Contemporary Clinical Trials Communications 7 (2017) 11–17

Contents lists available at ScienceDirect

Contemporary Clinical Trials Communications

journal homepage: www.elsevier.com/locate/conctc

A placebo-controlled Bayesian dose finding design based on continuous reassessment method with application to stroke research

Chunyan Cai[a,b,*], Mohammad H. Rahbar[a,b,c], Md Monir Hossain[d], Ying Yuan[e], Nicole R. Gonzales[f]

[a] Division of Clinical and Translational Sciences, Department of Internal Medicine, McGovern Medical School, The University of Texas Health Science Center at Houston, Houston, TX 77030, USA
[b] Biostatistics/Epidemiology/Research Design Component, Center for Clinical and Translational Sciences, The University of Texas Health Science Center at Houston, Houston, TX 77030, USA
[c] Division of Epidemiology, Human Genetics, and Environmental Sciences, University of Texas School of Public Health at Houston, Houston, TX 77030, USA
[d] Division of Biostatistics and Epidemiology, Cincinnati Children's Hospital Medical Center, Cincinnati, OH 45229, USA
[e] Department of Biostatistics, Division of Quantitative Sciences, The University of Texas MD Anderson, Houston, TX 77030, USA
[f] Department of Neurology, McGovern Medical School, The University of Texas Health Science Center at Houston, Houston, TX 77030, USA

ARTICLE INFO

Keywords:
Continuous reassessment method
Dose finding
Placebo-controlled study
Stroke research

ABSTRACT

Traditional dose-finding designs do not require assignment of patients to a control group. Motivated by SHRINC (Safety of Pioglitazone for hematoma resolution in intracerebral hemorrhage), we developed a placebo-controlled dose-finding study to identify the maximum tolerated dose for pioglitazone in stroke patients with spontaneous intracerebral hemorrhage. We designed an extension of the continuous reassessment method that allowed to incorporate information from the control group (i.e., the standard of care), and utilized it to determine the maximum tolerated dose in the SHRINC trial. We evaluated the operating characteristics of our design by conducting extensive simulation studies. Our findings from the simulation studies demonstrate that our proposed design is robust and performs well. By estimating the toxicity rate in the control group, we were able to obtain more accurate information about the natural history of the disease and identify appropriate dose for the next phase of this study. The proposed design provides a tool to incorporate the information from the control group into the dose-finding framework for trials with similar objectives.

1. Introduction

Clinical trials are considered as the most reliable method for evaluation of safety and efficacy of new drugs and other clinical interventions. Phase I dose-finding trials are designed to identify the maximum tolerated dose (MTD) of a new drug, defined as the highest dose within a tolerable dose-limiting toxicity (DLT). After the MTD is determined, the drug will be carried forward for subsequent assessments through phase II and III trials. Inaccurate evaluation of the MTD can lead to waste of resources if an inappropriate dose is moved to subsequent phases. Therefore, it is important to design efficient dose-finding trials to determine the most appropriate dose before the drug is tested in future phases of drug development.

Broadly, dose-finding trial designs are classified into two types, algorithm-based designs, and model-based designs. The algorithm-based designs, also known as up-and-down designs, are used often in practice due to their simplicity in implementation. The most popular algorithm-based design is the "3 + 3" design [1]. Although it is popular in practice, the reported shortcomings of this design include unreliable estimation of the MTD [2], a significant large proportion of patients treated at subtherapeutic dose levels [3], and the restricted choice of the target DLT rate [4]. Several investigators attempted to develop improved up-and-down designs to identify the MTD, including the accelerated titration design [5], the biased coin design [6] and its extension with isotonic regression [7], the k-in-a-row design [8], the up-and-down design based on isotonic regression [9], the modified toxicity probability interval design [10], and the Bayesian optimal interval design [11]. Comprehensive reviews of up-and-down designs are provided by Ivanova [12] and Liu et al. [4]. For model-based designs, the most popular one is the continual reassessment method (CRM) [13]. In contrast to 3 + 3 design, CRM design provides a more accurate estimation of toxicity probability of the MTD and a more flexible setup of target DLT rate. Due to the popularity of CRM, a variety of extensions have been proposed to improve its practical implementation and

* Corresponding author. Division of Clinical and Translational Sciences, Department of Internal Medicine, McGovern Medical School, The University of Texas Health Science Center at Houston, Houston, TX 77030, USA.
E-mail address: chunyan.cai@uth.tmc.edu (C. Cai).

https://doi.org/10.1016/j.conctc.2017.05.002
Received 20 September 2016; Received in revised form 6 April 2017; Accepted 3 May 2017
Available online 06 May 2017

operating characteristics [14–24]. Other model-based designs include designs based on Bayesian decision-theoretic approach [25], the escalation with overdose control [26], Bayesian nonparametric approach [27], and stochastic approximation [28]. Comprehensive reviews for dose-finding designs can be found in Rosenberger and Haines [29], Chevret [30], Ting [31], and Le Tourneau et al. [3].

In stroke trials, DLT usually is defined as a clinical event with substantial morbidity and mortality such as a cerebral hemorrhage [32]. Therefore, the acceptable target rate of patients undergoing such events is quite small. With the development of aforementioned dose-finding designs, CRM is the most promising detection method to address this challenge. There has been some effective utilization of dose-finding designs in stroke trials, especially the application of CRM [33–35]. In addition to the utilization of CRM design, other dose-finding designs were also applied to stroke trials. For example, Krams et al. [36] used a Bayesian adaptive dose finding design based on normal dynamic linear model in an acute ischemic stroke study. Whelan et al. [37] described the utilization of the Bayesian phase I/II design proposed by Thall and Cook [2] to find an optimal dose for treatment of ischemic stroke in children.

All the aforementioned dose-finding designs do not require assignment of patients to a control group. However, for an acute intracerebral hemorrhage (ICH) trial, SHRINC (Safety of Pioglitazone for hematoma resolution in intracerebral hemorrhage), the investigators were interested in finding a dose with the target toxicity rate dependent upon the rate in the concurrent control group. That means, the target toxicity rate in the SHRINC study is unknown before the initiation of the trial and needs to be determined based on the toxicity data collected from the concurrent control group. Motivated by SHRINC, we developed a placebo-controlled dose-finding study based on CRM to identify the MTD. The inclusion of a control group allowed us to study the natural history of the disease and co-morbidities. Therefore, our proposed design can lead to more meaningful MTD identification than traditional designs. The purpose of this paper is to describe the unique features of our proposed design, and share our experience on its application to the SHRINC study. Using a stroke study as a motivating example, our proposed design provides a tool to incorporate the information from the control group into the dose-finding framework for trials in other diseases with similar objectives. We evaluate the operating characteristics of the proposed design by conducting extensive simulation studies.

2. Methods

2.1. Rational of the inclusion of a control group

SHRINC study was designed to assess the safety of PIO in spontaneous ICH compared with the standard of care. It is a prospective, randomized, blinded, placebo-controlled, and dose-escalation safety trial in which patients are randomly allocated to control or treatment groups. The primary objective of the SHRINC trial is to determine the MTD of PIO, a dose with DLT rate closest to a target rate. More detailed information including inclusion/exclusion criteria, informed consent, safety outcomes, and clinical and radiographic outcomes can be found in Gonzales et al. [38]. The SHRINC study is registered at http://www.clinicaltrials.gov/(registration number NCT00827892).

The inclusion of a control group is not a characteristic of the traditional dose finding trials. In SHRINC, we included a control group for the following reasons: first, many expected adverse events (AEs) and serious AEs (SAEs) in the PIO group were also expected in the ICH population as part of the natural history of the disease and co-morbidities. These AEs and SAEs were not captured in our prospective stroke registry [39] and are difficult to collect accurately in a retrospective manner. Therefore, we planned to collect this information prospectively and compared the rates of AE/SAEs between the PIO-treated and control groups to obtain a more accurate measure of safety with the use of PIO in our patient population. Second, in our preclinical work, more rapid hematoma resolution was correlated with improved neurologic recovery. This finding has not been demonstrated in the clinical setting. In addition, the control group data related to the rate of hematoma resolution would help to determine an optimal duration of PIO. Third, retrospective studies usually overestimate the expected benefit of treatment. Therefore, the inclusion of a concurrent control group would improve our knowledge of toxicity and efficacy profile of the study population and guide us to choose a meaningful target DLT for MTD identification. Thus, instead of specifying a fixed target DLT rate before the initiation of the trial, we consider the target rate dependent up on the rate in the concurrent control group. As a result, many existing methods, such as the CRM, cannot be directly applied. Toward this goal, we propose a modified CRM dose-finding design with a control group for our study.

2.2. Dose-finding method

In this section, we propose an extension of the CRM that allows incorporating information from the control group for dose finding. Suppose K dose levels, denoted as $d_1, d_2, \ldots, d_K$, have been chosen for the investigation with the true toxicity probability P_k for dose d_k, $k = 1, \ldots, K$. Let Y be a binary variable to denote whether a patient has experienced the prespecified DLT event, with 1 denoting an event and 0 otherwise, and let x denote the dose level for this patient. Usually, a one-parameter model $\psi(x,\alpha)$ is proposed to model the relationship between dose level x and its toxicity probability $P(Y = 1|x)$ with the unknown parameter α. There are three popular dose-toxicity models of $\psi(x,\alpha)$ proposed for CRM: logistic model, power model, and hyperbolic tangent model. In our proposed design, we utilize the following one-parameter logistic model,

$$P(Y = 1|x) = \psi(x,\ \alpha) = \frac{\exp(c + \alpha x)}{1 + \exp(c + \alpha x)}, \tag{1}$$

where c is a constant and recommended to be 3 [40]. As part of the design, we also need a "skeleton" for the CRM, which is the investigator's prior estimates of DLT at each dose level, denoted as P_k^0 for dose level d_k. By plugging the toxicity probability in equation (1) with each value of the "skeleton", we obtain the standardized dose $d_k' = \frac{logit(P_k^0) - c}{\hat{\alpha}}$, $k = 1, \ldots, K$, where $\hat{\alpha}$ is the mean or median of the prior estimate of the parameter and $logit(p) = \log\left(\frac{p}{1-p}\right)$.

During the course of the trial, if n patients have been enrolled into the study and assigned to different dose levels, we denote the standardized dose for ith patient, $i = 1, \ldots, n$ as x_i with his/her observed outcome y_i, where $x_i \in \{d_1', d_2', \ldots, d_K'\}$. Denoting the observed data $\{x_i, y_i, i = 1, \ldots, n\}$ as D, the likelihood function of observed data D can be written as $L(D|\alpha) \propto \prod_{i=1}^{n} \psi(x_i, \alpha)^{y_i}(1 - \psi(x_i, \alpha))^{1-y_i}$. Denoting the prior density of α as $\pi(\alpha)$, the posterior density of α can be written as $f(\alpha|D) \propto L(D|\alpha)\pi(\alpha)$. The prior distribution for α is set to unit exponential distribution. The other choice can be uniform distribution with $0 < \alpha < 3$, and the log-normal distribution. The Gibbs sampler [41] can be used to obtain the posterior samples for α. After that, the posterior mean for P_i at dose d_i, $i = 1, \ldots, K$, can be calculated as $\hat{P}_i = \int \psi(d_i',\ \alpha) f(\alpha|D) d\alpha$.

In the traditional CRM, if the target DLT rate is set at a fixed value of θ, then the MTD d^* is taken to be one of the specified dose in the set of $d_1, d_2, \ldots, d_k$ which satisfies the following criteria $d^* = argmin|\hat{P}_i - \theta|$, $i \in 1, \ldots, k$. Without losing the generality of assumption, we consider a design to identify the MTD with the DLT rate closest to a target rate, defined as a rate higher than the rate in the control group by a constant magnitude of δ. We denote the toxicity rate in the control group as P_0. To reduce the variability of the toxicity rate in the control group among different cohorts, we estimate P_0 during the course of the trial using all enrolled controls so far. Suppose we observe m_0 toxicities among n_0 patients in the control group and assume P_0 follows a Beta distribution $Beta$(a, b) with two shape parameters a and

b. Under the beta-binomial model, the posterior mean of P_0 is estimated as $\hat{P}_0 = \frac{m_0 + a}{n_0 + a + b}$. Then the modified MTD d^* in our proposed design is identified as the dose which satisfies the following modified criteria

$$d^* = argmin|\hat{P}_i - \hat{P}_0 - \delta|, \quad i \in 1, \ ..., \ k. \tag{2}$$

2.3. Dose escalation/de-escalation rule

We assume patients are treated in cohorts at each dose level and the concurrent control group. To ensure patient safety, we do not allow skipping more than one dose level for dose escalation, but for de-escalation, we allow more than one skip in dose levels. Our proposed dose-finding algorithm is described as follows:

1. Start the trial at the lowest dose level and randomize a cohort of patients to the control and treatment groups based on a 1:1 allocation ratio.
2. After obtaining the toxicity outcome of the enrolled patients, update the dose-toxicity relationship for the treatment group using model (1), estimate the toxicity rate of the control group using all enrolled cohorts in the control group and identify the MTD based on the criteria (2).
   - 2.1. If the estimated MTD is lower than the current dose level, de-escalate the dose to the estimated MTD.
   - 2.2. If the estimated MTD is higher than the current dose level, escalate the dose to the estimated MTD if no more than one dose level was skipped; Otherwise, escalate the dose to the next higher dose level.
   - 2.3. If the estimated MTD is the current dose level, assign the next cohort to the same dose level.
3. Repeat the above steps until reaching the sample size N and select as the MTD the dose satisfying the criteria (2).

3. Simulation studies

Simulation studies have been conducted to evaluate the performance of our proposed design. To mimic our SHRINC study, we consider a sample size of 84 and 11 dose levels for investigation. The investigator's prior estimates of DLT at each dose level are 10%, 12%, 15%, 18%, 21%, 25%, 26%, 27%, 28%, 29%, 30%. We set the target DLT rate the same as the rate in the control group, i.e., $\delta = 0$. In addition, we consider the cases with $\delta = 0.10$. The logistic model (1) is utilized to estimate the dose-toxicity relationship with unit exponential prior for α. To estimate the toxicity rate in the control group, i.e., P_0, we assume it follows a weakly informative prior $Beta(0.1, 0.6)$ with mean at 0.14.

We consider ten scenarios to evaluate the performance under various dose-toxicity profiles and different values of δ. We display the true toxicity rates and the value of δ under each scenario in Table 1. In scenario 1, we consider the toxicity rate of the control group P_0 as 0.10 and the toxicity rates of 11 PIO dose levels ranging from 0.01 to 0.46. Since the MTD is defined as a dose with a toxicity rate the same as P_0, i.e., $\delta = 0$, the target MTD in scenario 1 is the third dose level, i.e., 0.4 mg/kg/day. We consider P_0 as 0.20 in scenario 2 and 0.30 in scenario 3, respectively. In scenarios 4 to 6, we consider $\delta = 0.10$. In scenarios 7 to 10, we consider the scenarios with the target toxicity rate of MTD exactly equal to 15% and 25% under different choices of P_0 and δ. The target MTDs are presented in boldface in Table 1 under each dose-toxicity profile.

We evaluate and compare the performance of our proposed study design with the traditional CRM and 3 + 3 designs. For the traditional CRM design, the MTD is usually defined as a dose with a prespecified target toxicity rate. However, in the SHRINC study, the target toxicity rate of the MTD depends on P_0. Without knowing P_0, we encounter a challenge to select an appropriate target rate when the traditional CRM design is applied. Therefore, we consider two choices of the target toxicity rate for the traditional CRM design: 1) set the target rate at 15%, denoted as $CRM_{15\%}$; 2) set the target toxicity rate at 25%, denoted as $CRM_{25\%}$. Second, our proposed design includes a control group and would assign half of patients (i.e., 42 patients) to the control group, but the traditional CRM design does not require the control group. For a fair comparison, we consider two different sample sizes, i.e., 42 and 84, for the traditional CRM design.

Simulation results are displayed in Table 2, which compares the correct selection percentage of the MTD and the percentage of patients assigned to over-toxic doses, i.e., dose levels above the MTD, among the proposed, the traditional CRM, and 3 + 3 designs. In scenario 1, our proposed design successfully selects the MTD (i.e, the 3rd dose) with the percentage of 24.6%, which is higher than that using the traditional CRM and the 3 + 3 designs. Although the selection percentage of the MTD using $CRM_{15\%}$ is close to that of our proposed design, it assigns almost half the number of patients to over-toxic doses. $CRM_{25\%}$ has the worst performance and incorrectly considers a higher dose as MTD most of the time. In scenario 2, our proposed design has similar performance as $CRM_{15\%}$ with sample size of 42, but performs worse than $CRM_{25\%}$ and the 3 + 3 designs. This is reasonable because the target toxicity rates of $CRM_{25\%}$ and the 3 + 3 design are both close to the true toxicity rate of the MTD. We observe similar findings in scenario 3, in which the proposed design outperforms $CRM_{15\%}$ and 3 + 3 designs, but not $CRM_{25\%}$. In scenarios 4 to 6, we consider $\delta = 0.10$, i.e., the MTD is defined as a dose with a toxicity rate higher than P_0 by a constant magnitude of 0.10. In scenario 4, our proposed design has a higher percentage of selecting MTD than that using $CRM_{15\%}$ and the 3 + 3 designs. $CRM_{25\%}$ with sample size of 84 selects the MTD with a higher percentage than our proposed design, but it assigns more than half number of patients to over-toxic doses. In scenario 5, our proposed

Table 1
True DLT rates at eleven PIO dose levels as well as the toxicity rate in the control group under six scenarios (the target MTD are in boldface).

Scenario	δ	Toxicity rate in control	True DLT rate										
			PIO dose level (mg/kg/day)										
			0.1	0.2	0.4	0.6	0.8	1.0	1.2	1.4	1.6	1.8	2.0
1	0	0.10	0.01	0.04	**0.09**	0.15	0.20	0.28	0.33	0.37	0.39	0.43	0.46
2	0	0.20	0.06	0.10	0.14	**0.20**	0.28	0.36	0.42	0.46	0.50	0.53	0.58
3	0	0.30	0.14	0.21	**0.28**	0.35	0.42	0.48	0.52	0.59	0.62	0.65	0.68
4	0.10	0.10	0.01	0.04	0.09	0.15	**0.20**	0.28	0.33	0.37	0.39	0.43	0.46
5	0.10	0.20	0.06	0.10	0.14	0.20	**0.28**	0.36	0.42	0.46	0.50	0.53	0.58
6	0.10	0.30	0.14	0.21	0.28	0.35	**0.42**	0.48	0.52	0.59	0.62	0.65	0.68
7	0	0.15	0.01	0.04	0.09	**0.15**	0.22	0.29	0.35	0.39	0.43	0.46	0.49
8	0	0.25	0.03	0.08	0.13	0.18	**0.25**	0.31	0.36	0.42	0.46	0.50	0.53
9	0.10	0.05	0.01	0.04	0.09	**0.15**	0.20	0.28	0.33	0.37	0.39	0.43	0.46
10	0.10	0.15	0.06	0.10	0.14	0.19	**0.25**	0.33	0.38	0.44	0.50	0.53	0.58

Table 2
Percentage of the correct identification of MTD (Average percentage of patients assigned to the doses above the MTD) based on the proposed, traditional CRM, and 3 + 3 designs under different scenarios based on 1000 simulations.

Scenario	Proposed[c] Design	$CRM_{15\%}$[a]		$CRM_{25\%}$[b]		3 + 3 Design
		N = 42[d]	N = 84[e]	N = 42[d]	N = 84[e]	
1	24.6 (18.3)	21.6 (54.6)	23.4 (62.8)	2.5 (75.7)	0.5 (87.3)	17.6 (49.3)
2	22.9 (17.6)	22.9 (14.7)	26.1 (10.9)	26.3 (44.9)	33.4 (51.4)	24.4 (17.9)
3	22.0 (21.4)	6.4 (5.2)	2.8 (3.0)	33.3 (24.8)	43.6 (21.7)	17.7 (10.6)
4	29.8 (17.2)	24.3 (17.1)	26.8 (11.9)	26.6 (45.5)	38.3 (51.6)	24.5 (16.5)
5	28.9 (20.2)	8.4 (5.3)	4.1 (3.1)	37.7 (24.6)	46.5 (21.1)	16.3 (7.1)
6	23.7 (17.6)	0.2 (0.5)	0 (0.4)	5.6 (4.0)	1.3 (2.3)	2.1 (0.7)
7	22.6 (17.6)	34.4 (32.0)	48.3 (30.1)	11.2 (61.6)	8.8 (73.9)	25.7 (29.9)
8	26.7 (16.8)	14.1 (9.3)	9.2 (5.3)	33.8 (33.4)	47.3 (33.3)	18.3 (10.3)
9	26.0 (18.4)	34.3 (33.9)	46.8 (33.0)	7.8 (63.0)	5.4 (76.5)	23.8 (31.0)
10	27.7 (14.5)	10.6 (7.2)	6.5 (4.2)	35.0 (28.8)	48.4 (27.8)	17.6 (8.7)

[a] Traditional CRM design with the target toxicity limit at 15%.
[b] Traditional CRM design with the target toxicity limit at 25%.
[c] We considered a sample size of 84 patients and randomized half patients to control group in the proposed design.
[d] We considered a sample size of 42 patients for the traditional CRM design.
[e] We considered a sample size of 84 patients for the traditional CRM design.

design outperforms $CRM_{15\%}$ and the 3 + 3 designs, but worse than $CRM_{25\%}$ design of which the chosen target rate is close to the true toxicity rate of the MTD. In scenario 6, only our proposed design is able to identify the MTD and selects it with the highest percentage among all designs. In scenarios 7 and 9, the toxicity rate of the MTD is 15%, which exactly equals to the prespecified target rate in $CRM_{15\%}$ design. Therefore, the $CRM_{15\%}$ design performs best among all designs. Our proposed design has a lower correct selection percentage than the $CRM_{15\%}$ design, but performs better than the $CRM_{25\%}$ design. Similarly, in scenarios 8 and 10, with the toxicity rate of the MTD at 25%, the $CRM_{25\%}$ design has the best performance, and the $CRM_{15\%}$ design has the worst performance. These results show that arbitrarily choosing a target rate for the CRM design may not be appropriate especially when the selected rate deviates from the real value a lot, resulting in an incorrect identification of MTD. Overall, based on the findings from these scenarios, if the target rate of the MTD depends on an unknown P_0, we need to estimate it and incorporate it into the dosing finding procedure. Otherwise, we may have high chance to miss the MTD. The 3 + 3 design works well if the toxicity rate of the MTD is close to 20–25%, e.g., scenarios 2 and 4. Similarly, the traditional CRM design also has a good performance if the prespecified rate is close to the toxicity rate of the MTD. In summary, our proposed design is robust and works well in most of the scenarios when the toxicity rate in the control group needs to be estimated and incorporated into the MTD finding procedure. To evaluate the robustness of the proposed design, we conduct a sensitivity analysis to evaluate its performance with a different choice of the prior specification for the parameter α. We consider a log-normal distribution with mean of 0 and variance of 1 for α and display the simulation results under the ten scenarios in Table 3. We compare these results to that in Table 2, in which the proposed design specifies a unit exponential prior for the parameter α. The percentages of the correct identification of MTD are similar for each scenario with these two priors. These results demonstrate that the performance of our design is robust on the specification of the prior distribution for the parameter α as long as it covers a reasonable range of the parameter estimated. In addition, we conduct a sensitivity analysis with a different choice of the allocation ratios between the treatment and the control groups. Different from the allocation ratio 1:1 used in Table 2, we consider the allocation ratios of 3:2 and 3:1, and summarize these results in Table 3. The selection percentage of the correct MTD identification with 3:2 ratio is slightly lower than the results with a ratio at 1:1 (see Table 2), but it utilizes a smaller sample size. However, the ratio 3:1 may not be a good choice in which the correct selection percentage is much lower than that with a ratio at 1:1. Therefore, we suggest to evaluate the performance of the proposed design with different choices of allocation ratio to select a one which provides a desirable performance with a reasonable sample size.

4. Application of the proposed design to the SHRINC study

We implemented the proposed placebo-controlled CRM design for the MTD determination in the SHRINC study. There were 11 dose levels proposed in the study, which are 0.1, 0.2, 0.4, 0.6, 0.8, 1.0, 1.2, 1.4, 1.6, 1.8, 2.0 mg/kg/day. A total of 84 patients were recruited and randomized to two study arms. Patients were enrolled in a cohort of 6 with 3 patients randomized to the treatment group and 3 patients randomized to the control group. Patients assigned to the treatment group were administered the escalated dose for 3 days followed by a daily maintenance dose of 30 mg. DLT was defined as in-hospital death (discharge or day 14 whichever occurs first) and the target rate was defined as a rate the same as the toxicity rate in the control group. That means, if the toxicity rate in control group is 15%, then the target rate for the MTD identification is set at 15%, i.e., $\delta = 0$ in criteria (2).

Our trial successfully enrolled 84 patients with 42 assigned to the placebo group and 42 assigned to the PIO group. Table 4 shows the number of patients and the observed in-hospital deaths at each dose level. At the end of the trial, we observed two in-hospital deaths, with one in the control group and the other in the PIO group. The course of

Table 3
Sensitivity analysis with a lognormal prior distribution for alpha and different allocation ratios between treatment and control groups based on the proposed design under ten scenarios.

Setting	Scenario									
	1	2	3	4	5	6	7	8	9	10
Lognormal prior for α	23.4 (15.6)	22.0 (15.8)	23.5 (18.6)	26.3 (15.2)	29.3 (17.9)	21.9 (15.3)	22.5 (17.2)	25.6 (16.3)	26.2 (17.7)	28.7 (13.8)
Ratio 3:2	21.8 (15.6)	19.8 (16.0)	21.7 (19.3)	29.3 (15.3)	25.3 (18.8)	23.0 (16.1)	19.6 (17.4)	23.2 (17.1)	27.7 (18.5)	26.4 (18.2)
Ratio 3:1	13.5 (13.8)	15.9 (14.8)	16.1 (18.5)	20.4 (14.9)	20.9 (17.1)	17.1 (15.6)	16.8 (15.3)	16.7 (15.6)	25.1 (17.9)	20.3 (21.4)

Data are represented as the percentage of the correct identification of MTD (average percentage of patients assigned to the doses above the MTD).

Table 4
Number of patients and the observed deaths at different PIO dose levels.

Pioglitazone (PIO)		
PIO daily dose for three days (mg/kg/day)	# of treated patients	# of deaths at discharge or day 14
0.1	4	0
0.2	0	0
0.4	2	0
0.6	3	0
0.8	3	0
1.0	4	0
1.2	3	0
1.4	3	0
1.5	3	0
1.6	3	0
1.7	6	0
1.8	6	1
1.9	3	0
2.0	0	0

patient flow including the dose escalation and de-escalation is shown in Fig. 1. Three patients out of the first cohort were randomized to the starting dose, 0.1 mg/kg/day. Subsequent cohorts were randomized to doses after obtaining in-hospital data of all previous cohorts. The doses for cohorts 2 to 7 were escalated as follows: 0.4, 0.6, 0.8, 1.0, 1.2, and 1.4 mg/kg/day. On the completion of cohort 7, we observed one death in the control group. The trial was continued with dose escalation for cohorts 8, 9, 10, and 11. However, we observed one death at cohort 11 in the PIO group. Subsequently, the dose levels for the next two cohorts were de-escalated to 1.7 and 1.6 mg/kg/day for patient safety. The dose level was then re-escalated to 1.7 mg/kg/day for the last cohort. After reaching the planned sample size, the study was completed. According to the proposed dose-finding algorithm, we updated the dose-mortality curve (Fig. 2) and estimated the in-hospital mortality rate for each dose level. The posterior mean of the mortality rate at the highest dose level 2.0 mg/kg/day is 4.1% and the posterior mean of the mortality rate in the control group is 2.6%. The dose level 1.0 mg/kg/day is the one that satisfies the proposed MTD criteria defined in (2), which is the dose with the mortality rate closest to that in the control group.

5. Discussion

In this manuscript, we proposed a placebo-controlled dose-finding design with an application to a stroke study and shared our experiences at various faces of its designing and implementations. The motivating study was not only a dose-finding study for PIO treating ICH patients, but also studying the natural history of the disease and identifying whether the speed of hematoma/edema resolution in ICH represents a biological marker actively correlated with radiographic and clinical effect of PIO. Considering the overall objectives of the study as well as the MTD identification, there was a need for having a control group and leveraging this additional information available to us by incorporating it into dose-finding procedure. The proposed design provides a tool to incorporate the control data into the dose-finding framework for studies with similar objectives, and can be applied to other diseases other than stroke. Through simulation studies, we evaluated the performance of our proposed design by comparing it with the traditional CRM and 3 + 3 designs. Our simulation results demonstrate that our proposed design is robust and works well when the toxicities of the treatment can overlap with the natural history of the studied disease. Therefore, we suggest including a concurrent control group for more meaningful MTD identification under this situation. The data from the control group contain important information to study the natural history of the disease, such as the incidence of AE/SAEs and the radiographic information. In addition, the two-week mortality rate in our concurrent control group is different from the one we estimated from the historical data using our registry data, which also demonstrates the importance of the inclusion of concurrent control group. These differences between the concurrent and historical controls could be due to the possible guidelines or standard of care changes.

A number of lessons we have learned from our SHRINC study. First, we have enrolled the planned 84 patients and observed two DLTs. With

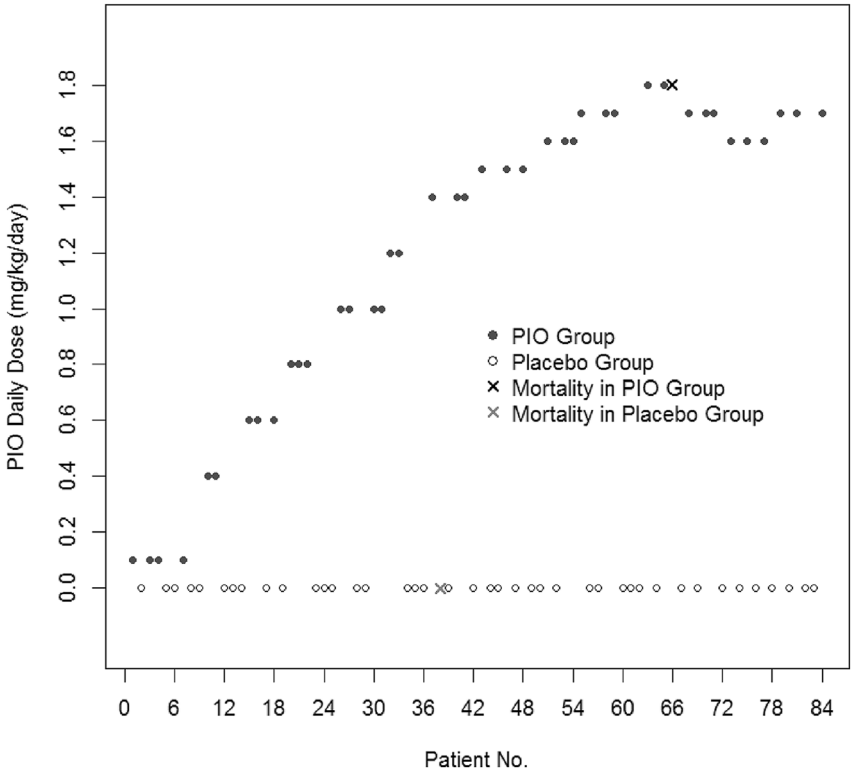

Fig. 1. Does escalation/de-escalation throughout the course of the trial for the PIO and placebo groups. Patients are presented in cohort order from left to right. Open circles indicate patients in the PIO group and solid grey circles indicate patients in the placebo group. Patients who experienced DLT, i.e., mortality within two weeks or before discharge, are denoted as crosses.

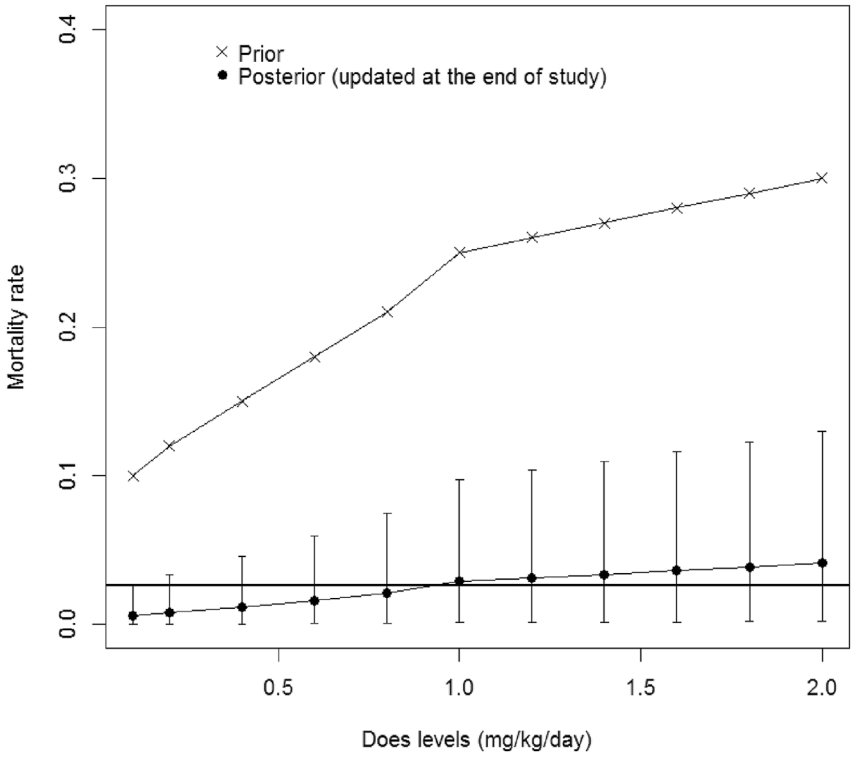

Fig. 2. Model-based probabilities of the mortality. Crosses represent a priori assumptions about the defined mortality from the investigator, i.e., "Skeleton". Filled circles and bars represent the posterior means as well as the 95% credible intervals of the mortality rates at each dose level based on the data collected at the end of the study. The horizontal line indicates the estimated mortality rate in the concurrent control group.

11 dose levels proposed in our study, nearly 1 or 2 cohorts were assigned to each dose. After the interim analysis, due to safety reasons, three more doses were inserted into the prespecified 11 dose levels. Although this provides an estimation of dose-toxicity relationship under different dose levels, it limits the number of cohorts we could assign to the target MTD or its neighborhood. Therefore, for designing a similar dose-finding study in future, we suggest investigating fewer dose levels or starting the trial from the dose in the middle of dose range. For appropriate dose insertion, simulation studies are recommended to evaluate the performance before implementing it. Second, the DLT in our SHRINC study is defined as a death before discharge or at two weeks which occurs first. Therefore, we need to wait for several days or even two weeks to observe outcomes from all previous cohorts. During the course of our trial, if an eligible patient arrived before the outcomes were observed for the current cohort, we allowed enrollment of this patient into the study and randomized this patient to either the current dose or control group, e.g., 4 patients assigned to 0.1 mg/kg/day. After this patient was enrolled, we did not randomize any new patients until observing outcome for all the patients enrolled to that point. However, it is possible that we still miss eligible patients. To avoid this, we can apply the methods described in Cai et al. [42] to impute the missing outcomes before decision making. Third, we realized the choice of in-hospital mortality as primary endpoint may not be optimal for this study after we finished the trial. The observed mortality rate is lower than our expected rate, which is not as informative as we expected for our dose escalation and dose finding procedure, but fortunately it demonstrates our proposed treatment is safe to patients. Our results show that the prespecified skeleton in our SHRINC study may deviate from the truth. To overcome the arbitrariness and further enhance the robustness of the design, one possible approach is to specify a different set of skeletons and adopt Bayesian model averaging method proposed in Yin and Yuan (2009) [20]. Further exploration of the optimal dose identification for the subsequent phase based on other information such as serious adverse events as well as considering both toxicity and efficacy data may be beneficial [43]. This may require a specific definition of the AEs and the determination whether the AE is part of the natural history of the disease or due to the treatment.

Declaration of conflicting interests

The authors declare that there is no conflict of interest.

Acknowledgements

This research is funded by the National Institute of Neurological Disorders and Stroke (NINDS) through a SPOTRIAS program grant P50NS044227 awarded to the University of Texas Health Science Center at Houston (UTHealth). We also acknowledge the support provided by the Biostatistics/Epidemiology/Research Design (BERD) component of the Center for Clinical and Translational Sciences (CCTS) for this project. CCTS is mainly funded by the NIH Centers for Translational Science Award (NIH CTSA) grant UL1 RR024148 (KL2 RR0224149 for the K12 program), awarded to University of Texas Health Science Center at Houston in 2006 by the National Center for Research Resources (NCRR) and its renewal (UL1 TR000371) by the National Center for Advancing Translational Sciences (NCATS). The content is solely the responsibility of the authors and does not necessarily represent the official views of the NINDS, NCRR, or the NCATS.

References

[1] B.E. Storer, Design and analysis of phase I clinical trials, Biometrics 45 (3) (1989 Sep) 925–937.
[2] P.F. Thall, S.J. Lee, Practical model-based dose-finding in phase I clinical trials: methods based on toxicity, Int. J. Gynecol. Cancer 13 (3) (2003 May) 251–261.
[3] C. Le Tourneau, J.J. Lee, L.L. Siu, Dose escalation methods in phase I cancer clinical trials, J. Natl. Cancer Inst. 101 (10) (2009 May 20) 708–720.
[4] S. Liu, C. Cai, J. Ning, Up-and-down designs for phase I clinical trials, Contemp. Clin. Trials 36 (1) (2013 Sep) 218–227.
[5] R. Simon, B. Freidlin, L. Rubinstein, S.G. Arbuck, J. Collins, M.C. Christian, Accelerated titration designs for phase I clinical trials in oncology, J. Natl. Cancer Inst. 89 (15) (1997 Aug 6) 1138–1147.
[6] S.D. Durham, N. Flournoy, Random Walks for Quantile Estimation. Statistical Decision Theory and Related Topics V, Springer, 1994, pp. 467–476.
[7] M. Stylianou, N. Flournoy, Dose finding using the biased coin up-and-down design and isotonic regression, Biometrics 58 (1) (2002 Mar) 171–177.
[8] M. Gezmu, The Geometric Up and Down Design for Allocating Dosage Levels, American University, Washington, DC, 1995.

[9] D.H. Leung, Y. Wang, Isotonic designs for phase I trials, Control Clin. Trials 22 (2) (2001 Apr) 126–138.
[10] Y. Ji, P. Liu, Y. Li, B.N. Bekele, A modified toxicity probability interval method for dose-finding trials, Clin. Trials 7 (6) (2010 Dec) 653–663.
[11] S. Liu, Y. Yuan, Bayesian optimal interval designs for phase I clinical trials, J. R. Stat. Soc. Ser. C Appl. Stat. 64 (3) (2015) 507–523.
[12] A. Ivanova, Dose Escalation and Up-and-Down Designs. Wiley Encyclopedia of Clinical Trials, John Wiley & Sons, Inc, 2007.
[13] J. O'Quigley, M. Pepe, L. Fisher, Continual reassessment method: a practical design for phase 1 clinical trials in cancer, Biometrics 46 (1) (1990 Mar) 33–48.
[14] S.N. Goodman, M.L. Zahurak, S. Piantadosi, Some practical improvements in the continual reassessment method for phase I studies, Stat. Med. 14 (11) (1995 Jun 15) 1149–1161.
[15] S. Moller, 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 (9–10) (1995 May 15) 911–922.
[16] S. Piantadosi, J.D. Fisher, S. Grossman, Practical implementation of a modified continual reassessment method for dose-finding trials, Cancer Chemother. Pharmacol. 41 (6) (1998) 429–436.
[17] J.M. Heyd, B.P. Carlin, Adaptive design improvements in the continual reassessment method for phase I studies, Stat. Med. 18 (11) (1999 Jun 15) 1307–1321.
[18] T.M. Braun, The bivariate continual reassessment method. extending the CRM to phase I trials of two competing outcomes, Control Clin. Trials 23 (3) (2002 Jun) 240–256.
[19] D.H. Leung, Y.G. Wang, An extension of the continual reassessment method using decision theory, Stat. Med. 21 (1) (2002 Jan 15) 51–63.
[20] G. Yin, Y. Yuan, Bayesian model averaging continual reassessment method in phase I clinical trials, J. Am. Stat. Assoc. 104 (487) (2009) 954–968.
[21] T. Daimon, S. Zohar, J. O'Quigley, Posterior maximization and averaging for Bayesian working model choice in the continual reassessment method, Stat. Med. 30 (13) (2011 Jun 15) 1563–1573.
[22] Y. Yuan, G. Yin, Robust EM continual reassessment method in oncology dose finding, J. Am. Stat. Assoc. 106 (495) (2011 Sep 1) 818–831.
[23] S. Liu, G. Yin, Y. Yuan, Bayesian data augmentation dose finding with continual reassessment method and delayed toxicity, Ann. Appl. Stat. 7 (4) (2013 Dec 1) 1837–2457.
[24] H. Pan, Y. Yuan, A default method to specify skeletons for Bayesian model averaging continual reassessment method for phase I clinical trials, Stat. Med. 36 (2) (2017) 266–279.
[25] J. Whitehead, H. Brunier, Bayesian decision procedures for dose determining experiments, Stat. Med. 14 (9–10) (1995 May 15) 885–893.
[26] J. Babb, A. Rogatko, S. Zacks, Cancer phase I clinical trials: efficient dose escalation with overdose control, Stat. Med. 17 (10) (1998 May 30) 1103–1120.
[27] M. Gasparini, J. Eisele, A curve-free method for phase I clinical trials, Biometrics 56 (2) (2000 Jun) 609–615.
[28] Y.K. Cheung, M.S. Elkind, Stochastic approximation with virtual observations for dose-finding on discrete levels, Biometrika 97 (1) (2010 Mar) 109–121.
[29] W.F. Rosenberger, L.M. Haines, Competing designs for phase I clinical trials: a review, Stat. Med. 21 (18) (2002 Sep 30) 2757–2770.
[30] S. Chevret, Statistical Methods for Dose-finding Experiments, 24 ed., John Wiley & Sons, Ltd, Hoboken, NJ, USA, 2006.
[31] N. Ting, Dose Finding in Drug Development, Springer Science & Business Media, New York, NY, USA, 2006.
[32] M. Fisher, K. Cheung, G. Howard, S. Warach, New pathways for evaluating potential acute stroke therapies, Int. J. Stroke 1 (2) (2006 May) 52–58.
[33] M.S. Elkind, R.L. Sacco, R.B. MacArthur, D.J. Fink, E. Peerschke, H. Andrews, et al., The Neuroprotection with Statin Therapy for Acute Recovery Trial (NeuSTART): an adaptive design phase I dose-escalation study of high-dose lovastatin in acute ischemic stroke, Int. J. Stroke 3 (3) (2008 Aug) 210–218.
[34] S.C. Fagan, J.L. Waller, F.T. Nichols, D.J. Edwards, L.C. Pettigrew, W.M. Clark, et al., Minocycline to improve neurologic outcome in stroke (MINOS): a dose-finding study, Stroke 41 (10) (2010 Oct) 2283–2287.
[35] M. Selim, S. Yeatts, J.N. Goldstein, J. Gomes, S. Greenberg, L.B. Morgenstern, et al., Safety and tolerability of deferoxamine mesylate in patients with acute intracerebral hemorrhage, Stroke 42 (11) (2011 Nov) 3067–3074.
[36] M. Krams, K.R. Lees, W. Hacke, A.P. Grieve, J.M. Orgogozo, G.A. Ford, Acute Stroke Therapy by Inhibition of Neutrophils (ASTIN): an adaptive dose-response study of UK-279,276 in acute ischemic stroke, Stroke 34 (11) (2003 Nov) 2543–2548.
[37] H.T. Whelan, J.D. Cook, C.M. Amlie-Lefond, C.A. Hovinga, A.K. Chan, R.N. Ichord, et al., Practical model-based dose finding in early-phase clinical trials: optimizing tissue plasminogen activator dose for treatment of ischemic stroke in children, Stroke 39 (9) (2008 Sep) 2627–2636.
[38] N.R. Gonzales, J. Shah, N. Sangha, L. Sosa, R. Martinez, L. Shen, et al., Design of a prospective, dose-escalation study evaluating the Safety of Pioglitazone for Hematoma Resolution in Intracerebral Hemorrhage (SHRINC), Int. J. Stroke 8 (5) (2013 Jul) 388–396.
[39] M.H. Rahbar, N.R. Gonzales, M. Ardjomand-Hessabi, A. Tahanan, M.R. Sline, H. Peng, et al., The University of Texas Houston Stroke Registry (UTHSR): implementation of enhanced data quality assurance procedures improves data quality, BMC Neurol. 13 (2013) 61.
[40] J. O'Quigley, S. Chevret, Methods for dose finding studies in cancer clinical trials: a review and results of a Monte Carlo study, Stat. Med. 10 (11) (1991 Nov) 1647–1664.
[41] G. Casella, E.I. George, Explaining the Gibbs sampler, Am. Stat. 46 (3) (1992) 167–174.
[42] C. Cai, S. Liu, Y.A. Yuan, Bayesian design for phase II clinical trials with delayed responses based on multiple imputation, Stat. Med. 33 (23) (2014 Oct 15) 4017–4028.
[43] M.M. Hossain, Z. Chen, Comments on 'Application of an adaptive design to a randomized phase II selection trial in gastric cancer: a report of the study design' by Satoshi Morita and Junichi Sakamoto, Pharm. Stat. Pharm. Stat. 11 (3) (2012) 267–268.