Published in final edited form as:
J Biopharm Stat. 2012 ; 22(6): 1206–1219. doi:10.1080/10543406.2011.582972.

Bayesian Continual Reassessment Method for Dose-Finding Trials Infusing T Cells with Limited Sample Size

Yuan Ji[1], **Lei Feng**[1], **Ping Liu**[1], **Elizabeth J. Shpall**[3], **Partow Kebriaei**[3], **Richard Champlin**[3], **Donald Berry**[1], and **Laurence J.N. Cooper**[2]

[1]Department of Biostatistics, U.T. M. D. Anderson Cancer Center, Houston, TX

[2]Division of Pediatrics, U.T. M. D. Anderson Cancer Center, Houston, TX

[3]Stem Cell Transplantation and Cellular Therapy Program, U.T. M. D. Anderson Cancer Center, Houston, TX

Abstract

We consider the design of dose-finding trials for patients with malignancies when only a limited sample size is available. The small sample size may be necessary because 1) the modality of treatment is very expensive, and/or 2) the disease under investigation is rare, requiring a lengthy period to enroll a target patient population. Both of these are common in the field of adoptive immunotherapy, in which T cells are infused to prevent and treat infections and malignancies. The clinical trial described in this paper investigates a novel therapy to adoptively transfer genetically modified T cells in small pilot protocols enrolling patients with B-lineage malignancies. Due to the constraints of cost and infrastructure, the maximum sample size for this trial is fixed at 12 patients distributed among four doses of T cells. Given these limitations, an innovative statistical design has been developed to efficiently evaluate the safety, feasibility, persistence, and toxicity profiles of the trial doses. The proposed statistical design is specifically tailored for trials with small sample sizes in that it uses the toxicity outcomes from patients treated at different doses to make dose-finding decisions. Supplementary materials including an R function and a movie demo can be downloaded in the websites listed in the first two sections of the paper.

Keywords

Adaptive designs; phase I; Toxicity

Introduction

We introduce a Bayesian continual reassessment method (B-CRM) for the statistical design of proof-of-concept Phase I dose-finding trials with small sample sizes. We demonstrate that traditional algorithmic designs such as the 3+3 are not applicable for these trials because of the limited sample size. We advocate the use of model-based methods[1,2,3,4,5] that allow patients' response information at different dose levels to be shared in future decision making. We suggest to examine the performance of various statistical designs using three specific gate-keeping scenarios, and highlight the importance of having extra safety rules in the design to protect patients from exposure to toxic doses.

Correspondence: Dr. Yuan Ji, Department of Biostatistics, U.T. M.D. Anderson Cancer Center, Houston, TX 77030, Phone: (713) 794-4153; Fax: (713) 792-4262, yuanji@mdanderson.org, Dr. Laurence J.N. Cooper, Pediatrics - Research, U.T. M.D. Anderson Cancer Center, 1515 Holcombe Blvd. Unit 907, Houston, TX 77030, Phone: (713) 563-3208; Fax: (713) 792-9832, ljncooper@mdanderson.org.

The B-CRM is applied to the development of a T cell trial at The University of Texas M. D. Anderson Cancer Center for patients with B-lymphoid malignancies undergoing allogeneic umbilical cord blood (UCB) transplantation.[6] A detailed description of the T cell trial can be found at the end of the article. In this trial, one dose of donor-derived T cells that have been genetically modified to express a cluster of differentiation 19 (CD19)-specific chimeric antigen receptor (CAR) will be infused to redirect specificity to CD19 on malignant (and normal) B cells.[7,8,9,10,11] The trial will enroll patients to one of four T cell dose levels, $10^6/m^2$, $10^7/m^2$, $10^8/m^2$, or $10^9/m^2$, based on patient body surface area. These T-cell dose levels were chosen based on preclinical knowledge to allow investigators to not only test safety and feasibility, but to determine a dose of infused T cells that result in sustained *in vivo* persistence.

Traditional 3+3 design is not appropriate for the T cell trial due to the small sample size. For example, if the highest dose level $10^9/m^2$ is the MTD, the 3+3 design would need to treat at least nine patients before it could reach this dose level. A suitable design for this type of trials must be able to escalate quickly and also to control for extreme toxicity. However, one usually has to trade off between fast escalation and control for toxicity. That is, faster escalation often leads to greater chance of toxicity. To this end, we propose a model-based Bayesian extension of the continual reassessment method (therefore the name B-CRM) that borrows strength across different doses in making dose escalation decision. This new design B-CRM is described in detail in the Method section, which is, required by the journal editorial office, placed as the last section of this paper. However, we recommend reading it first before moving onto the next section. An R computer program that implements the proposed design can be downloaded at the website http://odin.mdacc.tmc.edu/~ylji/bcrm.R.

Results

A typical approach to examining the operating characteristics of a Bayesian design for dose-finding trials is to simulate trials many times on a computer according to pre-specified clinical scenarios. Summary statistics, such as the percentage of times a true MTD is selected or the average numbers of patients treated at the MTD, can be used to evaluate the performance of the design. However, in current practice little attention is directed to the construction of clinical scenarios that critically examine the proposed designs in these computer simulations. Scenarios often seem to be selected arbitrarily, which makes the evaluations based on the simulation results hard to interpret and dubious.

We propose three types of clinical scenarios that examine a design's performance to cover diverse and yet practically important situations. In the first type of scenario, all of the doses are excessively toxic. Therefore, no dose should be selected as the MTD and the trial should be terminated quickly. We name this type of scenario the **ES** scenario to represent early stopping. See Table 1 for an example. In the second type of scenario, all of the doses are lower than the MTD. Therefore, appropriate designs should be able to quickly escalate to the highest dose without treating too many patients at lower doses. This type of scenario is specifically important to trials with small sample sizes like the T cell trial here. We name this type of scenario the **FE** scenario to represent fast escalation. See Table 2 for an example. In the third type of scenario, the MTD is bracketed by two adjacent doses, with one dose level much lower than the MTD and the other much higher. Desirable designs should recognize that the higher dose is too toxic and assign most patients to the lower dose. We name this type of scenario the **BR** scenario to indicate that the MTD is bracketed. See Table 3 for an example.

Collectively, these three types of scenarios will tell if a design will 1) stop early when all of the doses are too toxic; 2) escalate quickly when most doses are lower than the MTD; and 3)

stop escalating when higher doses are excessively toxic. The three scenarios will help investigators to filter out any designs that fail to perform well under the critical toxicity configurations, thereby helping to protect patients from being exposed to inferior and toxic doses.

We now demonstrate the performance of the proposed B-CRM under these three types of scenarios in comparison to two well-known designs, the continual reassessment method (CRM) and the 3+3 design. Tables 1-3 contain the operating characteristics of the three designs under the ES, FE, and BR scenarios based on 2,000 computer-simulated trials.

Table 1 presents an ES scenario in which the first dose is already too toxic with a 0.5 probability of toxicity. Due to the early stopping rule (rule 2) in our proposed algorithm, the proposed B-CRM terminates the trial 88% of the time and terminates the trial early when less than 6 patients are enrolled. The CRM with the same stopping rule terminates the trial 82% of the time. We also implemented a version of the CRM without the early stopping rule. This method treats all of the 12 patients at excessively toxic doses and selects the first dose as the MTD over 90% of the time. These results clearly indicate the importance of an early stopping rule to a model-based dose-finding method.

Table 2 demonstrates how important it is for a dose-finding method to be able to escalate quickly under an FE scenario in which the highest dose is still below the MTD (with a toxicity rate of 0.3, for example). Due to the small sample size, the 3+3 design is unable to escalate quickly to the highest dose under this design. On average, less than two patients are treated at the highest dose. The B-CRM and CRM both perform well under the FE scenario. The results in Table 2 indicate that algorithmic designs such as the 3+3 are not suitable for trials with small sample sizes because they do not combine information on different doses to estimate the toxicity rates, and therefore inefficiently use information.

Table 3 examines the performance of the B-CRM and CRM under a BR scenario. Due to the early exclusion rule (rule 3), the B-CRM is able to recognize that dose 2 is too toxic and avoids selecting this dose as the MTD. Because the CRM does not have such a rule, the B-CRM selects dose 2 as the wrong MTD almost 20% fewer times than the CRM.

In addition to these three recommended scenarios, investigators may construct additional scenarios to further examine the operating characteristics of the designs. In our example, we created a total of five scenarios (Figure 1) and compared the proposed B-CRM design with the CRM and 3+3 design under all of them.

We base the simulation on the setup of the aforementioned T cell trial. Specifically, the maximum sample size equals 12. The MTD is defined as the highest dose with a toxicity rate close to 17%. For the CRM, we assume that the prior probabilities of toxicity are 0.05, 0.10, 0.15, and 0.20. We tried several sets of prior probabilities of toxicity for the CRM and presented the one with the best performance. The software for the CRM can be obtained at http://biostatistics.mdanderson.org/SoftwareDownload/.

Examining the results summarized in Figure 2, we list the following observations:

1. Due to its inferior performance in Scenarios 1 and 4, the 3+3 design should not be recommended for this type of trial. In Scenario 1, the 3+3 design identifies dose 2 as the estimated MTD 14% of the time, compared to 49% by the CRM and 53% by the proposed B-CRM. In Scenario 4, the 3+3 design cannot escalate quickly due to its fixed cohort size of 3. Therefore, with a maximum sample size of 12 patients, the 3+3 design cannot determine if the highest dose is an estimated MTD.

2. The CRM performs reasonably well in Scenarios 1-3. Its selection percentages of the MTD in Scenarios 4 and 5 are lower than those of the proposed B-CRM. Specifically, in Scenario 4, the CRM selects dose 4 as the estimated MTD 49% of the time, compared to 69% by the B-CRM. In Scenario 5, the CRM selects dose 1 as the estimated MTD 67% of the time, compared to 85% by the B-CRM.

3. The proposed B-CRM performs well in all of the scenarios and is the best of the three methods under comparison. It is recommended for designing the T cell phase I trial.

Figure 3 further demonstrates how the proposed B-CRM estimates the dose-response curve based on the observed toxicity data across all of the dose levels. The six plots present the toxicity responses for each of the six cohorts during the course of a computer simulated trial with 12 patients. Each plot also contains an estimated dose-response curve based on the data seen within the plot. An animated documentation to further illustrate these results can be downloaded at: http://odin.mdacc.tmc.edu/~ylji/BCRM_demo.wmv.

Discussion

Statistical remarks

The challenges presented by dose-finding trials such as a T cell infusion trial necessitate careful statistical modeling and designs. We have presented a Bayesian design in line with the reassessment thinking of the CRM as a desirable solution for dose-finding problems with small sample sizes. Our proposed design is favorable over algorithmic methods such as the 3+3 design, as well as model-based approaches such as the CRM.

We have also proposed three types of clinical scenarios (ES, FE, and BR) that should be used to critically evaluate a dose-finding method in computer simulations. These three scenarios require reasonable methods for stopping the trial early when all of the doses are too toxic, escalating quickly when all of the doses are below the MTD, and excluding toxic doses during the course of the trial. For trials with a small sample size, these three properties are particularly desirable and important for any dose-finding method to perform well. For example, in our simulation the 3+3 design could not escalate quickly in an FE scenario, thus making it inapplicable to our T cell trial. Although one may reduce the cohort size (e.g., by inventing a 2+2 design) to achieve fast escalation, it seems difficult to determine a set of algorithmic dose-escalation rules with only two patients per cohort. Model-based methods are more efficient at using the information in the toxicity outcomes of patients treated at different doses, and therefore should be recommended for trials with a small number of patients.

An alternative approach is to design a phase I/II type of trial[12,13] in which the toxicity outcome and a certain type of activity outcome (such as some anti-tumor measure) are monitored simultaneously during the dose finding. However, special effort is needed to modify the statistical models of these approaches so that they will perform well with small sample sizes. This is one of our future research projects.

It is important to monitor the toxicity of each T cell dose level in trials infusing T cells that are numerically expanded *ex vivo* to achieve clinically-sufficient numbers. This is the case when genetically modified T cells are derived from UCB, as only a limited number of mononuclear cells are available from the donor for the generation of CAR^+ T cells without compromising hematopoietic engraftment in the patient (Figure 4).

Medical remarks

Adoptive immunotherapy with genetically modified T cells can provide a beneficial anti-tumor effect in clinical trials.[14,15,16] One approach to improving *in vivo* persistence is to adoptively transfer large numbers of T cells to the patient. This is founded on the premise that the more T cells present in the inoculum, the greater the chances of infusing a subset of cells with the potential to survive *in vivo*, and therefore the greater chance of a sustained anti-tumor effect.[17] The trade-off is that a larger number of clinical-grade cells requires the time and expense of prolonged tissue culture, and the very act of culturing may render the propagated T cells vulnerable to replicative senescence[18,19,20] and differentiate the cells into terminal effectors with a limited capacity for self-renewal despite engaging antigen via the introduced CAR. Since these genetically modified T cells also express a functioning endogenous αβTCR, an important safety consideration is whether the infused T cells will lead to graft-versus-host-disease and/or aplasia. Therefore, a desirable dose level of T cells to be infused should lead to prolonged persistence without causing an exacerbation of existing toxicities or the development of new toxicities.

It is anticipated that achieving and sustaining a therapeutic anti-tumor effect will depend on the dose of genetically modified T cells to be infused. Thus, the hypothesis to be tested is whether infusing more T cells results in a beneficial effect. However, since time in tissue culture is needed to propagate a larger number of UCB-derived T cells, more may not necessarily be better, and the following variables need to be considered.

(i) The cost of tissue culture in the good manufacturing practice facility— Prolonged *in vitro* propagation consumes resources and is expensive in terms of reagents, salaries, and the use of space within a facility operating in compliance with current good manufacturing practice. Prolonged culture time also increases the possibility of inadvertent contamination of the product. In addition, the Food and Drug Administration is expected to require that early phase trials observe a monitoring period after each T-cell infusion and before a subsequent patient receives cells. Thus, long culturing periods further prolong the period of enrollment since simultaneous infusions are not currently considered safe.

(ii) The potential for deleterious alloreactivity—Cryopreserved allogeneic UCB contains (in addition to hematopoietic stem cells, HSCs) donor-derived T cells expressing an endogenous αβ T cell receptor (TCR). While the specificity of the TCR is unknown, there is a chance that these donor-derived T cells will participate in undesired alloreactivity caused by the TCR recognizing haplo-human lymphocyte antigen manifested by graft-versus-host-disease.

(iii) The potential for the T cells to differentiate into effector cells with limited in vivo persistence—Most of the thawed UCB unit is infused to ensure hematopoietic engraftment in the patient, and thus only a limited amount of neonatal blood is available for *ex vivo* manipulation such as the gene transfer of T cells. Thus, the UCB-derived T cells will need to be numerically expanded in tissue culture to achieve a clinically-significant dose. However, repetitive stimulation of these naïve cells to enter into proliferative cycles may cause them to differentiate into effector T cells with a potentially limited capacity for continued persistence *in vivo*. Yet it is the naïve or memory pools of T cells that are expected to persist after adoptive transfer.

(iv) The percentage expression of CAR—To generate large numbers of genetically modified T cells from limited numbers of UCB mononuclear cells, we have developed an artificial antigen presenting cell that triggers and sustains T-cell proliferation through an introduced CAR.[21,22] Over time, we observed an increase in the percentage of T cells

expressing the CAR. Thus, the T cells that are infused early in the culturing process may express lower levels of CAR than the T cells that are infused after prolonged *in vitro* selection. This may impact the therapeutic effect. In the initial trail, we anticipate infusing genetically modified T cells about 100 days after infusion of HSCs, at a time point at which recipients of UCB transplantation have typically recovered their circulating $CD19^+$ B cells.[23] Thus, the infused CAR^+ T cells may undergo a selective survival advantage and continue to undergo *in vivo* selection if the CAR recognizes an antigen, such as CD19, after infusion.

The B-CRM for the trial takes into account the dosing and timing of T cells so that the number of T cells can be used to detect the long-term persistence of the infused cells while avoiding excessive toxicity.

Method

In the T cell dose-finding trial, patients will be assigned to four doses of T cells: dose level I ($10^6/m^2$), dose level II ($10^7/m^2$), dose level III ($10^8/m^2$), and dose level IV ($10^9/m^2$). Patients met eligibility criteria for the trial will be assigned to a dose level, and will be based on the following statistical design.

The total sample size is 12 patients and the cohort size is 2. A cohort size of 3 is deemed too large because we only have 12 patients and need to evaluate up to four doses.

Denote $p(d)$ as the probability of toxicity at dose level d. We assume that

$$p(d)=\frac{e^{int+\alpha\times d}}{1+e^{int+\alpha\times d}},$$

where int is a fixed known constant and α is an unknown parameter. For the T cell trial, we take $int = -10$, which implies that at dose $d = 0$, the probability of toxicity equals $e^{-10}/(1+e^{-10}) \approx 0$. In general, the value of int is determined based on the prior information of toxicity at dose 0 and by calibration with respect to the operating characteristics of the design. For this trial, we found that $int = -10$ gave the best performance among the values we tried ($int = -1, -3, -10$, and -20). The parameter α is assumed to be positive and follows a prior unit exponential distribution. Therefore, the probability of toxicity is assumed to increase with the dose level. We $\log_{10}$ transformed the four dose levels to obtain the transformed dose levels $d = 6, 7, 8$, and 9, respectively. Since the prior mean of α equals 1, the prior toxicity rates at the four dose levels are 0.02, 0.05, 0.12, and 0.27, respectively, computed by plugging $int = -10$ and $\alpha = 1$ into the formula for $p(d)$. These rates were considered to be reasonable estimates based on our clinical experience. For other trials, one can rescale the dose levels so that desirable prior toxicity estimates can be obtained based on the prior mean of α and the dose-response curve $p(d)$. Note that the above rates only represent the prior mean estimates. The uncertainty about these estimates is reflected in the prior variance associated with α. For example, the prior standard deviations for $p(1) - p(4)$ are 0.35, 0.39, 0.41, 0.43, respectively.

Denote d_1, d_2, d_3, and d_4 as the four transformed dose levels. Define the maximum tolerated dose (MTD) as the highest dose at which the probability of toxicity is close to p_T, a target probability of toxicity (here, $p_T = 0.17$). For $i = 1,\ldots,4$, the trial data consist of $x(d_i)$, the number of observed toxicities at dose d_i, and $n(d_i)$, the number of patients treated at dose d_i. For each patient, the toxicity outcome is either 0 (no toxicity) or 1 (toxicity). The likelihood function is defined as the product of the binomial probability mass functions over all the dose levels and all the patients treated.

With the definition of the likelihood and the unit exponential prior for α, we can estimate the posterior distribution of α via a random walk Markov chain using a Metropolis-Hastings proposal. Statistical inference can then be based on the posterior samples of α. We used the R software for computation and the R program is available at http://odin.mdacc.tmc.edu/~yuanj/bcrm.R/.

We propose a Bayesian dose-finding algorithm as follows.

1. *The first cohort is assigned to dose level I (10^6/m^2). No untried dose will be skipped in the dose escalation.*
2. ***[Early Stopping]*** *If at any point during the trial* $\Pr(p(d_1) > p_T \mid data) > 0.9$, *the trial is terminated and no dose is selected as the MTD.*
3. ***[Early Exclusion]*** *If at any point during the trial* $\Pr(p(d_i) > p_T \mid data) > 0.9$ *for* $i = 1, \ldots, 4$, *dose levels* d_i *and higher will never again be used to treat patients in the trial. The highest dose level available for the remainder of the trial will be* d_{i-1}.
4. *Subject to rules 1, 2, and 3, the next cohort is treated at dose j with the largest posterior probability*

$$\Pr\left[abs\left(p(d_j) - p_T\right) = \min_i \left(abs(p(d_i) - p_T)\right) \middle| data \right]$$

until the maximum sample size is reached (Here abs() represent the absolute value)

The algorithm consists of four decision rules. Rule 1 is standard to most dose-finding trials. Rule 2, the early stopping rule, is a safety rule that terminates the trial when the first dose is shown to be too toxic. Specifically, the rule states that when there is a high posterior probability that the first dose's toxicity rate is higher than the target rate p_T, the trial will be terminated and no dose will be selected. Rule 3 is another safety rule that prevents treating patients at toxic doses. The rule states that when there is a large posterior probability that any dose's toxicity rate is higher than the target rate p_T, that dose and all of the higher level doses will be excluded from the remainder of the dosefinding trial. Rule 4 is the core rule of our proposed algorithm and sequentially determines the most appropriate dose for the next cohort of patients. This rule compares all of the doses and selects the one closest to the MTD in terms of the absolute difference in their probabilities of toxicity. Note that unlike most decision rules in the literature,[1,2] rule 4 is based on posterior probabilities, and therefore automatically accounts for the variability in the estimation of the dose toxicity rates. In addition, the posterior probability in rule 4 compares the doses internally rather than comparing them to a specific threshold. This type of internal comparison has been shown in the literature[24] to perform well.

Note that the single-parameter logistic model specified as $p(d)$ has already been discussed in the literature, e.g., by Shen and O'Quigley (1996). The contribution of this paper is on 1) a simple Bayesian implemention and calibration procedure that helps design small-size trials under the model, 2) a full probability-based decision rule with complementary and important safty rules, and 3) the introduction of a set of scenarios that can used to fully evaluate a given dose-finding approach for various properties.

Acknowledgments

Support from Cancer Center Core Grant (CA16672); RO1 (CA132897, CA124782, CA120956); R21 (CA129390, CA116127); DOD PR064229; The Alliance for Cancer Gene Therapy; The Alex Lemonade Stand Foundation; The Carl C. Anderson, Sr. and Marie Jo Anderson Charitable Foundation; The Gillson Longenbaugh Foundation; The

J.P. McCarthy Fund Developmental Grant Program; The Leukemia and Lymphoma Society; The Lymphoma Research Foundation; The Miller Foundation; The National Foundation for Cancer Research; The National Marrow Donor Program; The Pediatric Cancer Research Foundation.

References

1. Berry DA. A guide to drug discovery: Bayesian clinical trials. Nature Reviews Drug Discovery. 2006; 5:27–36.
2. O'Quigley J, Pepe M, Fisher L. Continual reassessment method: A practical design for Phase I clinical trials in cancer. Biometrics. 1990; 46(1):33–48. [PubMed: 2350571]
3. Ji Y, Li Y, Bekele BN. Dose-finding in oncology clinical trials based on toxicity probability intervals. Clinical Trials. 2007 Jun.4:235–44. [PubMed: 17715248]
4. Cheung YK, Chappel R. Sequential designs for phase I clinical trials with late-onset toxicities. Biometrics. 2000 Sep; 56(4):1177–82. [PubMed: 11129476]
5. Rosenberger W, Haines L. Competing designs for phase I clinical trials: a review. Stat Med. 2002; 21(18):2757–70. [PubMed: 12228889]
6. Serrano LM, Pfeiffer T, Olivares S, Numbenjapon T, Bennitt J, Kim D, Smith D, McNamara G, Al-Kadhimi Z, Rosenthal J, Forman SJ, Jensen MC, Cooper LJ. Differentiation of naive cord-blood T cells into CD19-specific cytolytic effectors for posttransplantation adoptive immunotherapy. Blood. 2006 Apr 1; 107(7):2643–52. [PubMed: 16352804]
7. Cooper LJ, Topp MS, Serrano LM, Gonzalez S, Chang WC, Naranjo A, Wright C, Popplewell L, Raubitschek A, Forman SJ, Jensen MC. T-cell clones can be rendered specific for CD19: toward the selective augmentation of the graft-versus-B-lineage leukemia effect. Blood. 2003 Feb 15; 101(4): 1637–44. [PubMed: 12393484]
8. Brentjens RJ, Santos E, Nikhamin Y, Yeh R, Matsushita M, La Perle K, Quintás-Cardama A, Larson SM, Sadelain M. Genetically targeted T cells eradicate systemic acute lymphoblastic leukemia xenografts. Clin Cancer Res. 2007 Sep 15; 13(18 Pt 1):5426–35. [PubMed: 17855649]
9. Rössig C, Pscherer S, Landmeier S, Altvater B, Jürgens H, Vormoor J. Adoptive cellular immunotherapy with CD19-specific T cells. Klin Padiatr. 2005 Nov-Dec;217(6):351–6. [PubMed: 16307422]
10. Cheadle EJ, Gilham DE, Thistlethwaite FC, Radford JA, Hawkins RE. Killing of non-Hodgkin lymphoma cells by autologous CD19 engineered T cells. Br J Haematol. 2005 May; 129(3):322–32. [PubMed: 15842655]
11. Cooper LJ, Al-Kadhimi Z, DiGiusto D, Kalos M, Colcher D, Raubitschek A, Forman SJ, Jensen MC. Development and application of CD19-specific T cells for adoptive immunotherapy of B cell malignancies. Blood Cells Mol Dis. 2004 Jul-Aug;33(1):83–9. [PubMed: 15223016]
12. Yin G, Li Y, Ji Y. Bayesian dose-finding in phase I/II clinical trials using toxicity and efficacy odds ratio. Biometrics. 2006 Sep; 62(3):777–84. [PubMed: 16984320]
13. Thall PF, Cook JD. Dose-finding based on efficacy-toxicity trade-offs. Biometrics. 2004 Sep. 60:684–93. [PubMed: 15339291]
14. Rosenberg SA, Restifo NP, Yang JC, Morgan RA, Dudley ME. Adoptive cell transfer: a clinical path to effective cancer immunotherapy. Nat Rev Cancer. 2008 Apr; 8(4):299–308. [PubMed: 18354418]
15. Morgan RA, Dudley ME, Wunderlich JR, Hughes MS, Yang JC, Sherry RM, Royal RE, Topalian SL, Kammula US, Restifo NP, Zheng Z, Nahvi A, de Vries CR, Rogers-Freezer LJ, Mavroukakis SA, Rosenberg SA. Cancer regression in patients after transfer of genetically engineered lymphocytes. Science. 2006 Oct 6; 314(5796):126–9. [PubMed: 16946036]
16. June CH. Adoptive T cell therapy for cancer in the clinic. J Clin Invest. 2007 Jun; 117(6):1466–76. [PubMed: 17549249]
17. Powell DJ Jr, Dudley ME, Robbins PF, Rosenberg SA. Transition of late-stage effector T cells to CD27+ CD28+ tumor-reactive effector memory T cells in humans after adoptive cell transfer therapy. Blood. 2005 Jan 1; 105(1):241–50. [PubMed: 15345595]
18. Pawelec G, Rehbein A, Haehnel K, Merl A, Adibzadeh M. Human T-cell clones in long-term culture as a model of immunosenescence. Immunol Rev. 1997 Dec.160:31–42. [PubMed: 9476663]

19. Mazzatti DJ, White A, Forsey RJ, Powell JR, Pawelec G. Gene expression changes in long-term culture of T-cell clones: genomic effects of chronic antigenic stress in aging and immunosenescence. Aging Cell. 2007 Apr; 6(2):155–63. [PubMed: 17286612]

20. Zhou J, Shen X, Huang J, Hodes RJ, Rosenberg SA, Robbins PF. Telomere length of transferred lymphocytes correlates with in vivo persistence and tumor regression in melanoma patients receiving cell transfer therapy. J Immunol. 2005 Nov 15; 175(10):7046–52. [PubMed: 16272366]

21. Numbenjapon T, Serrano LM, Chang WC, Forman SJ, Jensen MC, Cooper LJ. Antigen-independent and antigen-dependent methods to numerically expand CD19-specific CD8+ T cells. Exp Hematol. 2007 Jul; 35(7):1083–90. [PubMed: 17588477]

22. Numbenjapon T, Serrano LM, Singh H, Kowolik CM, Olivares S, Gonzalez N, Chang WC, Forman SJ, Jensen MC, Cooper LJ. Characterization of an artificial antigen-presenting cell to propagate cytolytic CD19-specific T cells. Leukemia. 2006 Oct; 20(10):1889–92. [PubMed: 17041638]

23. Komanduri KV, St John LS, de Lima M, McMannis J, Rosinski S, McNiece I, Bryan SG, Kaur I, Martin S, Wieder ED, Worth L, Cooper LJ, Petropoulos D, Molldrem JJ, Champlin RE, Shpall EJ. Delayed immune reconstitution after cord blood transplantation is characterized by impaired thymopoiesis and late memory T-cell skewing. Blood. 2007 Dec 15; 110(13):4543–51. [PubMed: 17671230]

24. Huang X, Biswas S, Oki Y, Issa JP, Berry DA. A parallel phase I/II clinical trial design for combination therapies. Biometrics. 2007 Jun; 63(2):429–36. [PubMed: 17688495]

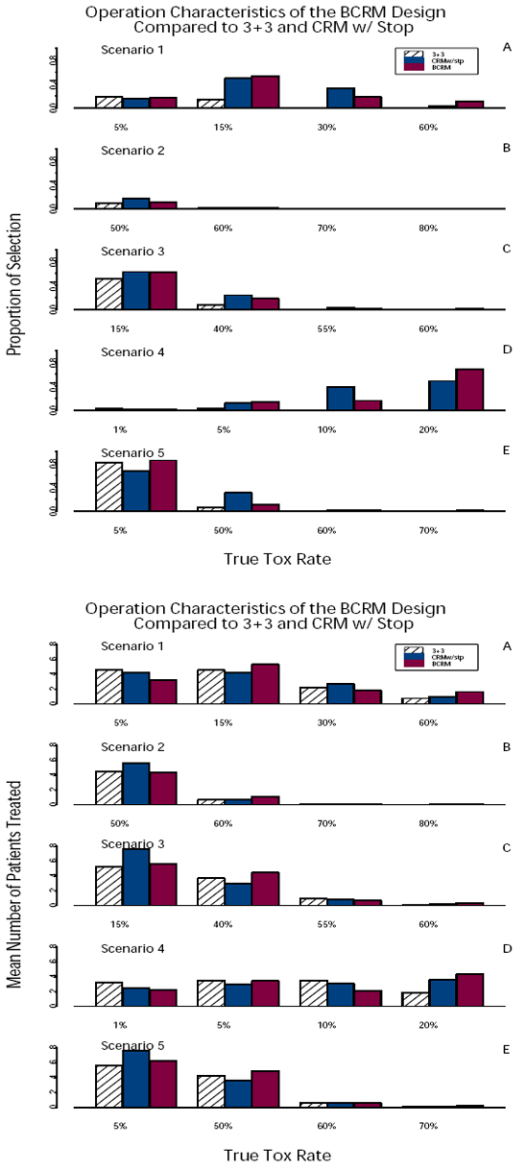

Figure 1.
Simulation results comparing the B-CRM, the CRM (with early stopping), and the 3+3 design based on the T cell dose-finding trial. The maximum sample size is 12 and the target toxicity rate is 17%.

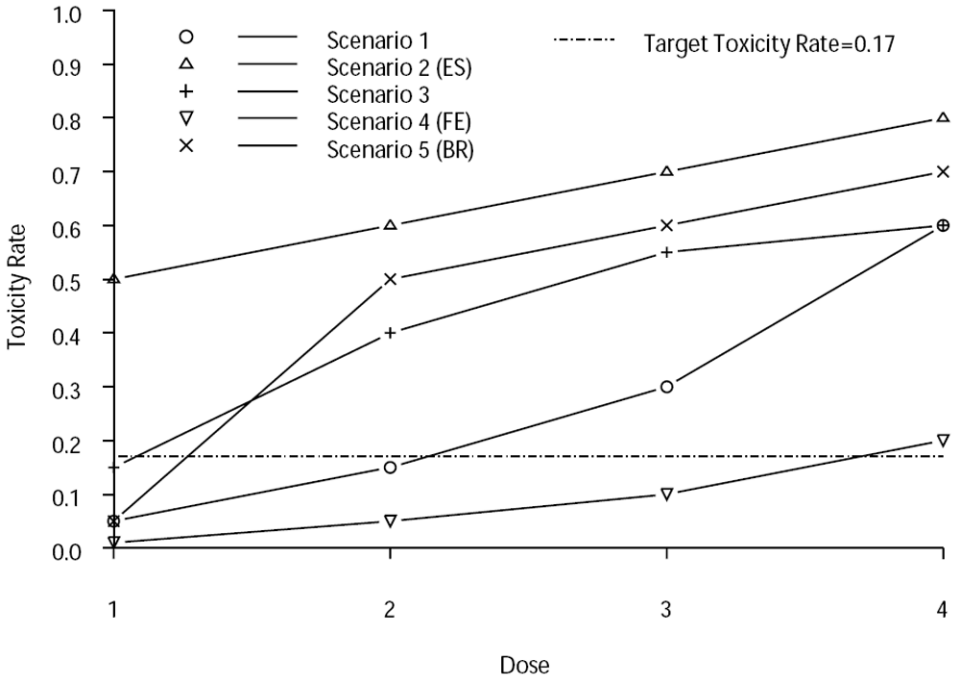

Figure 2.
Dose-response curves of the five scenarios in the simulation study.

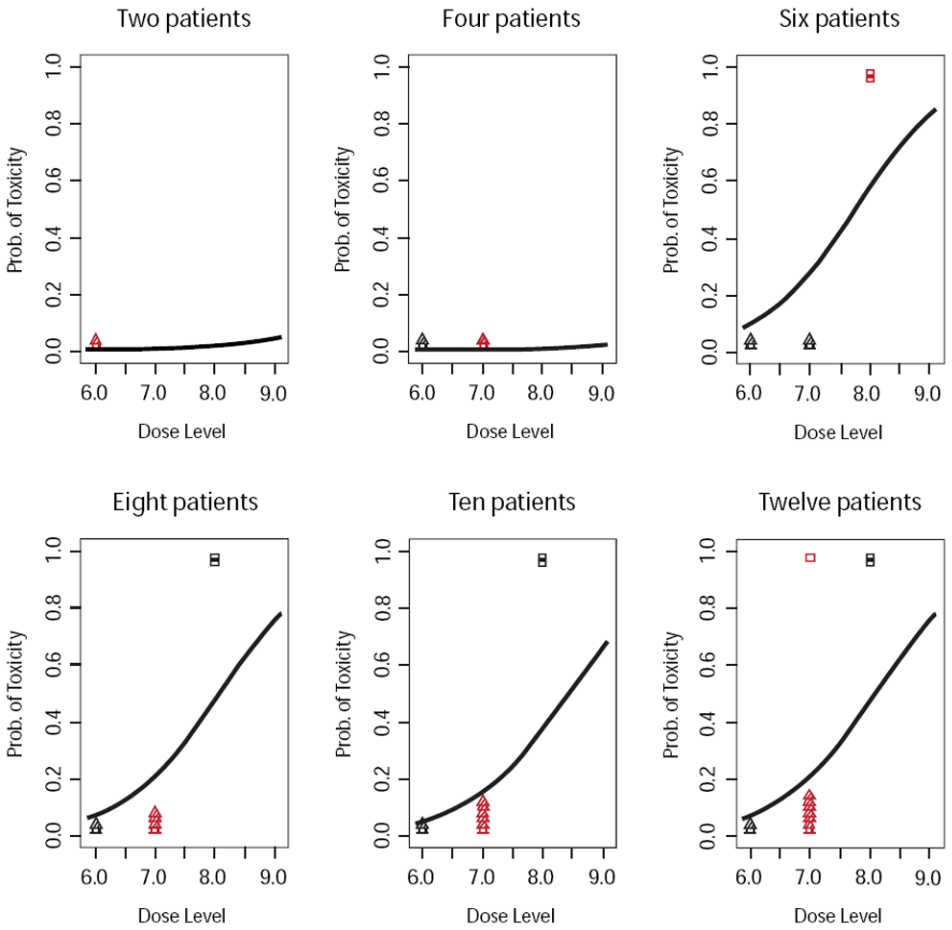

Figure 3.

A computer-simulated trial example with estimated dose-response curves by the B-CRM. The estimated curves were obtained by plugging the posterior mean of α into the formula for $p(d)$. In this simulation, the four dose levels were assumed to have toxicity rates of 0.05, 0.15, 0.50, and 0.60, respectively. A square indicates a dose limiting toxicity (DLT) event and a triangle indicates no DLT. The red color implies that the corresponding dose level is used to treat patients during the trial. The first cohort of two patients treated at dose level 6.0 did not have DLT, and therefore the second cohort was treated at dose level 7.0. No DLT was observed for the second cohort as well. The third cohort was treated at dose level 8.0, where both patients had DLTs. Then the fourth cohort was treated back at dose level 7.0 and no DLT was observed. Because both patients treated at dose level 8.0 had DLTs, the fifth and sixth cohorts of patients were still treated at dose level 7.0 instead of 8.0 (due to the Early Exclusion rule). In the end, one of eight patients treated at dose level 7.0 had a DLT, and 7.0 was selected as the MTD.

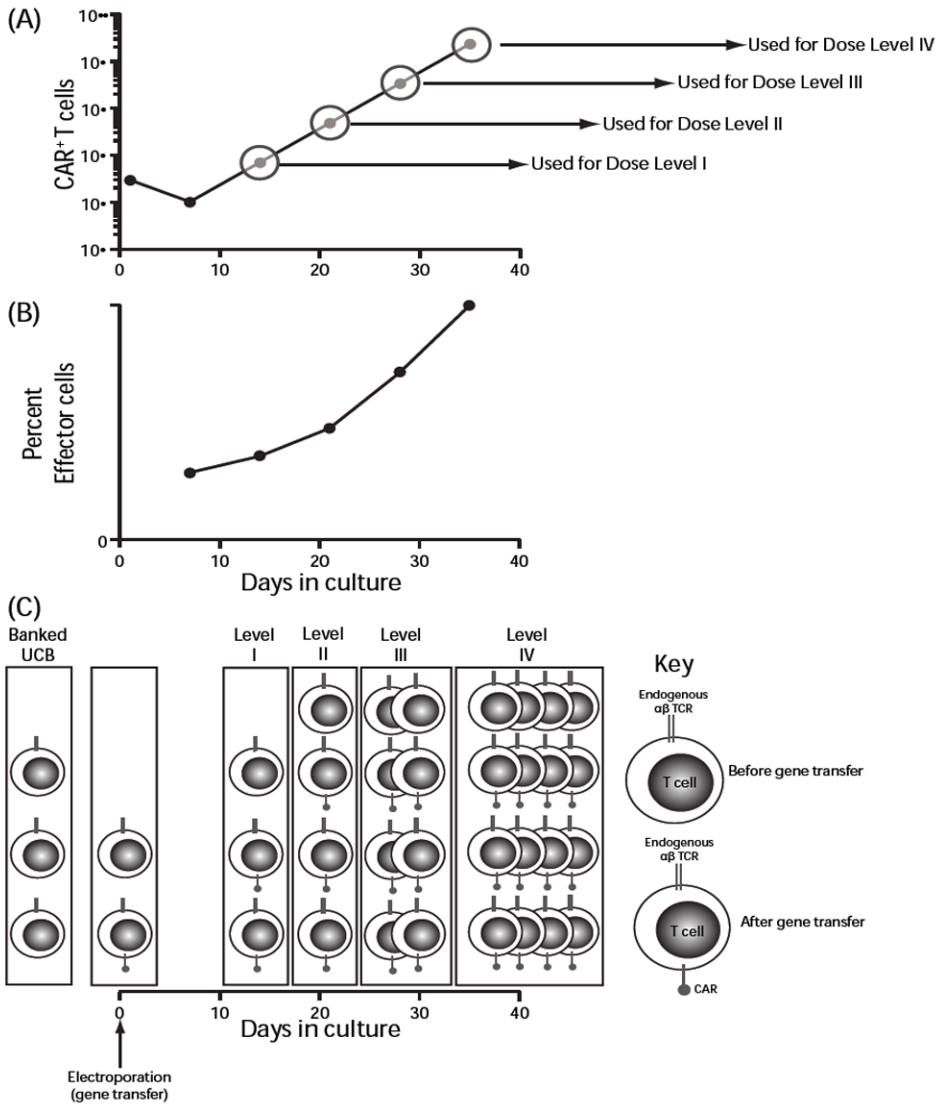

Figure 4.

Schematic of culturing clinical-grade T cells derived from umbilical cord blood.

(A) Hypothetical propagation of UCB-derived CD19-specific CAR^+ T cells after electrotransfer DNA into 10^7 total nucleated cells on day 0 of tissue culture, assuming a 7-fold expansion every 7 days beginning on day 7. On average we observe approximately 30% CAR expression on day 1, and the loss of CAR^+ T cells through day 7 reflects that not all of these initially CAR^+ T cells have integrated transgene. We estimate that electroporating 10^7 T cells (e.g. taken from ~1.43 mL of a 25 mL standard UCB unit cryopreserved at MDACC, with post-process banked UCB unit $CD3^+$ T-cell count of ~7×10^6/mL) will numerically expand to >2×10^{10} cells after 28 to 35 days of continuous co-culture on $CD19^+$ artificial antigen presenting cells added every 7 days. The use of ~1.5 mL of banked UCB for the manufacture of T cells is not predicted to harm hematopoietic engraftment, as we will require that the total (if necessary, combining up to two units) amount of UCB infused be ≥3×10^7 TNC/kg recipient weight, approximately equivalent to >0.8 mL/kg (assuming 3.8×10^7 TNC/mL UCB banked at MDACC) of the total post-processed UCB unit(s). Note that T cells are dosed based on total nuclear cell count and not percentage of CAR^+ T cells.

(B) Hypothetical graph showing the differentiation of T cells into effector cells over time and accompanying loss of naïve/memory cell phenotype.
(C) Schematic of the cell surface phenotype of genetically modified T cells over time.

Table 1

Simulation results under an early-stopping ES scenario for the B-CRM and the CRM.

Method		Dose 1	Dose 2	Dose 3	Dose 4	None Selected	Total # of Patients Treated	Overall Toxicity rate
Scenario	True Toxicity rate	0.50	0.60	0.70	0.80			
3+3	Pobability(selection)	0.09	0.00	0.00	0.00	**0.89**	5.20	0.51
	# of patients treated	5.55	4.21	0.53	0.02			
CRM without stop	Probability(selection)	0.94	0.05	0.01	0.00	**0.01**	12.00	0.52
	# of patients treated	10.56	1.27	0.16	0.01			
CRM with stop	Probability(selection)	0.17	0.01	0.00	0.00	**0.82**	6.37	0.52
	# of patients treated	5.61	0.64	0.11	0.01			
	# of patients treated	5.55	4.21	0.53	0.02			
B-CRM	Probability(selection)	0.11	0.01	0.00	0.00	**0.89**	5.63	0.53
	# of patients treated	4.39	1.14	0.09	0.02			

Table 2

Simulation results under a fast escalation (FE) scenario for the 3+3, CRM, and B-CRM.

Method		Dose 1	Dose 2	Dose 3	Dose 4	None Selected	Total # of Patients Treated	Overall Toxicity rate
Scenario	True Toxicity rate	0.01	0.05	0.10	0.20			
3+3	Probability(selection)	0.01	0.04	0.06	**0.00**	0.89	11.93	0.08
	# of patients treated	3.11	3.47	3.48	**1.86**			
CRM	Probability(selection)	0.04	0.15	0.26	**0.55**	0.00	11.97	0.09
	# of patients treated	2.61	3.23	3.23	**2.93**			
B-CRM	Probability(selection)	0.02	0.14	0.16	**0.69**	0.00	12.00	0.11
	# of patients treated	2.25	3.31	2.12	**4.33**			

Table 3

Simulation results under a bracket (BR) scenario for the 3+3, CRM, and B-CRM.

Method		Dose 1	Dose 2	Dose 3	Dose 4	None Selected	Total # of Patients Treated	Overall Toxicity rate
Scenario	True Toxicity rate	0.05	0.50	0.60	0.70			
3+3	Probability(selection)	**0.81**	0.06	0.01	0.00	0.03	10.31	0.27
	# of patients treated	**5.55**	4.21	0.53	0.02			
CRM	Probability(selection)	**0.67**	0.31	0.01	0.00	0.01	11.95	0.22
	# of patients treated	**7.59**	3.62	0.64	0.10			
B-CRM	Probability(selection)	**0.83**	0.12	0.01	0.02	0.02	11.85	0.27
	# of patients treated	**6.28**	4.82	0.55	0.20			