Published in final edited form as:

Clin Trials. 2020 April ; 17(2): 157–165. doi:10.1177/1740774519890145.

Adaptive dose finding based on safety and feasibility in early-phase clinical trials of adoptive cell immunotherapy

Nolan A Wages¹, Camilo E Fadul²

¹Division of Translational Research & Applied Statistics, University of Virginia, Charlottesville, VA, USA

²Division of Neuro-oncology, Department of Neurology, University of Virginia, Charlottesville, VA, USA

Abstract

Background/aims—Dose feasibility is a challenge that may arise in the development of adoptive T cell therapies for cancer. In early-phase clinical trials, dose is quantified either by a fixed or per unit body weight number of cells infused. It may not be feasible, however, to administer a patient's assigned dose due to an insufficient number of cells harvested or functional heterogeneity of the product. The study objective becomes to identify the maximum tolerated dose with high feasibility of being administered. This paper describes a new dose-finding method that adaptively accounts for safety and feasibility endpoints in guiding dose allocation.

Methods—We propose an adaptive dose-finding method that integrates accumulating feasibility and safety data to select doses for participant cohorts in early-phase trials examining adoptive cell immunotherapy. We sequentially model the probability of dose-limiting toxicity and the probability of feasibility using independent beta-binomial models. The probability model for toxicity borrows information across all dose levels using isotonic regression, allowing participants, infused at a lower dose than his or her planned dose, to contribute safety data to the dose-finding algorithm. We applied the proposed methodology in a single simulated trial and evaluated its operating characteristics through extensive simulation studies.

Results—In simulations conducted for a phase I study of adoptive immunotherapy for newly diagnosed glioblastoma, the proposed method demonstrates the ability to identify accurately the feasible maximum tolerated doses and to treat participants at and around these doses. Over ten hypothesized scenarios studied, the percentage of correctly selecting the true feasible and maximum tolerated dose ranged from 50% to 90% with sample sizes averaging between 21 and 24 participants. A comparison to the only known existing method accounting for safety and feasibility yields competitive performance.

Conclusion—We have developed a new practical adaptive dose-finding method to assess feasibility in early-phase adoptive cell therapy trials. A design that incorporates feasibility, as a function of the quantity and quality of the product manufactured, in addition to safety will have an impact on the recommended phase II doses in studies that evaluate patient outcomes.

Keywords

Dose finding; phase I; safety; feasibility; immunotherapy; adoptive T cell therapy

Introduction

A current challenge in phase I trials of cancer treatment is determining a recommended phase II dose for further testing of adoptive cell immunotherapies. In these trials, the number of cells examined relies on either a fixed or a per body weight dose. It may not be feasible to administer a patient's assigned dose due to an insufficient number of cells harvested or functional heterogeneity of the product. This makes it difficult to identify a feasible and safe dose to carry forward into later phases of development, hindering the evaluation of the therapeutic intervention. Historically, the primary objective of phase I trials has been to identify the maximum tolerated dose among a set of dose levels. The maximum tolerated dose is defined as the highest dose that can be administered to patients with an acceptable level of toxicity, and it is often the recommended phase II dose under the assumption that higher doses are likely to be more effective. Accepted dose finding designs for clinical trials assume that each patient is able to receive his or her intended dose based on the dose-finding algorithm. In adoptive cell therapy, however, the number of cells grown may be below the assigned infusion dose identified by the dose-finding design for a given participant. This raises questions of how to assess the feasibility of the approach, while also evaluating safety. Integrating accumulated feasibility and safety data to select sequential doses for participant cohorts during dose-finding trials remains a challenge.

Typically, data observed in patients deemed not feasible for his or her assigned dose are not factored into determining the maximum tolerated dose. In a phase I trial of chimeric antigen receptor T cells, Lee et al.¹ states, "Patients whose chimeric antigen receptor T cell product did not meet the dose to which they were assigned did not inform dose escalation." Ignoring safety information on patients treated away from their assigned dose could lead to biased estimates of the maximum tolerated dose, inflate the overall sample size, and increase the cost of conducting the study. A model-based method that adaptively borrows information across all available dose levels could use the toxicity data from each participant accrued to the study, regardless of which dose level he or she received because of feasibility. If a participant must receive a lower dose than his or her recommended dose, the observed dose-limiting toxicity data at the lower dose could contribute to the sequential updating of a dose-toxicity model and recommended planned dose for the next accrual.

Our interest on this subject arose when designing a phase I study for newly diagnosed glioblastoma using epidermal growth factor receptor bispecific antibody armed activated T cells in combination with radiation and chemotherapy. The primary objective of the trial was to determine the feasible maximum tolerated dose, defined by acceptable toxicity and high feasibility, by examining four dose levels of epidermal growth factor receptor bispecific antibody armed activated T cell infusions. Dose levels consist of infusions of activated T cells, with each level defined by the number of cells received by the patient; i.e., eight infusions of 5, 10, 20, or 40×10^9 activated T cells per infusion for a total dose of 40, 80,

160 and 320×10^9 cells, respectively (Table 1). To address this issue, we herein propose a new phase I design for implementation into dose-finding studies with dose-feasibility constraints. To our knowledge, the only existing dose-finding method that accounts for safety and feasibility is an extension of the continual reassessment method² proposed by Thall, Sung, and Choudhury³ and we compare our proposed method to this approach in simulation studies.

Methods

We propose the following dose-finding design with the objective of determining the feasible maximum tolerated dose from pre-specified set of dose levels $d_1 < d_2 < \dots < d_I$. The primary endpoints that guide accrual decisions include frequency of protocol-specific, treatment-related dose-limiting toxicities and feasibility of generating enough cells to administer a planned dose. The feasible maximum tolerated dose is defined as the dose with dose-limiting toxicity rate closest to the target rate of π^* and achieves a threshold of at least θ^* feasibility.

Feasibility

Prior to assignment of a recommended dose, each eligible participant has leukapheresis obtain an enriched population of peripheral blood T cells needed to prepare the immunotherapy. Lymphocytes are cultured and expanded in the laboratory for 14 days, and then are harvested and armed with a bispecific antibody. The epidermal growth factor receptor bispecific antibody armed activated T cells are released for clinical use after Quality Control testing and counted in order to assess the feasibility of administering each dose to the eligible participant. If the number of armed T cells obtained for an eligible participant is at least as high as 80% of the number of cells per patient for a dose level, then that dose is considered feasible for that participant. If the number of armed T cells harvested fails to reach the threshold associated with a dose level, then that dose is considered infeasible for that patient. For each eligible participant the feasibility (yes/no) of administering each dose level is evaluated and recorded before assignment of a dose. Some patients may have to receive doses below that of the design recommendation, or may not be able to receive any of the available doses. Let X_k denote the number of cells grown for a participant k and denote the number of study dose levels by I . The most relevant information is the relationship between X_k and the study dose levels $d_1 < d_2 < \dots < d_I$ evaluated in order to determine which dose levels it is feasible for each participant to receive. The event $[X_k \geq d_j]$ represents that enough cells have been grown to infuse the participant at any dose level up to d_j but not above d_j and $[X_k < d_1]$ is the event that there are not enough cells to infuse the participant even at the lowest dose level. A participant for whom X_k is below the recommended dose level d_R , based on updated dose-limiting toxicity data, is still infused at his or her highest feasible dose level since this is the participant's best hope of therapeutic benefit. Participants for whom $[X_k < d_1]$ are not treated under study protocol since d_1 is considered to be the lowest dose having potential therapeutic benefit. For a maximum target accrual of M participants who have cells extracted, the number of participants who are infused at some dose level and evaluated for toxicity is $N = \sum_{j=1}^I I(X_k \geq d_j); N \leq M$. The random variable Z_j takes value 1 in the case that an accrued participant has enough cells extracted to be

infusible at dose level i (i.e., $X_k = d_i$), and denote the probability of feasibility at dose level d_i as $\theta_i = \Pr(X_k = d_i) = \Pr(Z_i = 1)$, which are monotonically decreasing with dose. Note that since X_k is observable quantity that is compared to every dose level prior to infusion of each participant, then Z_i is observable at every dose level i for each participant k . For each participant, we will know prior to dose assignment which dose levels are feasible for him or her to receive. Therefore, each participant is assessed for feasibility at every dose level, not just the dose level he or she actually receives. Using terminology of O'Quigley, Paoletti, and Maccario,⁶ this is akin to having a vector of complete information with regards to feasibility data for each participant. In Table 1, for example, if the number of epidermal growth factor receptor bispecific antibody armed T cells harvested for eligible participant k is approximately 90×10^9 , it is feasible for participant k to receive dose levels 1 and 2, and it is not feasible for the participant to receive dose levels 3 and 4. In this case, $Z_k = (1, 1, 0, 0)$. At any point in the trial, suppose z_i participants have been observed to be feasible to receive dose level d_i in a total of $m, m = 1, \dots, M$ participants whom have been accrued and have had cells obtained, which does not vary with dose. To model the probability of feasibility θ_i at each dose level, we will assume a beta-binomial model

$$z_i | \theta_i \sim \text{Binomial}(\theta_i); \theta_i \sim \text{Beta}(\tau_i, \nu_i),$$

where $\text{Beta}(\tau_i, \nu_i)$ is a beta distribution with parameters τ_i and ν_i . Based on the accumulated feasibility data $\mathbb{D}_i = (z_i, m)$ at dose level d_i , the posterior distribution of θ_i follows a beta distribution

$$\theta_i | \mathbb{D}_i \sim \text{Beta}(\tau_i + z_i, \nu_i + m - z_i).$$

Based on this distribution and the pre-specified desired benchmark for feasibility θ^* , we can calculate the posterior probability that dose level d_i is not feasible

$$\Pr(\theta_i < \theta^* | \mathbb{D}_i) = \mathcal{B}(\theta^*; \tau_i + z_i, \nu_i + m - z_i)$$

where $\mathcal{B}(\theta^*; \tau_i + z_i, \nu_i + m - z_i)$ is the cumulative density function of the beta distribution with parameters $\tau_i + z_i$ and $\nu_i + m - z_i$. We compare this probability to an upper probability cutoff p_F . If $\Pr(\theta_i < \theta^* | \mathbb{D}_i) > p_F$, we say that dose level d_i is not a feasible dose level. Based on this criterion, we establish a set F of feasible doses so that $F = \{d_i : \Pr(\theta_i < \theta^* | \mathbb{D}_i) \leq p_F\}$.

Toxicity

The procedure for updating the recommended dose based on toxicity mimics that described in Conaway, Dunbar, and Peddada.⁷ At any point in the trial, for those participants with $[X_k = d_1]$, the dose-limiting toxicity data for dose d_i is of the form $\Omega_i = \{(y_i, n_i) : i = 1, \dots, I\}$ with y_i equal to the number of observed dose-limiting toxicities and n_i equal to the number of participants who have been infused and evaluated for toxicity at dose level d_i . Let A denote the set of doses that have been tried thus far in the trial such that $A = \{d_i : n_i > 0\}$. To model the probability of dose-limiting toxicity π_i at each dose level, we will assume a beta-binomial model

$$y_i | \pi_i \sim \text{Binomial}(\pi_i); \quad \pi_i \sim \text{Beta}(\alpha_i, \beta_i),$$

where $\text{Beta}(\alpha_i, \beta_i)$ is a beta distribution with parameters α_i and β_i . Based on the accumulated toxicity data $\Omega_i = (y_i, n_i)$ at dose level d_i , the posterior distribution of π_i follows a beta distribution

$$\pi_i | \Omega_i \sim \text{Beta}(\alpha_i + y_i, \beta_i + n_i - y_i).$$

Based on this distribution, the updated dose-limiting toxicity probabilities, calculated only for $d_i \in A$, are given by the following posterior mean.

$$\hat{\pi}_i = \frac{y_i + \alpha_i}{n_i + \alpha_i + \beta_i}$$

To impose monotonicity with respect to the dose-toxicity relationship and borrow information across dose levels, we apply isotonic regression to the $\hat{\pi}_i$ using the pool adjacent violators algorithm⁸ with weights n_i , denoting the resulting estimates by $\tilde{\pi}_i$. This algorithm replaces adjacent estimates that violate the monotonicity assumption with their weighted average, where the weights are the current sample size at each dose level.

Updating the planned dose

1. For all $d_i \in A$, compute the loss $L_i(\tilde{\pi}_i, \pi^*) = |\tilde{\pi}_i - \pi^*|$ associated with each dose level based on the target dose-limiting toxicity rate π^* .
2. Let $l_{min} = \min_{d_i \in A} L_i(\tilde{\pi}_i, \pi^*)$, and let T be the set of doses with losses equal to the minimum observed loss so that $T = \{d_i: L_i(\tilde{\pi}_i, \pi^*) = l_{min}\}$.
3. If T contains more than one dose, then we choose from among them according to the rules:
 - a. If $\tilde{\pi}_t > \pi^* \forall t \in T$, the suggested dose is the lowest dose in T .
 - b. If $\tilde{\pi}_t \leq \pi^*$ for at least one $t \in T$, the suggested dose is the highest dose in T with $\tilde{\pi}_t \leq \pi^*$.
4. If the suggested dose level has an estimated dose-limiting probability that is less than the target, then the next highest dose level will be chosen if it has not yet been tried.

Trial conduct

Each accrued participant is to be allocated to the dose indicated by the estimation to be the closest to the target dose-limiting toxicity rate of π^* , conditional on the recommended dose being feasible for the next eligible participant. Let $d_R \in \{d_1, \dots, d_J\}$ be the recommended dose for the next entered cohort according to the above algorithm. Throughout the trial, each

eligible participant k is sequentially allocated to (1) his or her recommended dose d_R if X_k or (2) the highest feasible dose he or she is feasible to receive if $X_k < d_R$, or (3) untreated on study protocol if $X_k < d_1$. Based on the posterior distribution $\pi_1|\Omega_1$ and the pre-specified target dose-limiting toxicity rate π^* , we can calculate the posterior probability that dose level d_1 is too toxic

$$\Pr(\pi_1 > \pi^*|\Omega_1) = 1 - \mathcal{B}(\pi^*; \alpha_1 + y_1, \beta_1 + n_1 - y_1)$$

where $\mathcal{B}(\pi^*; \alpha_1 + y_1, \beta_1 + n_1 - y_1)$ is the cumulative density function of the beta distribution with parameters $\alpha_1 + y_1$ and $\beta_1 + n_1 - y_1$. We compare this probability to an upper probability cutoff p_T . If $\Pr(\pi_1 > \pi^*|\Omega_1) > p_T$, we say that dose level d_1 is not a safe dose level. In such a case, the trial will stop for safety, and no dose will be recommended as the feasible maximum tolerated dose. Otherwise, accrual to the study continues until either (1) N participants have been infused and evaluated for toxicity or (2) M participants have had cells extracted, regardless of whether or not they were feasible to receive any dose. The value of M represents a strict upper limit on the sample size based on available resources. Based on the posterior calculations from the final data, the feasible maximum tolerated dose is the highest dose in F that is d_R .

Results

Application to the glioblastoma phase I trial

The primary objective of the trial was to identify the feasible maximum tolerated dose, defined by the dose with dose-limiting toxicity rate closest to the target rate of $\pi^* = 25\%$ and achieves at least $\theta^* = 80\%$ feasibility. Prior information for the dose-limiting toxicity probabilities π_i is expressed through a distribution of the form $\pi_i \sim \text{Beta}(\alpha_i, \beta_i)$. Based on the expected value of π_i and a 95% upper limit u_i on the dose-limiting toxicity probability, the equations

$$\mathbb{E}(\pi_i) = \frac{\alpha_i}{\alpha_i + \beta_i} \quad \text{and} \quad \Pr(\pi_i \leq u_i) = 0.95,$$

are solved to obtain prior specifications for α_i and β_i . In the absence of prior information, a practical prior specification can be acquired by setting the prior mean equal to the target dose-limiting toxicity rate π^* , and setting the 95% upper limit u_i equal to $2 \times \pi^*$ at each dose level. This prior specification is recommended by Wages and Conaway⁹ to avoid the problem of rigidity¹⁰ in which allocation can become confined to a sub-optimal dose level regardless of the ensuing observed data. Based on a target dose-limiting toxicity rate of 25%, we used a Beta (2.34, 7.02) prior for π at each dose level. A similar process is followed for calibrating a prior for the feasibility probabilities. Based on the expected value of θ_i and a 95% upper limit c_i on the feasibility probability, the equations

$$\mathbb{E}(\theta_i) = \frac{\tau_i}{\tau_i + \nu_i} \quad \text{and} \quad \Pr(\theta_i \geq c_i) = 0.95,$$

are solved to obtain prior specifications for τ_i and v_i . These equations were solved using prior means $\mathbb{E}(\theta_i) = (0.90, 0.85, 0.80, 0.75)$ for $i = 1, \dots, 4$ and $c_i = \theta^* \div 2$ for all i , yielding $\tau_i = (1.46, 2.14, 2.66, 3.11)$ and $v_i = (0.16, 0.38, 0.67, 1.04)$. The maximum target accrual was $M = 30$ participants who will have cells extracted, and the number of participants who will be infused at some dose level and evaluated for toxicity is $N = 24$. The trial will begin on the lowest dose level d_1 , and the upper probability cutoffs used to define the set of feasible doses and the safety stopping rule were $p_F = 0.70$ and $p_T = 0.90$, respectively.

We illustrate the behavior of the described method under a set of true (dose-limiting toxicity, feasibility) probabilities (π_i, θ_i) are $\{(0.07, 0.90), (0.12, 0.80), (0.25, 0.70), (0.36, 0.60)\}$, indicating d_2 to be the feasible maximum tolerated dose with a true dose-limiting toxicity probability closest to 25% among those with feasibility probabilities of at least 80%. The data from the entire simulated trial are provided in Table 2. In simulating toxicity outcomes, we generated a latent toxicity tolerance⁶ t_k from a *Uniform*(0,1) distribution for participant k . If t_k is smaller than or equal to the true dose-limiting toxicity probability of the dose d_i administered to participant k (i.e., $t_k \leq \pi_i$), then this participant has a dose-limiting outcome; otherwise this participant does not have a dose-limiting toxicity outcome. Similarly, in simulating feasibility outcomes, we generated a latent feasibility outcomes s_k from a *Uniform*(0,1) distribution that represents the amount of dose it is feasible for participant k to receive. In this case, since we can observe the number of cells extracted for each participant and compare this value to each of the study dose levels, we compare s_k to each of the true feasibility probabilities $\theta_i, i = 1, \dots, I$. This process generates the vector of complete data Z_k for participant k .

The first participant in the simulated trial receives d_1 , and he/she does not have a dose-limiting toxicity. The set A of tried doses consists only of d_1 , so based on this observed non-dose-limiting toxicity, we update $\hat{\pi}_1 = \frac{0 + 2.34}{1 + 2.34 + 7.02} = 0.2259$. Following the above algorithm for updating the planned dose, we compute the loss for each dose in A . In this case, the set will T consist only of d_1 so it will be the suggested level. However, according to Step 4 of the algorithm, since the suggested dose has an estimated dose-limiting toxicity probability less than the target (i.e., $0.2259 < 0.25$), then the next highest dose level is to be chosen if it has not yet been tried. Since d_2 has not been tried, it is recommended for the second participant. This illustration demonstrates that since the computation of posterior estimates are made only on the doses that have been tried, the algorithm for updating the planned dose does not skip an untried dose level. Participants 2 and 3 receive d_2 and d_3 , respectively, and the first dose-limiting toxicity is observed in participant 3 on d_3 . For each of these participants, enough cells were harvested for him or her to receive the respective planned dose. Participant 1 would not have been infusible at d_4 . It is not feasible to administer any of the study dose levels to the fourth participant, since the number of cells grown for this participant did not reach the threshold of the lowest dose level. As the trial progresses, 3 of the first 7 participants are not infusible at d_4 , causing d_4 to be deemed not feasible and temporarily eliminated from consideration after participant 7. The posterior probability that d_4 is not feasible at this point is $\Pr(\theta_4 < \theta^* | \mathbb{D}_4) = 0.879$, which is larger than our pre-specified minimum threshold $p_F = 0.70$, so d_4 is eliminated from F . The classification of dose levels as feasible is adaptive, thus d_4 may return to F if a sufficient number of participants are deemed

feasible to receive d_4 . This is exactly what happens upon observing four consecutive participants who are infusible at d_4 after participant 11. At this point d_4 returns to F with a posterior probability of $\Pr(\theta_4 < \theta^* | \mathbb{D}_4) = 0.699$. However, the next two participants do not have enough cells harvested to receive d_4 and again it is eliminated from F when its posterior probability $\Pr(\theta_4 < \theta^* | \mathbb{D}_4) = 0.836$. Participant 17 is the first occurrence of a dose-limiting toxicity observed at a dose lower than the currently planned dose. It was recommended that this participant receive d_3 but it was only feasible for he or she to be administered d_1 based on leukapheresis. Prior to the accrual of participant 17, the estimated dose-limiting toxicity probabilities at $d_i, i = 1, \dots, 3$ were $\tilde{\pi}_i = (0.153, 0.153, 0.265)$ indicating d_3 to be the recommended dose. After observing a dose-limiting toxicity probability in participant 17 on d_1 , these probability estimates were updated to $\tilde{\pi}_i = (0.177, 0.177, 0.265)$. This illustrates the fact that the sequential borrowing of information across dose levels accounts for this observed dose-limiting toxicity away from the currently planned dose. After participant 18, both d_3 and d_4 are adaptively eliminated from the feasible set of doses with respective posterior probabilities of 0.74 and 0.996. Ultimately, the design settles on d_2 as the feasible and maximum tolerated dose after the accrual of $N = 24$ participants to the study who have been infused and evaluate for toxicity.

Simulation results were generated in R¹¹ to display the performance of the design operating characteristics. For each scenario in Table 3, we ran 10,000 simulated trials. The results from 10,000 simulated trials of each scenario take approximately 21 seconds to generate. The results report the true dose-limiting toxicity rate at each dose level, the true feasibility rate at each dose level, the percentage of trials in which each dose level was recommended as the feasible maximum tolerated dose, the average number of participants treated in a simulated study, and the percentage of times the trial was stopped early. The results are reported in Table 3, in which bold type indicates the true feasible maximum tolerated dose (when one exists) or the correct decision to stop the trial early. In Scenario 1, all dose levels are considered feasible and safe, and the proposed method performs very well, recommending the true feasible maximum tolerated dose in 90.7% of simulated trials and treating 15.3 of 23.7 (65%), on average, of accrued participants at this dose. Scenario 2 illustrates the impact on operating characteristics of the design when the feasibility probabilities are changed relative to Scenario 1. With the same toxicity profile, the design is able to account for a lack of feasibility at the highest dose level and identify d_3 as the true feasible maximum tolerated dose more than any other dose. Scenario 3 illustrates the impact on operating characteristics of the design when the dose-limiting toxicity probabilities are altered relative to Scenario 1. In Scenario 3, all dose level still have true dose-limiting toxicity probabilities less than or equal to the target rate of 25%, although the probabilities have increased from Scenario 1. In this case, the percentage of correct feasible maximum tolerated dose selection decreases to 61.3% from the 90.7% observed in Scenario 1. Similar conclusions can be made when examining the relationships between Scenarios 4 and 5, as well the relationships between Scenarios 6 and 7. In Scenario 8, d_1 is the true feasible and maximum tolerated dose, and the proposed method is able to recommend this dose level at the conclusion of the study in 56.8% of simulated trials, while treating 13.2 of 21.2 (62%), on average, of accrued participants at this dose. The final two scenarios illustrate the method's ability to stop in the absence of feasibility (Scenario 9) or in the presence of excessive toxicity (Scenario 10) in a

high percentage of simulated trials (i.e., > 95%). Overall the simulation results indicate that the design outlined in this article is a practical early-phase adaptive method for use in trials with safety and feasibility endpoints.

Comparison to a previous approach

To our knowledge, the only other existing method that adaptively accounts for both safety and feasibility in its design is the method of Thall et al.³ This method employs a continual reassessment method² modeling approach for both toxicity and feasibility, so we compare our beta-binomial approach to the operating characteristics reported in their paper. All scenarios and results provided in Table 4 are taken from Table 1 in Thall et al.³ The feasible maximum tolerated dose is defined by the dose with dose-limiting toxicity rate closest to the target rate of $\pi^* = 30\%$ and achieves at least $\theta^* = 50\%$ feasibility. Based on a target dose-limiting toxicity rate of 30%, we used a Beta (2.07, 4.83) prior for π at each dose level. The priors for feasibility probabilities were based on using prior means $E(\theta_i) = (0.90, 0.85, 0.80, 0.75, 0.70)$ for $i = 1, \dots, 5$ and $c_i = \theta^* \div 2$ for all i , yielding $\tau_i = (1.46, 2.14, 2.66, 3.11, 3.51)$ and $v_i = (0.16, 0.38, 0.67, 1.04, 1.51)$. The maximum target accrual was $M = 48$ participants who will have cells extracted, and the number of participants who will be infused at some dose level and evaluated for toxicity is $N = 24$. The upper probability cutoffs used to define the set of feasible doses and the safety stopping rule were $p_F = 0.70$ and $p_T = 0.70$, respectively, for the proposed method and $p_F = 0.90$ and $p_T = 0.90$ for the method of Thall et al. For each scenario in Table 4, 10,000 simulated trials were run.

In Scenario 1, our proposed method identifies the true feasible maximum tolerated dose in a higher percentage of trials than the method of Thall et al. (68.4% vs. 60.7%, respectively), while treating a higher number of participants on average at this dose (12.0 vs. 10.4, respectively). Additionally, the proposed method recommends the overly toxic dose level 3 as the feasible maximum tolerated dose and incorrectly stops the trial early in a smaller percentage of simulated trials. In Scenario 2, the method of Thall et al. displays better operating characteristics than our proposed method by recommending the true feasible maximum tolerated dose in a higher percentage of trials (74.5% vs. 69.5%), while treating an average of one additional participant at this dose. It is interesting to note that even though all dose levels in Scenario 2 are feasible and dose level 1 has a dose-limiting toxicity probability of 0.10, the Thall et al. method still terminates early in more than 11% of simulated trials. The proposed method does this less than 4% of the time. The performance of the two competing approaches is nearly identical across the remaining three scenarios. Both methods appropriately stop early in a high percentage of trials in Scenarios 3 and 5, and recommend the dose level 5 as the feasible maximum tolerated dose in approximately 60.0% of trials in Scenario 4. A similar number of participants are treated at dose level 5 on average in Scenario 4 (5.6 vs. 5.7). Overall, Table 4 indicates that the proposed method is competitive with that of Thall et al, making it practical alternative for designing and conducting early-phase dose-finding trials that must account for dose feasibility.

Conclusions

In this paper, we have proposed a new adaptive dose-finding method that accounts for both safety and dose feasibility in designing early-phase clinical trials. This method has applications in the early development of adoptive T cell therapies aiming to identify feasible maximum tolerated doses. The proposed framework models toxicity and feasibility independently, with formally accounting for the correlation between the two outcomes. We feel that this is reasonable given the limited amount of data available in early phase studies. Our simulation studies demonstrate desirable operating characteristics without the need for the modeling of additional parameters.¹² While this more under-parameterized approach is mathematically simple, it is feasible to implement. Specifically, the modeling of additional parameters in this setting requires strong assumptions about the correct specification of the additional parameter. In the clinical setting, such strong information is usually not available, and work has been done showing the large potential biases produced by such assumptions.¹³

Adoptive cell therapy is a promising cancer treatment that needs evaluation by properly designed clinical trials. This therapy entails manufacturing the cell product for each individual, a variable that becomes extremely significant in the design of dose finding trials and one that is usually irrelevant in phase I clinical trials for cytotoxic chemotherapy. In these studies, it is difficult to identify a feasible dose to carry forward into middle and late development, hindering the evaluation of the therapeutic intervention. A design that incorporates feasibility, as a function of the quantity and quality of the product manufactured, and safety carries the potential to lead to more optimal dosing and improved patient outcomes in adoptive cell therapies. We applied the proposed methodology to a phase I study targeting newly diagnosed glioblastoma using epidermal growth factor receptor bispecific antibody armed activated T cells in combination with radiation and chemotherapy, and compared it via simulation studies in Supplemental Material to the only existing alternative method to our knowledge. Software in the form of R code for simulation of the proposed design is freely available for download at http://faculty.virginia.edu/model-based_dose-finding/.

Acknowledgements

The authors would like to thank Drs. Mark Conaway and Gina Petroni for helpful discussions on this topic.

Funding: R03CA238966.

References

1. Lee DW, Kochenderfer JN, Stetler-Stevenson M, et al. T cells expressing CD19 chimeric antigen receptors for acute lymphoblastic leukemia in children and young adults: a phase I dose-escalation trial. *Lancet* 2015; 517–528.
2. O’Quigley J, Pepe M and Fisher L. Continual reassessment method: a practical design for phase I clinical trials in cancer. *Biometrics* 1990; 46: 33–48. [PubMed: 2350571]
3. Thall PF, Sung H-G and Choudhury A. Dose-finding based on feasibility and toxicity in T-cell infusion trials. *Biometrics* 2001; 57: 914–921. [PubMed: 11550945]
4. Lum LG, Thankur A, Al-Kadhimi Z, et al. Targeted T-cell therapy in stage IV breast cancer: a phase I clinical trial. *Clin Cancer Res* 2015; 21: 2305–2314. [PubMed: 25688159]

5. Lum LG, Thakur A, Liu A, Deol A, et al. CD20-targeted T cells after stem cell transplantation for high risk and refractory non-Hodgkin's lymphoma. *Biol Blood Marrow Transplant* 2013; 19: 925–933. [PubMed: 23529012]
6. O'Quigley J, Paoletti X and Maccario J. Non-parametric optimal design in dose finding studies. *Biostatistics* 2002; 3: 51–56. [PubMed: 12933623]
7. Conaway MR, Dunbar S and Peddada SD. Designs for single- or multiple-agent phase I trials. *Biometrics* 2004; 60: 661–669. [PubMed: 15339288]
8. Robertson T, Wright TF and Dykstra R. *Order restricted statistical inference*. New York: John Wiley, 1988.
9. Wages NA and Conaway MR. Revisiting isotonic phase I design in the era of model-assisted dose-finding. *Clin Trials* 2018; 15: 524–529. [PubMed: 30101616]
10. Cheung YK. On the use of nonparametric curves in phase I trials with low toxicity tolerance. *Biometrics* 2002; 58: 237–240. [PubMed: 11890321]
11. R Development Core Team. *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing, 2008.
12. Iasonos A, Wages NA, Conaway MR, et al. Dimension of model parameter space and operating characteristics in adaptive dose finding studies. *Stat Med* 2016; 35: 3760–3775. [PubMed: 27090197]
13. Cunanan K and Koopmeiners JS. Evaluating the performance of copula models in phase I-II clinical trials under model-misspecification. *BMC Med Res Methodol* 2014; 14: 51. [PubMed: 24731155]

Table 1.

Study dose levels for a phase I study targeting newly diagnosed glioblastoma using epidermal growth factor receptor bispecific antibody armed T cells.

Study dose level	Dose of activated T cells per IV infusion	Total dose level per patient
1 (start)	5×10^9	40×10^9
2	10×10^9	80×10^9
3	20×10^9	160×10^9
4	40×10^9	320×10^9

Table 2.

A simulated sequential trial illustrated the described dose-finding design. The feasible maximum tolerated dose recommendation is d_2 after $N=24$ participants have infused and evaluated for safety. DLT = dose-limiting toxicity.

Dose levels									
Participant	Feasibility tolerance	Indicator of feasibility				Planned dose	Highest feasible	Received dose	DLT
		Z ₁	Z ₂	Z ₃	Z ₄				
1	0.682	1	1	1	0	1	3	1	0
2	0.431	1	1	1	1	2	4	2	0
3	0.438	1	1	1	1	3	4	3	1
4	0.990	0	0	0	0	2	0	n/a	n/a
5	0.347	1	1	1	1	2	4	2	0
6	0.086	1	1	1	1	2	4	2	0
7	0.678	1	1	1	0	2	3	2	0
8	0.430	1	1	1	1	2	4	2	0
9	0.380	1	1	1	1	3	4	3	0
10	0.091	1	1	1	1	3	4	3	0
11	0.378	1	1	1	1	3	4	3	0
12	0.776	1	1	0	0	3	2	2	0
13	0.719	1	1	0	0	3	2	2	0
14	0.667	1	1	1	0	3	3	3	0
15	0.406	1	1	1	1	3	4	3	1
16	0.668	1	1	1	0	3	4	3	0
17	0.819	1	0	0	0	3	1	1	1
18	0.842	1	0	0	0	3	1	1	0
19	0.187	1	1	1	1	2	4	2	0
20	0.762	1	1	0	0	2	2	2	0
21	0.306	1	1	1	1	2	4	2	0
22	0.419	1	1	1	1	2	4	2	0
23	0.447	1	1	1	1	2	4	2	0

Dose levels									
Participant	Feasibility tolerance	Indicator of feasibility				Planned dose	Highest feasible	Received dose	DLT
		Z ₁	Z ₂	Z ₃	Z ₄				
24	0.782	1	1	0	0	2	2	2	0
25	0.449	1	1	1	1	2	4	2	0

Table 3.

Distribution of feasible maximum tolerated dose recommendation and participant allocation for the proposed design. For each scenario, the table reports the recommendation percentage for each dose level, the average number of patients treated on each dose level, the average sample size, and the percentage of trials stopped early. Bold type indicates the true feasible maximum tolerated dose (when one exists) or the correct decision to stop the trial early.

Scenario	Dose level				Size	None
	1	2	3	4		
1	(0.01, 0.99) 0.0 (1.7)	(0.04, 0.97) 0.6 (2.2)	(0.07, 0.95) 7.7 (4.5)	(0.12, 0.90) 90.7 (15.3)	23.7	1.07
2	(0.01, 0.95) 0.6 (2.7)	(0.04, 0.90) 20.3 (6.6)	(0.07, 0.80) 46.7 (7.8)	(0.12, 0.70) 25.8 (5.4)	22.5	6.55
3	(0.04, 0.99) 0.1 (2.1)	(0.07, 0.97) 4.0 (3.2)	(0.12, 0.95) 33.4 (8.0)	(0.25, 0.90) 61.3 (10.5)	23.8	1.23
4	(0.04, 0.95) 0.9 (3.0)	(0.07, 0.90) 22.5 (7.3)	(0.12, 0.80) 51.6 (8.3)	(0.25, 0.70) 18.1 (3.8)	22.4	6.86
5	(0.07, 0.99) 2.4 (3.0)	(0.12, 0.97) 26.6 (6.9)	(0.25, 0.95) 50.5 (8.9)	(0.36, 0.90) 19.3 (4.9)	23.7	1.21
6	(0.07, 0.90) 4.8 (4.2)	(0.12, 0.85) 61.3 (11.7)	(0.25, 0.68) 14.3 (3.3)	(0.36, 0.60) 1.3 (0.8)	20.0	18.24
7	(0.12, 0.95) 22.6 (7.7)	(0.25, 0.90) 53.5 (10.1)	(0.36, 0.80) 15.9 (3.8)	(0.50, 0.70) 1.1 (0.8)	22.4	6.89
8	(0.25, 0.95) 56.8 (13.2)	(0.36, 0.90) 23.3 (6.2)	(0.50, 0.80) 2.3 (1.5)	(0.65, 0.70) 0.1 (0.3)	21.2	17.5
9	(0.01, 0.65) 3.7 (2.6)	(0.04, 0.55) 0.8 (1.4)	(0.07, 0.45) 0.1 (0.5)	(0.12, 0.35) 0.0 (0.1)	4.6	95.36
10	(0.60, 0.99) 1.8 (7.6)	(0.75, 0.97) 0.0 (0.7)	(0.89, 0.95) 0.0 (0.1)	(0.95, 0.90) 0.0 (0.0)	8.4	98.19

Table 4.

Distribution of feasible maximum tolerated dose recommendation and participant allocation for the proposed design and the design of Thall et al.³ For each scenario, the table reports the recommendation percentage for each dose level, the average number of patients treated on each dose level, the average sample size, and the percentage of trials stopped early. Bold type indicates the true feasible maximum tolerated dose (when one exists) or the correct decision to stop the trial early.

Scenario	Dose level					
	1	2	3	4	5	
1	(0.10,0.99)	(0.30,0.95)	(0.50,0.90)	(0.70,0.75)	(0.80,0.50)	None
Proposed	14.2 (6.3)	68.4 (12.0)	13.1 (4.1)	0.7 (0.7)	0.1 (0.1)	3.47
Thall et al	2.9 (1.7)	60.7 (10.4)	27.8 (9.7)	0.1 (0.7)	0.0 (0.0)	8.5
2	(0.10,0.90)	(0.30,0.75)	(0.50,0.50)	(0.70,0.25)	(0.80,0.05)	
Proposed	13.3 (8.3)	69.5 (12.0)	13.4 (2.7)	0.0 (0.2)	0.0 (0.0)	3.77
Thall et al	3.8 (4.6)	74.5 (13.0)	10.3 (4.0)	0.1 (0.2)	0.0 (0.0)	11.3
3	(0.05,0.25)	(0.10,0.10)	(0.30,0.05)	(0.50,0.02)	(0.60,0.01)	
Proposed	0.0 (1.4)	0.0 (0.2)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	99.95
Thall et al	0.0 (2.6)	0.4 (1.2)	2.8 (0.5)	0.4 (0.0)	0.0 (0.0)	96.4
4	(0.01,0.99)	(0.05,0.95)	(0.07,0.90)	(0.10,0.75)	(0.30,0.50)	
Proposed	0.0 (2.1)	0.2 (2.6)	2.3 (5.0)	36.9 (8.6)	60.5 (5.6)	0.03
Thall et al	0.0 (1.0)	0.4 (1.5)	9.2 (7.3)	28.8 (8.5)	61.5 (5.7)	0.1
5	(0.50,0.90)	(0.60,0.75)	(0.70,0.50)	(0.75,0.25)	(0.80,0.05)	
Proposed	5.8 (6.2)	0.7 (1.5)	0.0 (0.3)	0.0 (0.0)	0.0 (0.0)	93.5
Thall et al	4.8 (3.2)	3.2 (3.6)	0.0 (1.7)	0.0 (0.0)	0.0 (0.0)	92