Biostatistics Department Technical Report

BST2015-001

Two-Group Time-To-Event Continual Reassessment Method Using Likelihood Estimation

Amber Salter
John O'Quigley
Gary R. Cutter
Inmaculada B. Aban

March 2015

Department of Biostatistics
School of Public Health
University of Alabama at Birmingham
asalter@uab.edu
caban@uab.edu

Two-Group Time-To-Event Continual Reassessment Method Using Likelihood Estimation

Amber Salter, MPH[1], John O'Quigley, PhD[2], Gary R. Cutter, PhD[1] Inmaculada B. Aban, PhD[1]

[1] Department of Biostatistics, University of Alabama at Birmingham, Birmingham, AL USA

[2] Universit Paris VI, Paris, France

February 24, 2015

Abstract

The presence of patient heterogeneity in dose finding studies is inherent. When not accounted for in the trial design, some subjects may be exposed to toxic or suboptimal doses. Options to handle patient heterogeneity include conducting separate trials or splitting the trial into arms; however, cost or lack of resources limit the feasibility of separate trials. If information is shared between the groups, then both of these options do not benefit from using the shared information. Extending current dose finding designs to handle patient heterogeneity maximizes the utility of existing methods within a single trial. We propose a modification to the time-to-event continual reassessment method to handle two groups by using a two-parameter model and maximum likelihood estimation method. The operating characteristics of the design are investigated through simulations under different scenarios including the scenario where one conducts two separate trials, one for each group, using the one-sample time-to-event continual reassessment method.

Keywords:dose finding, phase I trial, patient heterogeneity, continual reassessment method, time-to-event, maximum likelihood

1 INTRODUCTION

Dose finding trials are conducted in order to investigate the relationship between dose and safety of a drug. The goal of these studies is to find the optimal dose that is considered safe, usually based on the toxicity responses in those being tested. This dose is usually called the maximum tolerated dose (MTD) defined as the highest dose with the expected probability that is closest to the target rate of toxicity (TRT). The TRT is the level of risk researchers are willing to accept for how likely a patient is to experience a toxic event and is defined before the study begins. The TRT in cancer studies is between 0.2 to 0.3 while stroke studies use a TRT of 0.1 or less[1]. The MTD resulting from dose finding trials often provide the dose used in future trials. The impact of these studies is significant for the future of the drug.

To minimize heterogeneity, a trial sample is often selected using inclusion and exclusion criteria. However, situations where subject heterogeneity arises are known. Examples include being naive versus non-naive to treatment or the presence or absence of a biomarker. We consider having 2 groups which assume the relationship between the dose and the probability of toxicity to be different but some information is shared. An option available to the researcher when this heterogeneity is present is to conduct 2 separate trials or split the trial into 2 arms. However, the availability of resources and not utilizing the shared information may preclude their use. Another option is to simply use the one-sample CRM and therefore ignore any heterogeneity which may be present. Using the one-sample CRM in this case as previously described[2, 3] results in a single MTD being estimated that is an average of the two groups. Consequently, subjects are treated at overly toxic or suboptimal doses as shown by O'Quigley[4] using the one-sample CRM. Finally, an alternative option is to use a design that accounts for the heterogeneity. O'Quigley[4, 5] proposed an extension of the CRM, the two-group CRM, to estimate the MTD for both groups in the same trial using a model with two parameters.

The objective of this paper is to address the issue of patient heterogeneity based on the two-group two-parameter models used in the CRM applied to the time-to-event CRM (TITE-CRM) for

2 groups. Instead of estimating the next recommended dose at the end of follow-up, TITE-CRM estimates a recommended dose upon enrollment of a new subject utilizing any partial information available on subjects. We derive the estimating equations to obtain maximum likelihood estimates and use simulations to investigate the operating characteristics of the design. The outline of the paper is as follows. In section 2, we define notations, assumptions and the model. Section 3 derives the estimating equations for the maximum likelihood estimator (MLE) and describes the estimation process. Section 4 discusses the simulation design and the results of our evaluation of the operating characteristics. Finally, the last section provides some concluding remarks.

2 METHODS

2.1 Notation and Assumptions

We define n and m be the total number of subjects in each group so that the total sample size for the trial is $N = n + m$. Let Y_j, $j = 1, ..., n$ and V_l, $l = 1, ..., m$ be the binary outcomes (1=toxicity, 0=no toxicity) for group 1 and group 2, respectively where Y_j and V_l are independent for all j and l. The clinical doses, $d \in \{d_1, ..., d_r\}$, chosen for the trial (e.g. 10, 50, 100, 150, 300, 400 milligrams) are associated with a probability of toxicity and a dose-toxicity function, $f(x, \beta)$, is used to describe the relationship between them. The function should be strictly increasing and although the dose-toxicity function is continuous, the estimation is for a set of discrete doses. The common choices for $f(x, \beta)$ which satisfy the necessary conditions include the empiric function, $f(x, \beta) = x^{e^\beta}$, hyperbolic tangent function, $f(x, \beta) = \left(\frac{\tanh x + 1}{2}\right)^\beta$, and logistic function, $f(x, \beta) = \frac{\exp(a + \beta x)}{1 + \exp(a + \beta x)}$. To ensure the probabilities of toxicity range between 0 and 1, we solve for the x in $f(x, \beta)$ where x is the standardized dose units representing the clinical dose levels in the dose-toxicity function. For example, given a dose, d_3 of 100 milligram with an associated probability of toxicity of 0.2 and $f(x, \beta) = x^{e^\beta}$, the corresponding x_3 is $x_3 = \exp\{\frac{ln(0.2)}{e^\beta}\}$.

To incorporate partial information and to accommodate two groups, let the dose-toxicity function for the i^{th} group be denoted by $G^{(i)}(x, w, \beta^{(i)}) = wf^{(i)}(x, \beta^{(i)})$, $i = 1, 2$ where w is the weight which is a function of the time-to-event for the k^{th} patient and $0 \leq w \leq 1$. Thus, $G(x, w, \beta)$ is a weight-adjusted dose-toxicity function which is monotone increasing in w and has the marginal constraints $G(x, 0, \beta)$ = 0 and $G(x, 1, \beta)$ = $f(x, \beta)$ for all x, β [6]. Let w_j be the weight for the j^{th} subject in group 1 and w_l be the weight for the l^{th} subject in group 2 at the estimation before the enrollment of the $(k + 1)^{th}$ subject. We consider the case where the two groups have the same parametric form of the dose-toxicity function. Let x_j and x_l be the actual dose assigned to subjects in each group. The target rate of toxicity, θ, is assumed to be the same for both groups.

If no information is shared between the two groups, conducting two separate trials is reasonable. If the two groups share some information, one can take advantage of that shared information to conduct a single trial. How can this information be incorporated in the model? One way to handle this is to introduce two parameters in the model as done in the two-group CRM by O'Quigley [4]. One parameter describes the information they share and the other parameter measures the difference between the two groups. We refer to this as the two-parameter model. Another approach is to use a shift model as described by O'Quigley [7–9]. With this model, the differences between the groups are limited to a small finite set of differences. The second group is a shift (in either direction) of the dose-toxicity function specified for the first group. This results in the MTD being located at a higher, lower or the same dose level for the second group compared to the first. Specifying each shift allows for only one parameter to be estimated but requires estimation of the candidate models for each shift. The shift model with the highest estimated likelihood value is chosen and used to find the recommended dose for each group.

For this paper, we will be focusing on the two-parameter model. There are multiple specifications of the parameters that allow for information to be shared between the two groups. We chose to use the additive relationship and define the two parameters as $\beta^{(1)} = \beta$, $\beta^{(2)} = \beta + \tau$ for $-\infty < \beta < \infty, -\infty < \tau < \infty$. Consequently, the dose-toxicity function for group 1 is $G(x, w, \beta)$ and

the dose-toxicity function for group 2 is $G(x, w, \beta + \tau)$.

2.2 Likelihood Estimation

As with the one-sample likelihood-based TITE-CRM, the two-group likelihood-based TITE-CRM needs at least one toxic outcome and non-toxic outcome in each group to begin using the full likelihood estimation. Define the initial dose escalation (IDE) stage as the first stage employed until both outcomes (toxicity and no toxicity) have been observed in each group. There are multiple options for the IDE in the two-group setting [4]. These options involve choosing the dose level to start the trial followed by escalating each group independently or jointly to be able to move out of the IDE stage.

During the IDE stage, a group may use the one-sample likelihood TITE-CRM once both outcomes have been observed in that group while waiting for the other group to observe both outcomes. The second stage begins once both outcomes have been observed in both groups. At that time, the full two-group likelihood TITE-CRM can be employed.

Let $k^{(i)}, i = 1, 2$ denote the number of subjects from the i^{th} group and $k = k^{(1)} + k^{(2)}$. The likelihood of this model after enrolling k subjects is given by

$$\begin{aligned} L(\beta, \tau | y_j, v_l) &= \prod_{j=1}^{k^{(1)}} G(x_j, w_j, \beta)^{y_j} (1 - G(x_j, w_j, \beta))^{(1-y_j)} \\ &\times \prod_{l=1}^{k^{(2)}} G(x_l, w_l, \beta + \tau)^{v_l} (1 - G(x_l, w_l, \beta + \tau))^{(1-v_l)} \\ &= \prod_{j=1}^{k^{(1)}} (w_j f(x_j, \beta))^{y_j} (1 - w_j f(x_j, \beta))^{(1-y_j)} \\ &\times \prod_{l=1}^{k^{(2)}} (w_l f(x_l, \beta + \tau))^{v_l} (1 - w_l f(x_l, \beta + \tau))^{(1-v_l)} \end{aligned}$$

The MLE for β and τ based on the first k observations are solutions to the equations

$$\begin{aligned}
\frac{\partial \ell}{\partial \beta} &= \sum_{j=1}^{k^{(1)}} y_j \frac{\frac{\partial}{\partial \beta} f(x_j, \beta)}{f(x_j, \beta)} + \sum_{j=1}^{k^{(1)}} (1 - y_j) \frac{-w_j \frac{\partial}{\partial \beta} f(x_j, \beta)}{1 - w_j f(x_j, \beta)} \\
&+ \sum_{l=1}^{k^{(2)}} v_l \frac{\frac{\partial}{\partial \beta} f(x_l, \beta + \tau)}{f(x_l, \beta + \tau)} + \sum_{l=1}^{k^{(2)}} (1 - v_l) \frac{-w_l \frac{\partial}{\partial \beta} f(x_l, \beta + \tau)}{1 - w_l f(x_l, \beta + \tau)} = 0 \qquad (1) \\
\frac{\partial \ell}{\partial \tau} &= \sum_{l=1}^{k^{(2)}} v_l \frac{\frac{\partial}{\partial \tau} f(x_l, \beta + \tau)}{f(x_l, \beta + \tau)} + \sum_{l=1}^{k^{(2)}} (1 - v_l) \frac{-w_l \frac{\partial}{\partial \tau} f(x_l, \beta + \tau)}{1 - w_l f(x_l, \beta + \tau)} = 0 \qquad (2)
\end{aligned}$$

where $\frac{\partial f}{\partial \beta}$ and $\frac{\partial f}{\partial \tau}$ are the partial derivatives of $f(\cdot\,, \cdot\,)$ with respect to β and τ, respectively. Since the likelihood was based on the first k subjects, we denote the MLE at this stage as $\hat{\beta}_k$. The solutions to these equations typically have no closed analytical form even for the commonly chosen function. To illustrate, consider the empiric dose-toxicity function, $f(x, \beta) = x^{e^{\beta}}$. The estimating equations are then given by

$$\begin{aligned}
\frac{\partial \ell}{\partial \beta} &= \sum_{j=1}^{k^{(1)}} y_j \ln(x_j) + \sum_{j=1}^{k^{(1)}} (1 - y_j) \frac{-w_j \ln(x_j) x_j^{e^{\beta}} e^{\beta}}{1 - w_j x_j^{e^{\beta}}} \\
&+ \sum_{l=1}^{k^{(2)}} v_l \ln(x_l) + \sum_{l=1}^{k^{(2)}} (1 - v_l) \frac{-w_l \ln(x_l) x_l^{e^{\beta+\tau}} e^{\beta+\tau}}{1 - w_l x_l^{e^{\beta+\tau}}} = 0 \\
\frac{\partial \ell}{\partial \tau} &= \sum_{l=1}^{k^{(2)}} v_l \ln(x_l) + \sum_{l=1}^{k^{(2)}} (1 - v_l) \frac{-w_l \ln(x_l) x_l^{e^{\beta+\tau}} e^{\beta+\tau}}{1 - w_l x_l^{e^{\beta+\tau}}} = 0
\end{aligned}$$

which requires numerical approximation methods to obtain $\hat{\beta}_k$ and $\hat{\tau}_k$.

The revised dose-toxicity function is defined using $\hat{\beta}_k$ and $\hat{\tau}_k$ with the same target rate of toxicity, θ. The next dose for the $k + 1$ subject is determined by

$$x_{k+1}^{(1)} = \arg\min_x |G(x, w, \hat{\beta}_k) - \theta| \qquad \text{(dose for Group 1)}$$

or

$$x_{k+1}^{(2)} = \arg\min_x |G(x, w, \hat{\beta}_k + \hat{\tau}_k) - \theta| \qquad \text{(dose for Group 2)}$$

where $\arg\min_x$ is the argument of the minimum corresponding to the dose level for which $|G(x, w, \hat{\beta}_k - \theta|$ or $|G(x, w, \hat{\beta}_k + \hat{\tau}_k) - \theta|$ attain their respective minima. The next dose recommended depends on the group membership of the $k+1$ subject. This procedure continues in this manner to obtain $(\hat{\beta}_{k+1}, \hat{\tau}_{k+1}, x^{(1)}_{k+2}, x^{(2)}_{k+2})$, $(\hat{\beta}_{k+2}, \hat{\tau}_{k+2}, x^{(1)}_{k+3}, x^{(2)}_{k+3})$, $\ldots$, $(\hat{\beta}_N, \hat{\tau}_N, x^{(1)}_{N+1}, x^{(2)}_{N+1})$, where $\hat{\beta}_N$ and $\hat{\tau}_N$ are the MLEs of β and τ using data from all subjects. The doses, $(x^{(1)}_{N+1}, x^{(2)}_{N+1})$, defined by

$$x^{(1)}_{N+1} = \arg\min_x |G(x, w, \hat{\beta}_N) - \theta|$$

and

$$x^{(2)}_{N+1} = \arg\min_x |G(x, w, \hat{\beta}_N + \hat{\tau}_N) - \theta|$$

are the trial MTDs for each group. The trial ends when N subjects are enrolled.

3 OPERATING CHARACTERISTICS

Conventionally when working with MLEs the interest for finite samples is to assess the properties of the estimators. For dose finding methods, the primary goal is to find the MTD, and in the CRM framework, the dose-toxicity function assumed and the parameter(s) associated with this function are conduits to assist in finding the MTD. What is relevant are the operating characteristics of the dose finding method. The operating characteristics describe how well the method recommends the correct dose as the MTD, how well the method allocates the doses, and how well the method minimizes the proportion of persons who experience toxicity [10]. We use simulations to investigate the performance of the two-group TITE-CRM from its operating characteristics.

3.1 Simulation Method

Sample size, N, was fixed at 32 subjects and used 6 dose levels. The dose-toxicity function chosen was the empiric function, $f(x, \beta^{(i)}) = x^{e^{\beta^{(i)}}}, i = 1, 2$ where $\beta^{(1)} = \beta$ and $\beta^{(2)} = \beta + \tau$ and x

is the standardized units representing the dose levels in the dose-toxicity function. Parameter values for the dose-toxicity relationships (DTR) and used to obtain the probability of toxicity for group 2 based on group 1 were: DTR 1-3: $\beta = 1$ and $\tau = 0.5$; DTR 4: $\beta = 0.75$ and $\tau = 0.75$. Table 1True Dose-Toxicity Relationships (**True MTD**)table.caption.2 shows the probability of toxicity at each dose level for a given scenario where the MTD of each group is bolded. Each DTR used these probabilities as their initial working model or the initial estimates of the dose-toxicity relationship.

The target rate of toxicity was set at 0.2 which is within the range used in cancer trials as previously mentioned. Conditional on experiencing a toxicity, the time to toxicity follows a uniform distribution on the observation window interval $(0, T)$ where T is the length of follow-up for the trial. The uniform weight, $w(u, T) = \frac{u}{T}$, was used for the TITE-CRM where u is the amount of time the subject has been followed. The patient accrual was fixed such that 2 subjects were enrolled over the observation window. Note that with this mechanism there will be at most 1 subject who has not completed the study. The IDE stage of the simulations used the MTD of each group as the starting dose. Escalation proceeds separately for each group, and if there were 3 toxicities at a dose level within a group, the dose would de-escalate. A trial was considered as a failure if there were 3 toxicities at the lowest dose level for either group in the first stage. Failed trials were excluded. During the second stage, dose escalation and de-escalation were not restricted.

Subject recruitment is in many ways unpredictable. In practice, when subjects present for enrollment, the timing and number of subjects for each group is difficult to predict. In keeping with this, group allocation was randomly assigned using a Bernoulli random variable. For example, the 75/25 group allocation simulated samples made up of 75% from group 1 and 25% for group 2. The one-sample case was also simulated (100/0 and 0/100). In the analyses of the simulation results, the actual group sizes for the trials will be compared. The group sizes used include 0, 2-6, 7-12, 13-18, 19-24, 25-30 and 32 where a group size of 25 to 30 in one group represents a large imbalance in the groups in the trial as this groups size implies the other group only had 2 to 6 subjects. The

13-18 group size represents a balance in the group sizes in the sample. Trials that did not have an estimate for both parameters were not included.

The implementation of the estimation requires solving equations to find the MLEs using numerical methods. We used the NLP procedure (Nonlinear Programming) in SAS v9.3 to find solutions. The Newton-Raphson with ridging optimization method was used with starting parameter values of 1 for β and 0 for τ. This optimization method is recommended for small problems ($n \leq 40$) [11]. The random number generator used the *rand* function to determine the outcome, group membership and time to toxicity variables. Two thousand simulations were performed for each of the 20 scenarios tested looking at 5 different group allocations for 4 dose-toxicity relationships resulting in a total of 40,000 trials simulated.

3.2 Results

Of the 40,000 trials simulated, DTR 3 was the only scenario where failed trials were observed and this occurred in 0.1% of the simulated trials (Table 2Number of Failed Trials in Each Allocation for DTR 3table.caption.3). We believe that this is a result of the MTD for group 1 being at the lowest level in this scenario. When planning a trial, it is important to be cognizant that failed trials are more likely to occur when the MTD is set at the lowest dose level. A procedure for handling or avoiding these occurrences should be discussed before the study begins.

Additionally, we required the trials to be analyzed to have both parameters estimated in the two-group case (i.e. the trial made it to the second stage) and the single parameter estimated in the one-group case. This resulted in 1767 (4.4%) additional trials being excluded. The number of trials used for each group size are presented in Table 3Number (Col %) of Simulations from Each Group Allocation Used in Each Group Size by Dose Level and Dose-Toxicity Relationship (DTR)table.3 and are used for the subsequent results unless otherwise specified. The one-sample allocations represent separate trials conducted for each group using all N subjects and provide a comparison of performance.

3.2.1 Recommended MTD Proportion

Recall that for these simulations we started the trial at the MTD for each group based on the true DTR. We evaluate the proportion the correct MTD is recommended in the simulated trials. The corresponding one-sample proportions for each group is used as a benchmark of performance where the maximum number of subjects for a single group are used in the estimation.

Figure 1Recommended MTD Proportionfigure.caption.7 shows the proportion each dose level is recommended as the MTD in the simulations by group sizes. For the one-sample case ($n = 32$), the recommended MTD proportions range from 35% to 60% for the different DTRs. The higher the true MTD dose level the higher proportion of time the true MTD is selected. The exception being when the MTD is at dose level 1 (DTR 3) where group 1 has no other opportunity for de-escalation. As expected, there is a drop in the proportion the true MTD is recommended for both groups with decreasing sample size. Compared to the one-sample, the balanced group sizes (13-18) show a 7% to 17% decrease in proportion recommended. For other group sizes (i.e. imbalance of groups in the sample), the group with the larger sample size recommends the true MTD with only a slight reduction compared to the one-sample, but the loss for the other group compared to the one-sample is larger. This loss is expected compared to the one-sample as neither group use all 32 subjects to find the MTD. In the 50/50 allocation, the recommended proportion in the one-sample uses double the subjects of the two-sample estimation and yet, the recommended MTD proportion does not suffer as much as it could with the added advantage of getting information about another group.

3.2.2 In-trial Dose Allocation of Subjects

Another characteristic of interest is the distribution of doses the subjects treated at each dose. The aim is to treat as many subjects as possible at the MTD. While some subjects may be exposed to doses one level above the MTD, treating beyond that level should be minimized.

Table 4In-Trial Dose Allocation of Subjects by Group Size for Each Dose-Toxicity Relationship (DTR): Proportion of Subjects at Each Dose Level (**True MTD**)table.caption.5 shows the proportion of subjects allocated to each dose across all simulated trials. In general, allocations are similar to the one-sample TITE-CRM. For DTRs 1 and 2, 6% or less of the trials allocated subjects to doses two or more dose levels above the MTD. Group 1 in DTRs 3 and 4 treat more subjects at 2 or more dose levels higher than the MTD. However, this is consistent with the results for the one-sample case. The highest proportion of subjects are treated at the MTD for each DTR and group which is reassuring.

3.3 Two-Group TITE-CRM in the Case of a Homogeneous Sample

Recall that the design assumes the sample has two groups with different MTDs from the onset. However, this assumption can be incorrect and the sample is actually homogeneous (no heterogeneity exists). While the truth is usually unknown in practice, it is helpful to know the effect of this misspecification of group relationship in the event this design is used when the one-sample TITE-CRM may be more applicable. In this situation, the initial working model and the method used to estimate the MTD do not accurately reflect the underlying true dose-toxicity probabilities. We investigate the impact of this scenario on the recommended MTD proportions and in-trial dose allocations.

Assume that the common true dose-response probabilities for both groups is DTR 2 Group 1 in Table 1True Dose-Toxicity Relationships (**True MTD**)table.caption.2 and the initial working model used DTR 1 for each group, respectively. The initial dose level was the MTD of the working model where one group started one dose level below the MTD and the other group started one dose level above the MTD. Failed trials were excluded in the simulation summary as well as trials that did not reach the second stage. Two thousand simulations were performed for 3 different group allocations resulting in 6,000 trials simulated. A total of 5481 trials from the 75/25, 50/50 and 25/75 group allocations were used in the results after excluding the trials that did not reach the

second stage.

Since the primary concern in these trials is to accurately recommend the MTD at the end of the trial for each group, we compared the concordance of the recommended doses for each group after each simulated trial (percent of the total trials that fall in each cell, Table 5Concordance of the Recommended Maximum Tolerated Dose (MTD) in the Homogeneous Sample, N (%N)table.caption.6) regardless of the allocation of subjects to each group. Table 5Concordance of the Recommended Maximum Tolerated Dose (MTD) in the Homogeneous Sample, N (%N)table.caption.6 shows that the method recommends the true MTD (dose level 3) in both groups only 14% of the time. Almost 50% of the time, the recommended dose is correct for at least one group and the other group is one dose level above or below the correct MTD dose level. In 25% of trials, the same dose level is selected as the MTD (i.e. along the diagonal) which is lower than to the proportion of time the one-sample selects the correct MTD (33% and 36% for each one-sample scenario).

4 DISCUSSION

In this paper, we address the issue of patient heterogeneity in a single trial for dose finding studies using TITE-CRM. We showed how the TITE-CRM can accommodate two groups by extending the two-group CRM design of O'Quigley [4] using likelihood-based methods. The general form of the estimating equations provides a flexible framework which can be used to customize the design to a specific trial. We examined the performance of the two-group TITE-CRM design in the presence of patient heterogeneity, including imbalances in group size and varying magnitude and difference in the MTD between groups. The decision to conduct the design incorporating both groups into a single trial instead of conducting 2 separate trials has the benefit of maximizing the human and financial resources with an impact on the recommended MTD dependent on the sample size allocation between the two groups.

While there is benefit using the two-parameter two-group TITE-CRM when patient heterogeneity is present, when the sample truly is homogeneous this method has less than ideal performance. The common true MTD is not recommended frequently and the estimation often leads to two different MTDs. Careful consideration should be given to using this method if the confidence regarding the presence of patient heterogeneity is weak.

As a result of using likelihood estimation, additional challenges are presented in the two-group setting. Estimation requires the observation of at least one toxic and a non-toxic event in both groups before valid parameter estimates can be obtained. An optimal strategy to accomplish this is unclear. Of course, the Bayesian method could be used and eliminate the need for the two stages. Only one IDE algorithm was considered in the simulations where the IDE started at the true MTD of each group. Other IDE options could impact how subjects are used early in the trial and are expected to affect the operating characteristics of the design. Another issue resulting from using likelihood estimation stems from the need to use numerical methods to find solutions. The method has the potential to be sensitive to the choice of optimization technique or starting parameter value [11]. Choosing a weight which needs to estimate another parameter or specifying a multiplicative group relationship may cause more challenges in obtaining solutions. Potential issues that may result with other design specifications warrant further exploration in the planning phase of a trial.

References

[1] M. Fisher, K. Cheung, G. Howard, and S. Warach, "New pathways for evaluating potential acute stroke therapies," *International Journal of Stroke*, vol. 1, pp. 52–58, 2006.

[2] J. O'Quigley, M. Pepe, and L. Fisher, "Continual reassessment method: A practical design for phase 1 clinical trials in cancer," *Biometrics*, vol. 46, no. 1, pp. 33–48, 1990.

[3] J. O'Quigley and L. Shen, "Continual reassessment method: A likelihood approach," *Biometrics*, vol. 52, no. 2, pp. 673–684, 1996.

[4] J. O'Quigley, L. Shen, and A. Gamst, "Two-sample continual reassessment method," *Journal of Biopharmaceutical Statistics*, vol. 9, no. 1, pp. 17–44, 1999.

[5] J. O'Quigley and X. Paoletti, "Continual reassessment method for ordered groups," *Biometrics*, vol. 59, pp. 430–440, 2003.

[6] Y. K. Cheung and R. Chappell, "Sequential designs for phase I clinical trials with late-onset toxicities," *Biometrics*, vol. 56, pp. 1177–1182, 2000.

[7] J. O'Quigley and M. Conaway, "Continual reassessment and related dose-finding designs," *Statistical Science: a Review Journal of the Institute of Mathematical Statistics*, vol. 25, no. 2, pp. 202–216, 2010.

[8] J. O'Quigley and M. Conaway, "Extended model-based designs for more complex dose-finding studies," *Statistics in Medicine*, vol. 30, no. 17, pp. 2062–2069, 2011.

[9] J. O'Quigley and A. Iasonos, "Bridging solutions in dose-finding problems," *Statistics in Biopharmaceutical Research*, vol. 6, no. 2, pp. 185–197, 2014.

[10] X. Paoletti and A. Kramar, "A comparison of model choices for the continual reassessment method in phase I cancer trials," *Statistics in Medicine*, vol. 28, no. 24, pp. 3012–3028, 2009.

[11] Sas Institute Inc, *SAS/STAT 9.3 User's Guide*. Cary, NC: SAS Institute Inc, 2011.

	Dose Level					
	1	2	3	4	5	6
Dose-Toxicity Relationship 1 (DTR 1)						
Group 1	0.01	**0.18**	0.25	0.36	0.50	0.70
Group 2	0.02	0.06	0.10	**0.19**	0.32	0.56
Dose-Toxicity Relationship 2 (DTR 2)						
Group 1	0.05	0.10	**0.20**	0.35	0.50	0.70
Group 2	0.007	0.02	0.07	**0.18**	0.32	0.56
Dose-Toxicity Relationship 3 (DTR 3)						
Group 1	**0.18**	0.26	0.36	0.50	0.65	0.80
Group 2	0.06	0.11	**0.19**	0.32	0.49	0.69
Dose-Toxicity Relationship 4 (DTR 4)						
Group 1	0.08	**0.18**	0.24	0.35	0.46	0.65
Group 2	0.004	0.03	0.05	0.11	**0.19**	0.40

Table 1: True Dose-Toxicity Relationships (**True MTD**)

Group Allocation	Total Number of Failed Trials	Proportion of All Simulations
100/0	7	0.4%
75/25	2	0.1%
50/50	11	0.6%
25/75	18	0.9%
0/100	0	0%

Table 2: Number of Failed Trials in Each Allocation for DTR 3

	Group Allocation				
Group 1 Size	100/0	75/25	50/50	25/75	0/100
			DTR 1		
0	0 (0)	0 (0)	0 (0)	0 (0)	2000 (100)
2 - 6	0 (0)	0 (0)	0 (0)	390 (22)	0 (0)
7 - 12	0 (0)	0 (0)	208 (10)	1297 (74)	0 (0)
13 - 18	0 (0)	38 (2)	1429 (72)	72 (4)	0 (0)
19 - 24	0 (0)	1088 (59)	354 (18)	0 (0)	0 (0)
25 - 30	0 (0)	706 (39)	2 (0)	0 (0)	0 (0)
32	2000 (100)	0 (0)	0 (0)	0 (0)	0 (0)
			DTR 2		
0	0 (0)	0 (0)	0 (0)	0 (0)	2000 (100)
2 - 6	0 (0)	0 (0)	0 (0)	417 (23)	0 (0)
7 - 12	0 (0)	0 (0)	212 (11)	1337 (76)	0 (0)
13 - 18	0 (0)	33 (2)	1422 (71)	62 (3)	0 (0)
19 - 24	0 (0)	1054 (58)	363 (18)	0 (0)	0 (0)
25 - 30	0 (0)	724 (40)	2 (0)	0 (0)	0 (0)
32	2000 (100)	0 (0)	0 (0)	0 (0)	0 (0)
			DTR 3		
0	0 (0)	0 (0)	0 (0)	0 (0)	2000 (100)
2 - 6	0 (0)	0 (0)	1 (0)	367 (22)	0 (0)
7 - 12	0 (0)	0 (0)	220 (11)	1231 (73)	0 (0)
13 - 18	0 (0)	37 (2)	1400 (71)	93 (6)	0 (0)
19 - 24	0 (0)	1056 (59)	362 (18)	0 (0)	0 (0)
25 - 30	0 (0)	703 (39)	2 (0)	0 (0)	0 (0)
32	1993 (100)	0 (0)	0 (0)	0 (0)	0 (0)
			DTR 4		
0	0 (0)	0 (0)	0 (0)	0 (0)	2000 (100)
2 - 6	0 (0)	0 (0)	0 (0)	354 (21)	0 (0)
7 - 12	0 (0)	0 (0)	190 (10)	1250 (74)	0 (0)
13 - 18	0 (0)	28 (2)	1417 (71)	79 (5)	0 (0)
19 - 24	0 (0)	1058 (59)	380 (19)	0 (0)	0 (0)
25 - 30	0 (0)	711 (40)	0 (0)	0 (0)	0 (0)
32	2000 (100)	0 (0)	0 (0)	0 (0)	0 (0)

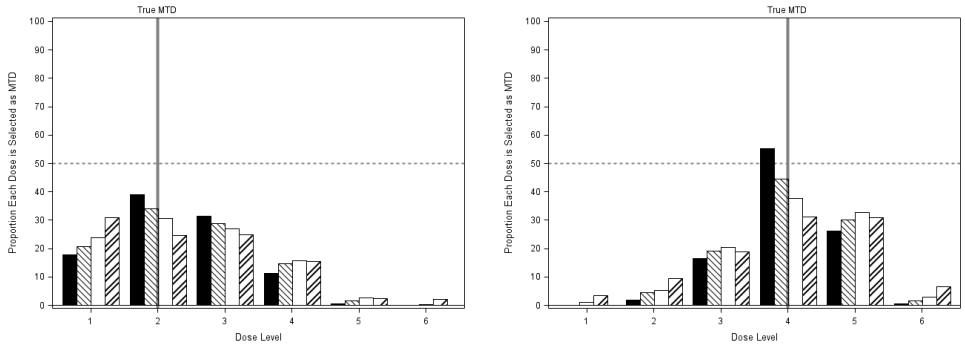
Table 3: Number (Col %) of Simulations from Each Group Allocation Used in Each Group Size by Dose Level and Dose-Toxicity Relationship (DTR)

	Group 1 Dose Level							Group 2 Dose Level					
Group Size	1	2	3	4	5	6		1	2	3	4	5	6
DTR 1													
							One Sample	3	5	16	**44**	28	4
7 - 12	25	**44**	23	7	1	0	19 - 24	3	6	15	**41**	29	5
13 - 18	23	**36**	25	14	3	0	13 - 18	4	7	14	**41**	29	5
19 - 24	20	**33**	26	17	4	0	7 - 12	7	7	12	**44**	25	4
One Sample	23	**35**	27	13	2	0							
DTR 2													
							One Sample	1	3	15	**48**	29	3
7 - 12	14	14	**47**	21	4	0	19 - 24	1	3	15	**47**	30	4
13 - 18	10	17	**43**	24	6	0	13 - 18	2	5	15	**45**	29	5
19 - 24	7	15	**45**	26	6	0	7 - 12	2	6	15	**48**	25	4
One Sample	7	20	**48**	22	4	0							
DTR 3													
							One Sample	10	17	**42**	26	5	0
7 - 12	**87**	12	0	0	0	0	19 - 24	10	15	**41**	27	6	0
13 - 18	**56**	26	15	3	0	0	13 - 18	12	15	**41**	26	6	0
19 - 24	**54**	26	15	4	0	0	7 - 12	15	12	**45**	23	5	0
One Sample	**58**	28	12	2	0	0							
DTR 4													
							One Sample	1	2	6	20	**51**	20
7 - 12	24	**44**	23	8	1	0	19 - 24	1	2	6	17	**50**	23
13 - 18	21	**36**	24	14	5	0	13 - 18	1	3	7	16	**49**	24
19 - 24	19	**33**	26	16	6	0	7 - 12	2	4	7	14	**49**	24
One Sample	20	**36**	27	13	3	0							

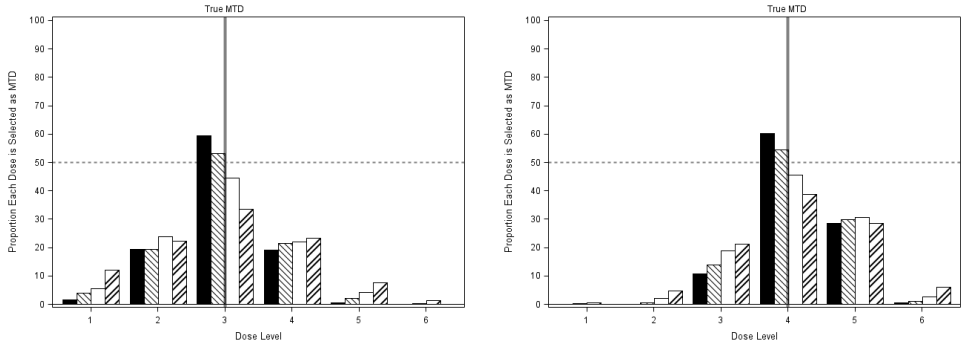
Table 4: In-Trial Dose Allocation of Subjects by Group Size for Each Dose-Toxicity Relationship (DTR): Proportion of Subjects at Each Dose Level (**True MTD**)

Recommended MTD Group 1	Recommended MTD Group 2						Total
	1	2	3 (True MTD)	4	5	6	
1	36 (1)	117 (2)	191 (3)	109 (2)	15 (0)	1 (0)	469 (9)
2	155 (3)	308 (6)	**491 (9)**	282 (5)	72 (1)	6 (0)	1313 (24)
3 (True MTD)	284 (5)	**553 (10)**	**786 (14)**	**427 (8)**	136 (2)	21 (0)	2207 (40)
4	156 (3)	313 (6)	**426 (8)**	219 (4)	73 (1)	1 (0)	1188 (22)
5	25 (0)	81 (1)	97 (2)	35 (1)	11 (0)	1 (0)	250 (5)
6	7 (0)	16 (0)	20 (0)	9 (0)	2 (0)	0 (0)	54 (1)
Total	663 (12)	1388 (25)	2011 (37)	1080 (20)	309 (6)	30 (1)	5481 (100)

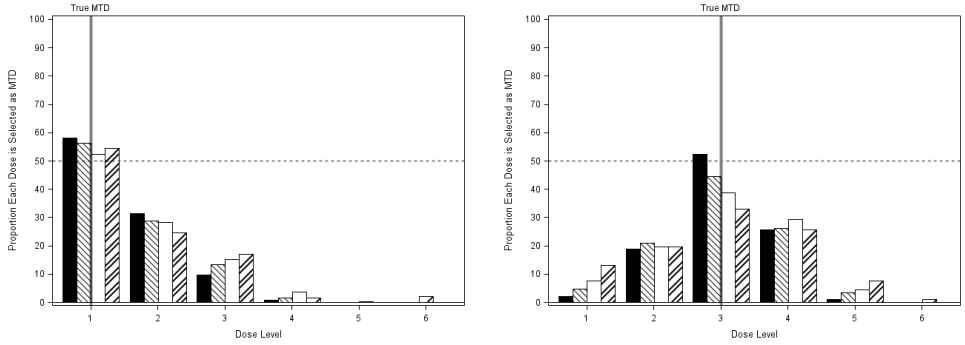
Table 5: Concordance of the Recommended Maximum Tolerated Dose (MTD) in the Homogeneous Sample, N (%N)



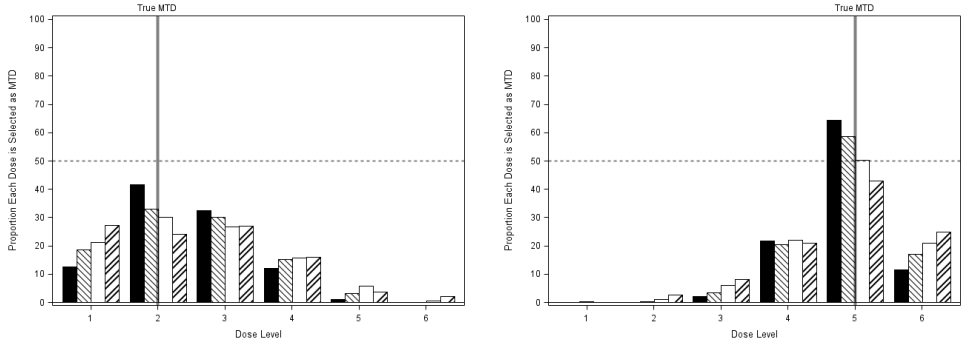
(a) DTR1: Group 1 **(b)** DTR1: Group 2



(c) DTR2: Group 1 **(d)** DTR2: Group 2



(e) DTR3: Group 1 **(f)** DTR3: Group 2



(g) DTR4: Group 1 **(h)** DTR4: Group 2

Figure 1: Proportion Each Dose Level is Recommended as the Maximum Tolerated Dose (MTD) for Varying Group Sizes by Dose-Toxicity Relationship (DTR)
Group Size: One-Sample 19-24 13-18 7-12