Published in final edited form as:

Clin Trials. 2011 August ; 8(4): 361–369. doi:10.1177/1740774511407533.

The Superiority of the Time-to-Event Continual Reassessment Method to the Rolling Six Design in Pediatric Oncology Phase I Trials

Lili Zhao¹, Julia Lee², Rajen Mody³, and Thomas M. Braun⁴

¹Biostatistics Unit, University of Michigan Comprehensive Cancer Center

²Survey Methodology, Institute of Social Research, University of Michigan

³Department of Pediatrics and Communicable Diseases, University of Michigan

⁴Department of Biostatistics, University of Michigan, Ann Arbor, MI

Abstract

Background—The Rolling Six Design (RSD) is currently being used by the Children’s Oncology Group (COG) as their standard design for Phase I trials. Because the COG has large multi-center trials with fast accrual, the motivation for adopting the RSD is to hasten accrual and shorten the duration of their trials. However, trial suspension due to completion of follow-up still cannot be entirely avoided by the RSD. Therefore, a design that allows continuous enrollment of patients throughout the entire trial is needed.

Purpose—To demonstrate the superior performance of the Time-to-Event Continual Reassessment Method (TITE-CRM) with continuous patient recruitment relative to the RSD, in terms of identifying the maximum tolerated dose (MTD) and reducing exposure of patients to toxic doses.

Methods—Using scenarios that were based on an actual pediatric Phase I trial at the University of Michigan, Monte Carlo simulations were used to investigate the operational characteristics of RSD and TITE-CRM.

Results—The TITE-CRM treated all available patients, identified the MTD more accurately than the RSD and did not increase the probability of exposing patients to toxic doses.

Limitations—Both the TITE-CRM and RSD assume that the probability of dose limiting toxicity increases with higher dose level.

Conclusions—The TITE-CRM, which allows for continual enrollment of patients, provides a safe design for pediatric oncology Phase I trials with better accuracy than the RSD.

Corresponding author: Lili Zhao, Biostatistics Unit, University of Michigan Comprehensive Cancer Center, 300 North Ingalls Building, Room 8D11, University of Michigan, Ann Arbor, MI 48109. Phone: 734-763-6898; Fax: 734-936-4540; zhaolili@umich.edu.

Disclosure of potential conflicts of interest: No

BACKGROUND

The 3+3 Method

The Children's Oncology Group (COG) is the largest multi-center consortium for all Phases of pediatric clinical trials, and most of the patients who participate in Phase I trials with COG are less than 21 years old. As a result, the COG has designed majority of its Phase I trials using the 3+3 design [1]. In the most-common form of the 3+3 method, a cohort of three patients is assigned to a dose. If none of the three patients experience a dose-limiting toxicity (DLT), the next cohort is assigned to the next highest dose. If at least two of the three patients experience a DLT, the trial is halted, and the next lowest dose is selected as the maximum tolerated dose (MTD). If one of the three patients experiences a DLT, an additional cohort of three patients is assigned to the same dose. If no additional DLTs occur, the trial continues on the next highest dose; otherwise, the trial is halted and the next lowest dose is selected as the MTD. There are other forms of the 3+3 method, such as assigning the next cohort of patients to the next lowest dose when 2/3 or 2/6 DLTs are observed, rather than halting the trial altogether. The two cohort sizes of three and three can also be altered to general values of A and B, leading to an entire family of designs known as A+B designs [2, 3].

Selection of the 3+3 method for COG Phase I trials was based primarily under the guise that the 3+3 method is safe and avoids exposure of patients to toxic doses. However, the 3+3 design, like nearly all outcome-adaptive designs based upon a binary outcome that takes time to occur, requires the complete observation of currently enrolled patients before the next cohort of patients can be enrolled in the trial. Thus the study will be closed to further accrual until the current cohort is completely evaluated. During periods of rapid accrual, the only recourse for most designs is that patients will either have their entry delayed or will be treated off protocol with some other agent if they cannot wait for the protocol treatment.

The Rolling Six Design

Because of this existing limitation for all Phase I trial algorithms, the COG has recently adopted an alternative design to the 3+3 design known as the Rolling Six Design (RSD) [4,5] in order to reduce how often patient accrual is suspended and thereby shorten the duration of COG Phase I trials. The RSD begins like the 3+3 design and enrolls a cohort of three subjects on a dose. Specifically, a fourth subject can be enrolled if at least one of the first three patients has not been fully followed and no more than one of the previous three subjects has experienced a DLT. Likewise, a fifth (sixth) patient can be enrolled if at least one of the first four (five) subjects has not been fully followed and no more than one of the previous four (five) subjects has experienced a DLT. If at any time all enrolled subjects receiving the same dose are fully followed without experiencing toxicity, the study enrolls a new cohort of subjects on the next highest dose. Likewise, if at any time two subjects receiving the same dose experience DLT, the study enrolls a new cohort of subjects on the next lowest dose. If six patients already entered at the next lowest dose, the study is halted and the next lowest dose is selected as the MTD; see Table 1 for an outline of the decision rules. Simply put, the RSD allows for temporal overlap of the two cohorts of three subjects

used in the 3+3 design. Hence, the probability of trial suspension to further accrual is lower in the RSD design as compared to the 3+3 design.

However, large multi-center trials often will have fast accrual, with an average of 4 or 5 patients enrolled per month, relative to a DLT observation period of 28 to 42 days. Thus, many available patients still cannot be enrolled when using the RSD because of forced closure of accrual after each cohort of six patients, leading to longer trial duration than a design with seamless recruitment. Therefore, a design that allows continuous enrollment of patients throughout the entire trial is needed, especially in large multi-center COG trials.

Although the RSD maintains the same safety profile as the 3+3 method, there are inherent weaknesses of the 3+3 method that also carry over to the RSD. First, all A+B designs lack an explicit targeted rate of DLTs that defines the MTD. The implicitly targeted DLT rate for a general A+B design is a function of the values of A and B, although the precise DLT rate is not transparent or straightforward to compute. For example, it is incorrectly believed that the 3+3 design targets a DLT rate of 0.33, but the implicitly targeted rate is much lower than 0.33 [2]. Thus, it is not obvious exactly what DLT rate is being targeted with the RSD and more importantly, the RSD is inappropriate if the desired DLT rate does not match the one implicitly targeted by the RSD. Second, the RSD, like all A+B designs, makes decisions for future patients that are based solely on the outcomes of the current cohort of patients. It is statistically more efficient to incorporate the outcomes of all patients, irrespective of their temporal proximity to the next cohort of patients to be enrolled.

The Time-to-Event Continual Reassessment Method

Both of the weaknesses of the RSD (and all A+B designs) described above are easily addressed by model-based, adaptive designs, the most common of which is the Continual Reassessment Method [6, 7]. It is already well-documented that the CRM with practical modifications [8, 9], such as enrolling the first patient at the first dose and limiting escalation to adjacent doses, identifies the MTD more accurately than the 3+3 method with no increased exposure of patients to overly toxic doses [10–12]. Furthermore, the CRM is able to assign more patients to the eventual MTD than the 3+3 method, thereby increasing the probability that patients receive an efficacious dose. A recent publication demonstrates these facts specifically in Phase I trials of pediatric patients [13]. However, studies designed with the CRM also must suspend accrual of new patients until all currently enrolled patients have completed follow-up.

The solution to this limitation, which was proposed by Cheung and Chapell [14] and predates the RSD by nearly a decade, is known as the Time-to-Event CRM (TITE-CRM). Like the original CRM, for dose k , $k=1, 2, \dots, K$, the TITE-CRM assumes a parametric model, $p(d_k; \beta)$, for the association of dose d_k with the probability of DLT. Two common models are the power model $\log\{p(d_k; \beta)\} = \beta \log(d_k)$ and the logistic regression model $\log\{p(d_k; \beta)/[1 - p(d_k; \beta)]\} = 3 + \beta d_k$. Although many other possible models could be used (see [7, 15] for examples), it has been shown that results are quite robust to the choice of dose-toxicity model [15, 16].

Before the trial begins, the investigator will state an *a priori* value for probability of DLT for each dose. These values (known as a “skeleton”), are used to determine a rescaled value to assign to each d_k , i.e. d_k is not the actual clinical dose but instead a numeric value determined to promote good fit of the model during the study.

The crucial aspect of the TITE-CRM is that it uses the actual amount of follow-up, T_i , collected on subject i , as well as an indicator as to whether the subject experienced DLT ($Y_i=1$) or is still being followed without DLT ($Y_i=0$). The likelihood used by the TITE-CRM is a weighted binomial likelihood

$$L_n(\beta) = \prod_{i=1}^n (p(d_{[i]}; \beta)^{Y_i} \{1 - w(T_i)p(d_{[i]}; \beta)\}^{1-Y_i})$$

in which $d_{[i]}$ is the value in $\{d_1, d_2, \dots, d_k\}$ of the dose assigned to subject i and $0 \leq w(T_i) \leq 1$ is a weight given to subject i based upon how long they have been followed. Theoretically, the weight function $w(T_i)$ should reflect the expected distribution of DLT times, which is often unknown and difficult to estimate. A simple solution, which is shown to work well in many settings, is to assume DLTs occur uniformly over the interval $[0, \tau]$, so that $w(T_i) = T_i/\tau$. Other weight functions could be used if DLTs are more likely to occur early or late during follow-up; the functional form of $w(T_i)$ could also be adaptively changed during the study [17].

Although maximum likelihood methods could be used to estimate β , a Bayesian formulation of the TITE-CRM is more common, requiring a prior distribution to be selected for β . The appropriate prior distribution, specifically the variance of the prior distribution, is often determined through simulation so that the resulting prior distribution is “informative enough” during the beginning of the study when little data exists, yet allows the data to drive decisions later in the study when enough data has been accrued. As each subject is enrolled in the study, an updated value, usually the posterior mean, is computed for β using the data from all patients in the study. This posterior mean is then used to update the probability of DLT for each dose and the dose with probability closest to a pre-defined target is assigned to the next patient.

By using a weighted likelihood, subjects can be enrolled continuously throughout the trial without recruitment pauses, thereby shortening the duration of the trial relative to the CRM, as well as allowing every patient to be enrolled as soon as they are eligible. To promote patient safety and ease concerns that the TITE-CRM may escalate too quickly, we have adopted a number of steps that limit escalation and obtain a safety profile comparable to the RSD. These steps include:

1. The first patient is treated at the first dose (or second dose if the first dose is included only as a “fall-back” for the unlikely event of DLTs occurring in the first or second patient);
2. The dose assigned to each patient has an estimated DLT rate closest to, but not greater than the target probability;

3. Dose escalation is restricted to one level between adjacent patients;
4. Escalation from the current dose is not allowed until at least one patient assigned to the current dose completes their follow-up.
5. Discontinue the trial when the probability of DLT, derived from the dose toxicity function using the posterior mean of β , at the lowest dose is larger than 25%.

Note that each Phase I trial is different, and the steps above were chosen for our specific setting. Step (4) above could be modified to further slow escalation by requiring more than one patient have complete follow-up before dose escalation can occur.

MOTIVATION FOR CURRENT RESEARCH

The TITE-CRM has already been used successfully in a published adult oncology trial [18], and a more generalized version of the TITE-CRM has been used successfully for simultaneous dose-and schedule-finding in adults [19]. However, because the TITE-CRM incorporates subjects with incomplete follow-up, i.e. who have not yet experienced a DLT but may later in the trial, the TITE-CRM has seen limited use in pediatric trials due to fears of rapid escalation and exposure of too many patients to overly toxic doses. To alleviate these fears, Normolle and Lawrence [20] compared the TITE-CRM to the 3+3 method in trials with delayed toxicity induced by a combined chemotherapy and radiation therapy. Their work showed that when the maximum time required to observe a DLT is significantly longer than the inter-arrival of patient times, trials designed with the TITE-CRM are significantly shorter and identify the MTD more accurately than trials designed with the 3+3 method. Most important, use of the TITE-CRM did not expose patients to significant excess risk beyond that seen with the 3+3 method.

Onar-Thomas and colleagues recently presented a comparison of the RSD to a maximum likelihood version of the CRM [21]. In this comparison, it was assumed that patients who were eligible for enrollment were lost entirely for trials using the CRM if current patients had not completed their follow-up. The authors presented results showing that, as expected, the RSD resulted in shorter trials than the CRM, but also that the CRM assigned more patients than the RSD to the eventual MTD, with no additional increase in observed toxicity.

Our work is presented as a complement to the work cited above by comparing the TITE-CRM directly to the RSD. We hypothesize that the TITE-CRM will show superior performance to the RSD, similar to the performance of the CRM relative to the 3+3 method. To that end, our manuscript compares, via Monte Carlo simulation, a modified version of the original TITE-CRM to the RSD in a real pediatric Phase I oncology trial. We evaluate the performance of the TITE-CRM and RSD in terms how many patients are assigned doses above the MTD and how often each design correctly identifies the MTD. Once we demonstrate the safety and excellent performance of the TITE-CRM relative to the RSD, we would recommend that the TITE-CRM should replace the RSD in pediatric oncology trials.

MONTE CARLO SIMULATIONS OF AN ACTUAL PEDIATRIC TRIAL

Description of Trial

A Phase I trial at the University of Michigan seeks to determine the MTD for intravenous irinotecan, given together with a fixed dose of bortezomib, for children with recurrent and/or resistant neuroblastoma. Bortezomib was administered at a fixed dose of $1.2\text{mg}/\text{m}^2$, given intravenously on days 1, 4, 8 and 11 of each cycle (21 days). Irinotecan was administered intravenously once daily from days 1–5 of each treatment cycle. Patients received a minimum of 2 cycles of therapy. A DLT is defined as one that occurs during the first two cycles (42 days from the therapy is initiated). The study will examine five dose levels of irinotecan: 30, 35, 40, 45 and 50mg. The trial will enroll 24 patients with an estimated accrual rate of ten children per year. In simulation studies, the accrual rate was assumed to be three children per month to mimic the accrual rate of a COG trial.

The investigators have defined the MTD as the dose with DLT probability of no more than 0.25. The *a priori* DLT probabilities (“skeleton”) for the five doses are, in dose order: 0.05, 0.10, 0.15, 0.25, and 0.35. The first child is assigned to 35mg, as the lowest dose of 20mg is included in the trial only as a “fall-back” dose should de-escalation be required early in the trial. We used the power model described earlier, except that we reparameterized β to $\exp(\beta)$, to model the DLT rate of each dose and assumed uniformly distributed DLT times for our weight function. A trial using the RSD also assigned the first child to 35mg of irinotecan; no other specifications were needed for the RSD.

Simulation Settings & Methods

The operating characteristics of the TITE-CRM and RSD were examined in five different scenarios, which vary by the true DLT rate of each dose. Those probabilities are displayed in Table 2. In Scenario 4, the true DLT rates are identical to the skeleton probabilities elicited from the investigators. Scenarios 2 and 3 set the true MTD at a dose below that initially believed to be the MTD; Scenario 5 sets the true MTD above the dose initially believed to be the MTD. Scenario 1, although is unlikely to be the true toxicity profile as claimed by the clinicians, is also examined to see if the trial can be discontinued quite often. These scenarios are important for gauging the ability of the TITE-CRM to realize that the MTD is at a dose lower or higher than expected and avoids assigning higher doses to patients as much as possible.

For each scenario, we ran 2,000 simulations of hypothetical Phase I trials using both the TITE-CRM and the RSD. In each simulation, we simulated patients to arrive via a Poisson process such that each potential patient arrived an average of every 10 days. An indicator of DLT for each patient was drawn from a Bernoulli distribution with the DLT probability specified in Table 2. If a patient had a DLT, the time to DLT was drawn uniformly between day 1 and the length of the follow-up period. In order to compare the operating characteristics of the TITE-CRM and the RSD, we have summarized the results of the 2,000 simulations with: (1) the percentage of trials correctly identifying the MTD, (2) the proportion of patients treated at or one dose lower than the true MTD, (3) the proportion of patients assigned to doses above the true MTD, and the proportion of worst-case DLTs (the

number of DLTs in patients treated above the MTD/ the number of enrolled patients), and (4) the sample size, trial duration and the proportion of available patients not treated due to the trial suspension.

We placed a normal distribution with mean 0 and variance 0.3 on the parameter β . The variance of 0.3 was determined through simulations. We compared the trial operating characteristics using a variance of 0.1, 0.2, 0.3, 0.5, 1.0 and 2.0. We found that values of 0.1 or 0.2 led to over-assignment of patients to overly toxic doses in Scenario 2, and 56% of trials failed to stop early in Scenario 1 with a value of 0.1. In contrast, values larger than 0.3 resulted in less accurate identification of the true MTD in Scenario 3 and 4. Thus a value of 0.3 struck a good balance between identification of the MTD and patient safety, thereby providing desired trial properties across all Scenarios.

We also performed a sensitivity analysis by comparing the two designs with several different inter-patient arrival times ranging from 5 to 60 days in Scenario 4.

Simulation Results

The average dose assignment for sequentially enrolled patients, using the TITE-CRM design, is presented in Figure 1. In Scenario 3–5, the average dose, starting at dose level 2, is gradually escalated towards the true MTD as more and more patients enrolled. Table 3 contains a summary of the operating characteristics of the TITE-CRM and the RSD. In Section A of Table 3, we see significant differences between the TITE-CRM and the RSD in their abilities to correctly identify the MTD. In scenarios 3–5, the TITE-CRM correctly identified the MTD much more often than the RSD and treated significantly more patients at or one dose lower than the true MTD. The TITE-CRM also demonstrates its ability in Scenario 1 to discontinue the trial more often when the lowest dose is very toxic. Furthermore, the RSD identified the MTD more often at one dose lower than the MTD than at the actual MTD in Scenarios 3–4. These results further emphasize the conservative nature of the RSD, which although may prove to be safe in terms of toxicity, is actually unsafe in terms of efficacy. The final dose selected by the RSD is also likely to be too low to effectively treat cancer and lead to negative Phase II trials. In Scenario 2, the difference between the TITE-CRM and the RSD is negligible, with the TITE-CRM equally divided between the MTD and the next highest dose, which has a DLT rate that is only 5% higher than the targeted DLT rate of 25%.

In Section B of Table 3, as well as in Figure 2, we find that the TITE-CRM and the RSD are equally likely to treat patients at doses above the MTD, while in scenario 2, the TITE-CRM is more likely to assign patients to doses above the MTD. This, nonetheless, does not result in higher rates of worst-case DLTs. This is mainly because the sequential dose assignments of the TITE-CRM tends to treat more patients at doses close to the actual MTD and to treat only 6.9% and 0.4% at doses 4 and 5 vs. 9.3% and 1.4% at doses 4 and 5 using the RSD. In summary, the TITE-CRM is as safe as the RSD, yet is far more likely to correctly identify the MTD. In Section C of Table 3, we see the number of available patients not treated due to the trial suspension in the RSD actually increases as the true MTD lies at higher and higher doses.

Both the TITE-CRM and RSD are very robust to a wide variety of accrual rates, based upon the percentage of simulations in which each dose is recommended as the MTD. The top left panel in Figure 3 demonstrates that the TITE-CRM consistently identifies the MTD correctly more often than the RSD, regardless of patient inter-arrival times. Moreover, the TITE-CRM assigns much more patients at or one dose lower than the true MTD across different accrual rates as shown in the top right panel. The middle left panel in Figure 3 indicates that when patients are accrued quickly (small inter-arrival times), the RSD tends to assign more patients to doses above the MTD than the TITE-CRM does. Although the TITE-CRM may assign slightly more patients than the RSD to doses above the MTD when patient accrual is low (large inter-arrival times), the differential between the RSD and the TITE-CRM is greater at small inter-arrival times than at large inter-arrival times. Furthermore, the worst-case DLT rate is smaller in the TITE-CRM than in the RSD until the inter-patient arrival time reaches 30 days where the DLT rates in two designs are very close (see the middle right panel). The remaining panels in Figure 3 demonstrate that for short inter-arrival times, significant number of patients would have to be turned away using the RSD, leading to longer trial duration than using the TITE-CRM designs. Therefore, the TITE-CRM is as safe, or perhaps slightly safer, and far more efficient and accurate than the RSD in trials with rapid accrual (such as 6–8 days inter-patient arrival times in COG trials).

We also did a sensitivity analysis in Scenarios 2–5 by simulating different times to DLT (TTD) using truncated normal distributions with different means and standard deviations. The (mean, standard deviation) combinations were (0, 21), (21, 1000), (21, 5), (25, 15), (31, 15), (41, 15) which lead to early-DLT (those before day 21) rates of 70%, 50%, 50%, 42%, 31%, and 17%, respectively. We found that both the TITE-CRM and RSD are very robust to different TTDs, based upon the percentage of simulations in which each dose is recommended as the MTD. The RSD is also robust based on the worst-case DLT rates, but trials could run longer (7–8 more days) when the early-DLT rate was decreased from 70% (early-onset DLT) to 17% (late-onset DLT). Conversely, the TITE-CRM is very robust based on the trial duration, and the worst-case DLT rate is not very sensitive to the TTDs except that in Scenario 2, the worst-case DLT rate was increased from 12% to 14% when the early-DLT rate was decreased from 70% to 17%, and an adaptive weight function would help mitigate this.

DISCUSSION

Investigators in pediatric oncology may prefer using the RSD because it is easy to implement and does not require the use of a computer [7]. Thus, they can make dose assignments at the clinic without assistance of a statistician. However, Phase I studies using the CRM have existed in adult Phase I trials for nearly 10 years and the TITE-CRM is now becoming a viable study option in the presence of rapid accrual.

The main reason that the TITE-CRM has not yet been more widely utilized in pediatric clinical trials is mostly due to its complexity relative to the RSD. Before running a TITE-CRM trial, one must select a dose-toxicity model, an initial guess of the expected DLT rates, and a degree of confidence regarding this initial guess. Nonetheless, it has been shown that the performance of the TITE-CRM is fairly robust to the choice of dose-toxicity model.

Although opponents of the TITE-CRM may claim it is difficult to accurately specify an initial guess for the expected DLT rates, our simulations have shown that even if the initial DLT rates are wrong, we still obtain better operating characteristics than the RSD. In addition, we can calibrate our belief in the initial DLT rates through the variance used in the prior distribution for the dose-toxicity parameter β . A small variance would put more weight on the prior guess of the DLT probabilities, thereby causing the trial to choose the hypothesized MTD more often. Conversely, a large variance will increase the chance of choosing the MTD when it is far from the hypothesized MTD. For example, the TITE-CRM in Scenario 2 recommended the actual MTD and the next highest dose equally, which is due to the prior belief that the fourth dose was the MTD; a slightly less informative prior for the dose-toxicity parameter β could mitigate this result. For any trial design, the variance, as well as different stopping rules, can be viewed as tuning parameters that are calibrated through extensive simulations in order to obtain desired operating characteristics.

A library of functions for the statistical package **R** now exist for the TITE-CRM, easily facilitating the use of the TITE-CRM with future COG trials. Furthermore, COG created the Remote Data Entry (RDE) system to facilitate dose assignment and avoid unnecessary trial stoppage due to slowed communication between data managers and statisticians. It seems probable that the TITE-CRM could be easily incorporate into the RDE, further increasing the use of the TITE-CRM with future COG trials.

Acknowledgments

We would like to thank the Editor and an Associate Editor, for their constructive comments and suggestions, which led to an improved paper.

Appendix: R programs

Manual and download information of dfcrm package can be found at <http://cran.r-project.org/web/packages/dfcrm>

```
# an example for simulations in Scenario 4
```

```
install.packages("dfcrm")
```

```
Library(dfcrm)
```

```
PI= c(0.05, 0.1, 0.15, 0.25, 0.35)
```

```
Prior= c(0.05, 0.1, 0.15, 0.25, 0.35)
```

```
foo<- titesim(PI, prior, target=0.25, nn=24, x0=2, nsim=2000, restrict=TRUE, obswin=42,
rate=42/10, scale = sqrt(0.3), accrual="poisson", scheme ="linear")
```

```
# the current titesim function allows stopping rule (1) and (3), and other stopping rules will
be incorporated at the next iteration of the dfcrm package. For right now, the stopping rules
can be easily added to the "restrict=TRUE" loop in the source code.
```

References

1. Dixon WJ, Mood AM. A method for obtaining and analyzing sensitivity data. *J Amer Statist Assoc.* 1948; 43:109–126.
2. Ivanova A. Escalation, group and A+B designs for dose-finding trials. *Statistics in Medicine.* 2006; 25:3668–3678. [PubMed: 16381057]
3. Lin Y, Shih WJ. Statistical properties of the traditional algorithm-based designs for phase I cancer clinical trials. *Biostatistics.* 2001; 2:203–215. [PubMed: 12933550]
4. Skolnik JM, Barrett JS, Jayaraman B, Patel D, Adamson PC. Shortening the timeline of pediatric phase I trials: the rolling six designs. *Journal of Clinical Oncology.* 2008; 26:190–5. [PubMed: 18182661]
5. Hartford C, Volchenboum SL, Cohn SL. 3 + 3 (Rolling) 6. *Journal of Clinical Oncology.* 2008; 26:170–171. [PubMed: 18182656]
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. [PubMed: 2350571]
7. Garrett-Mayer E. The continual reassessment method for dose-finding studies: a tutorial. *Clinical Trials.* 2006; 3:57–71. [PubMed: 16539090]
8. Faries D. Practical modifications of the continual reassessment method for Phase I cancer clinical trials. *J Biopharm Stat.* 1994; 4:147–164. [PubMed: 7951271]
9. Goodman SN, Zahurak ML, Piantadosi S. Some practical improvements in the continual reassessment method for Phase I studies. *Stat Med.* 1995; 14:1149–1161. [PubMed: 7667557]
10. Ahn C. An evaluation of Phase I cancer clinical trial designs. *Statistics in Medicine.* 1998; 17:1537–49. [PubMed: 9699228]
11. Storer BE. An evaluation of Phase I clinical trial designs in the continuous dose-response setting. *Stat Med.* 2001; 20:2399–408. [PubMed: 11512130]
12. Korn E, Midthune D, Chen T, et al. A comparison of two Phase I trial designs. *Stat Med.* 1994; 13:1799–806. [PubMed: 7997713]
13. Onar A, Kocak M, Boyett JM. Continual reassessment method vs. traditional empirically-based design: Modifications motivated by phase I trials in pediatric oncology by the pediatric brain tumor consortium. *J Biopharm Stat.* 2009; 19(3):437–455. [PubMed: 19384687]
14. Cheung K, Chappell R. Sequential designs for Phase I clinical trials with late-onset toxicities. *Biometrics.* 2000; 56:1177–1182. [PubMed: 11129476]
15. Paoletti X, Kramar A. A comparison of model choices for the continual reassessment method in phase I cancer trials. *Statistics in Medicine.* 2009; 28:3012–3028. [PubMed: 19672839]
16. Shen L, O'Quigley JO. Consistency of continual reassessment method under model misspecification. *Biometrika.* 1996; 83:395–405.
17. Braun T. Generalizing the TITE-CRM to adapt for early- and late-onset toxicities. *Statistics in Medicine.* 2005; 25:2071–2083. [PubMed: 16217853]
18. Muler JH, McGinn CJ, Normolle D, et al. Phase I trial using a time-to-event continual reassessment strategy for dose escalation of cisplatin combined with gemcitabine and radiation therapy in pancreatic cancer. *Journal of Clinical Oncology.* 2004; 22:238–243. [PubMed: 14665608]
19. de Lima M, Giralt S, Thall PF, et al. Maintenance Therapy with Low-Dose Azacitidine after Allogeneic Hematopoietic Stem Cell Transplantation for Relapsed AML or MDS: a Dose and Schedule Finding Study. *Journal of Clinical Oncology.* (in press).
20. Normolle D, Lawrence T. Designing dose escalation trials with late onset toxicities using the time-to-event continual reassessment method. *Journal of Clinical Oncology.* 2006; 24:4426–4433. [PubMed: 16983110]
21. Onar-Thomas A, Xiong Z. A simulation-based comparison of the traditional method, Rolling-6 design and a frequentist version of the continual reassessment method with special attention to trial duration in pediatric Phase I oncology trials. *Contemporary Clinical Trials.* 2010; 31:259–70. [PubMed: 20298812]

Figure 1.
Average dose assignment for sequentially enrolled patients using the TITE-CRM design.
Horizontal axis lists the ordered patient ID. Scenarios 1–5 are denoted by S1–S5.

Figure 2.

Boxplots summarize the safety and trial efficiency for the TITE-CRM and RSD. The top panels show the percentage of overdosing patients and worst-case DLT rate in each Scenarios 1–4; the bottom two panels present the trial duration and the number of available patients not treated due to trial suspension. Horizontal axis lists Scenario number S1–S5. The lower and upper boundaries of the box represent the first and third quartile, respectively; the middle line indicates mean, (◆) denotes median.

Figure 3.

The performance of the TITE-CRM relative to the RSD for various inter-arrival times in Scenario 4. The top left panel shows the percentages of selection at the true MTD given different inter-arrival times; the top right panel shows the proportion of patients assigned doses at the MTD or one dose lower than the MTD (denoted by MTD-). The middle panels show the percentage of patients treated at the overly toxic dose and the worst-case DLT rate for different inter-arrival times. The bottom panels present the trial duration and the number of available patients not treated due to trial suspension for different inter-arrival times. Horizontal axis lists patient inter-arrival times (days).

Table 1

Decision rules in RSD

# Pts Enrolled	# Pts with DLT	# Pts with Data Pending	Decision
1	–	–	Same dose level
2	2	–	De-escalate *
	else	–	Same dose level
3	2	–	De-escalate *
	0	0	Escalate **
	else	–	Same dose level
4	2	–	De-escalate *
	0	0	Escalate **
	else	–	Same dose level
5	2	–	De-escalate *
	0	0	Escalate **
	else	–	Same dose level
6	2	–	De-escalate *
	1	0	Escalate **
	0	1	Escalate **
	else	–	Suspend

* If six patients already entered at next lower dose level, the MTD has been defined; if de-escalation occurs at the first dose level, then the study is discontinued

** If final dose level has been reached, the MTD has been reached.

Table 2

True probability of DLT used in simulations

Scenario	Dose levels				
	1	2	3	4	5
1	0.40	0.50	0.60	0.70	0.80
2	0.15	0.22	0.30	0.40	0.50
3	0.08	0.15	0.22	0.30	0.40
4	0.05	0.10	0.15	0.25	0.35
5	0.02	0.05	0.10	0.15	0.22

Table 3

Operating characteristics of TITE-CRM and RSD

Scenario	Design	(A)					(B)		(C)		
		% of selection at MTD					% ppts treated at MTD or MTD- ^{**}		Observed DLT rates	Trial duration	Average number of ppts enrolled
1	TITE-CRM	8 (90 [*])	2	0	0	0	n/a				
	RSD	20 (68)	11	0	0	0	n/a		54 (54 [‡])	159	12
2	TITE-CRM	16 (11)	33	35	5	0	58	42	27 (13)	267	23
	RSD	29 (9)	35	21	6	1	66	35	26 (13)	214	13 (5)
3	TITE-CRM	3 (2)	19	51	22	3	77	17	21 (5)	279	24
	RSD	19 (2)	32	28	14	4	71	18	22 (6)	246	15 (6)
4	TITE-CRM	1 (0)	5	41	41	12	66	6	17 (2)	283	24
	RSD	12 (0)	21	31	25	10	49	9	18 (3)	285	17 (8)
5	TITE-CRM	0 (0)	0	16	40	44	45	n/a	n/a	283	24
	RSD	3 (0)	12	19	27	39	39	n/a	n/a	332	20 (9)

* average percentage of discontinued trials

** MTD- is one dose lower than the MTD

[‡] number of DLTs in patients treated above the MTD/number of patients enrolled, averaged across 2000 simulations.

[†] average number of available patients not treated

ppts = patients