., Rubin, E.H., Mehrotra, N., Pinheiro, J., Fernandes, L.L., Roy, A. & Alwis, D.P. (2016). Rendering the 3+3 design to rest: More efficient approaches to oncology dose-finding trials in the era of targeted therapy. Clinical Cancer Research: An Official Journal of the American Association for Cancer Research, **22**(11): 2623-2629.

O'Quigley, J., Pepe, M. & Fisher, L. (1990). Continual reassessment method: A practical design for phase I clinical trials in cancer. Biometrics, **46**(1):33-48.

O'Quigley, J. & Chevret, S. (1991). Methods for Dose Finding Studies in Cancer Clinical Trials: A Review and Results of a Monte Carlo Study. Statistics in Medicine, **10**(11), 1647-1664.

O'Quigley, J. & Shen, L.Z. (1996). Continual reassessment method: A likelihood approach. Biometrics, **52**(2):673-684.

O'Quigley, J. & Reiner, E. (1998). Stopping rule for the continual reassessment method. Biometrika, **85**(3): 741-748.

Storer, B.E. (1989). Design and analysis of phase I clinical trials. Biometrics, **45**(3): 925-937.

Sweeting, M., Adrian, M. & Tony, S. (2013). "BCRM: Bayesian continual reassessment method designs for phase I dose-finding trials. Journal of Statistical Software, **54**(13):1-26.

Ulas, E., & Karaman, F. (2018). A comprehensive comparison of phase I dose escalation methods. Kuwait Journal of Science, **45**(4): 26-36.

Yin, G. & Yuan, Y. (2009). Bayesian model averaging continual reassessment method in phase I clinical trials. Journal of the American Statistical Association, **104**(487):954-968.

Yin, G. (2012). Clinical trial design: Bayesian and frequents adaptive methods. John Wiley & Sons, Inc. Pp.330.

Submitted : 08/05/2017

Revised : 13/05/2018

Accepted : 17/09/2018

دراسة طرق البحث عن الجرعات في الأدوية وهياكل نموذجية مختلفة في طريقة إعادة التقييم المستمر البيزي

[1]طوبا كوك، [2]محمد جنكيز

[1]جامعة كانكيري كاراتكين، قسم الإحصاء، كانكيري، تركيا

[2]جامعة أوندوكوز مايز، قسم الإحصاء، سامسون، تركيا

الملخص

إن التجارب السريرية عبارة عن دراسات تستكشف ما إذا كان أي علاج أو دواء أو جهاز ما آمناً وفعالاً للاستخدامات البشرية أم لا. وتستطيع هذه الدراسات أن تبين الطريقة الوسيطة التي تعمل بشكل أفضل في بعض الأمراض. حيث أن التجربة السريرية التي يتم تصميمها بشكل جيد وتحليلها بصورة صحيحة تعتبر أداة قوية لتطوير عقاقير جديدة. الخطوة الأولى في اكتشاف الأدوية (المرحلة الأولى) مهمة للغاية لتحديد الجرعة القصوى المسموح بها (MTD).

في الجزء الأول من الدراسة، تم مقارنة التصميم الكلاسيكي 3 + 3، وطريقة إعادة التقييم المستمر (CRM)، وطريقة إعادة التقييم المستمر البيزي (B–CRM) من حيث احتمال اختيار MTD وعدد المرضى المُعالجين. من بين هذه التصاميم، قدمت طريقة B–CRM نتائج أفضل من الطريقتين الأخرتين. وفي الجزء الثاني من الدراسة، درسنا هياكل نموذجية مختلفة وتوزيعات سابقة في تصميم B–CRM. حيث نظرنا في ثلاثة هياكل للنماذج؛ هم: نموذج الطاقة، نموذج الظل الزائدي، ونموذج لوجيت وثلاثة توزيعات مختلفة؛ هي: جاما، والتوزيع المنتظم وتوزيعات طبيعية اللوغاريتم على التوالي. ووجدنا أنه إذا تم اختيار هيكل نموذج الطاقة أو نموذج الظل الزائدي والتوزيع المنتظم، فإن معدلات اختيار MTD كانت الأعلى في B–CRM.