Received: 23 January 2019 | Revised: 27 August 2019 | Accepted: 28 September 2019

DOI: 10.1002/sim.8402

RESEARCH ARTICLE

WILEY Statistics in Medicine

Incorporating patient-reported outcomes in dose-finding clinical trials

Shing M. Lee | **Xiaoqi Lu** | **Bin Cheng**

Department of Biostatistics, Mailman School of Public Health, Columbia University, New York, New York

Correspondence
Shing M. Lee, Department of Biostatistics, Mailman School of Public Health, Columbia University, New York, NY 10032.
Email: sml2114@cumc.columbia.edu

Funding information
American Cancer Society, Grant/Award Number: MRSG-13-146-01-CPHPS; National Center for Advancing Translational Sciences, National Institutes of Health, Grant/Award Number: UL1 TR001873,

Oncology dose-finding clinical trials determine the maximum tolerated dose (MTD) based on toxicity outcomes captured by clinicians. With the availability of more rigorous instruments for measuring toxicity directly from patients, there is a growing interest to incorporate patient-reported outcomes (PRO) in clinical trials to inform patient tolerability. This is particularly important for dose-finding trials to ensure the identification of a well-tolerated dose. In this paper, we propose three extensions of the continual reassessment method (CRM), termed PRO-CRMs, that incorporate both clinician and patient outcomes. The first method is a marginal modeling approach whereby clinician and patient toxicity outcomes are modeled separately. The other two methods impose a constraint using a joint outcome defined based on both clinician and patient toxicities and model them either jointly or marginally. Simulation studies show that while all three PRO-CRMs select well-tolerated doses based on clinician's and patient's perspectives, the methods using a joint outcome perform better and have similar performance. We also show that the proposed PRO-CRMs are consistent under robust model assumptions.

KEYWORDS

continual reassessment method, maximum tolerated dose, multiple toxicity constraints, phase I trial, robust modeling

1 | INTRODUCTION

The goal of dose finding clinical trials is to determine the maximum tolerated dose (MTD), defined as the dose associated with a prespecified target probability of toxicity. Conventionally, toxicities are captured by clinicians using the National Cancer Institute Common Terminology Criteria for Adverse Events (NCI-CTCAE) on a scale from 1 to 5 with 1 being a mild toxicity, 2 a moderate toxicity, 3 a severe toxicity, 4 a life-threatening toxicity, and 5 death.[1] In recent years, there has been increased interest in also including the patient's perspective in clinical trials with the understanding that patients provide a unique perspective on their overall health, function, and symptoms which cannot be captured by clinician-reported outcomes.[2,3] While patient-reported outcomes (PROs) have generally been included in randomized trials to complement the reporting of the NCI-CTCAE, they are increasingly also being included in phase 1/2 trials. One example is the CheckMate040, which collected patient-reported health status at baseline and every 6 weeks during their 25 week treatment.[4]

A growing literature also suggests disagreement between clinician- and patient-reported symptoms, with patients reporting more frequent and more severe symptoms.[5-8] The correlation between clinician- and patient-reported symptoms was evaluated in three breast cancer randomized clinical trials.[7] Among the patients for whom clinicians reported no toxicity, 13% to 50% reported "very much" toxicity. The disagreement was more pronounced when the symptoms

were not previously documented, and thus unexpected, suggesting that this is much more of a concern in early phase trials.[7] This degree of disagreement is concerning in the context of dose-finding trials because it implies that a dose identified solely based on clinician input may be intolerable from a patient's perspective. This calls for the inclusion of patient-reported symptoms in dose-finding trials. However, dose-finding designs incorporating both clinician and patient outcomes have not been proposed. One of the impediments has been the lack of a unified and reliable instrument for the collection of patient-reported outcomes in clinical trials. However, with the improved rigor of PRO measurements, such as the Patient-Reported Outcomes version of the CTCAE (PRO-CTCAE[9,10]), there is impetus to use PROs to inform patient tolerability and to incorporate PROs in the regulatory process of clinical trials.[11,12]

With the availability and implementation of the PRO-CTCAE, we expect to have a subset of toxicities that are rated by both clinicians and patients. One example is a dose-finding trial of cilengitide and paclitaxel which collected the PRO-CTCAE in cycle one along with the NCI-CTCAE, and although the PRO-CTCAE was not used for determining the MTD, it was able to capture symptoms not reported by clinicians. Moreover, patient reports indicated more severe symptoms.[13] Thus, it is timely to develop dose-finding methods that can also incorporate toxicities captured using the PRO-CTCAE in the determination of the MTD. This is of increased relevance and importance for clinical trials of novel anticancer therapies, given that they have been associated with low-grade toxicities that can impact patient tolerability while maintaining similar or lower rates of clinician-reported severe toxicity. With the increased number of dose-finding trials evaluating novel treatments,[14] it is important to ensure that doses selected can be well tolerated by patients before evaluation of efficacy. In this paper, we propose novel approaches for incorporating both clinician's and patient's perspectives in the determination of the MTD. The methods, termed patient-reported outcomes continual reassessment method (PRO-CRM), are extensions of the CRM[15] and allow for the specification of a target probability of toxicity based solely on clinician input and, in addition, allow for the specification of a toxicity threshold based on patient input. In Section 2, we describe the problem formulation for the proposed methods. In Section 3, we present a motivating example and the setup for the simulation studies. In Section 4, we present the simulation results evaluating the operating characteristics of the proposed methods under various scenarios of toxicity probabilities. We end with a discussion in Section 5.

2 | METHODS

2.1 | Outcome definition

Most commonly used dose-finding methods assume a binary outcome, conventionally defined as the presence or absence of a dose-limiting toxicity (DLT). In the context of cancer therapies, a DLT is generally defined as the presence of any grade 3 or higher nonhematological toxicity or any grade 4 or higher hematological toxicity using the NCI-CTCAE. Let Y_C be a clinician-reported binary toxicity outcome, where $Y_C = 1$ indicates the occurrence of DLT. Given m dose levels $d_1, \cdots, d_m$ under investigation with $d_1 < \cdots < d_m$, the MTD in traditional dose finding studies is defined as

$$d_{C,*} = \text{argmax}_{d \in \mathcal{D}} \{P(Y_C = 1|d) \leq \theta\}, \tag{1}$$

where $\mathcal{D} = \{d_1, \ldots, d_m\}$ and $\theta \in (0, 1)$ is a prespecified threshold for clinician-reported DLT rate, typically chosen as 0.25 or 0.33.

With the availability of the PRO-CTCAE, we will have a subset of symptomatic toxicities with ratings on severity, frequency, interference, amount, and presence. The PRO-CTCAE comprises 124 items covering 78 symptomatic toxicities from the CTCAE.[9] While many CTCAE categories cannot be evaluated by patients themselves (eg, laboratory values), the information obtained from the PRO-CTCAE can be complementary to the CTCAE in providing a better overall assessment of symptom tolerability. Some items have a single dimension, while others may have multiple dimensions. The dimensions are rated as follows: severity is graded very similar to the CTCAE with each toxicity being rated as none, mild, moderate, severe, and very severe; interference is rated on a scale of not at all, a little bit, somewhat, quite a bit, and very much; frequency is rated as none, rarely, occasionally, frequently, and almost constantly; and presence is a binary variable.[9] Thus, similar to clinician-reported toxicity, we can define patient-reported DLT based on the PRO-CTCAE as the presence of any toxicity with a rating of severe or worse in severity, quite a bit or worse in interference, frequent or more often in frequency, or present for presence, assuming these are not present at baseline. DLT from the patient's perspective can also be defined based on changes from baseline reported symptoms. Let Y_P to be the patient-reported binary toxicity outcome, where $Y_P = 1$ indicates the occurrence of a patient-reported DLT.

Based on the availability of these two binary outcomes, Y_C and Y_P, we can apply the multiple constraint framework. We consider them as separate outcomes and elicit the toxicity thresholds for clinician and patient separately and impose two separate constraints. In practice, eliciting an additional patient-reported toxicity threshold from clinicians should be simple given their existing familiarity with the specification of θ. However, most likely Y_C and Y_P are correlated. In that case, it may be preferable to redefine the two binary outcomes as Y_C and $\max\{Y_C, Y_P\}$. That is, we keep the constraint specified in traditional designs for Y_C, but impose an additional toxicity constraint on the probability that either $Y_C = 1$ or $Y_P = 1$. The elicitation of this threshold would require some knowledge regarding the correlation between Y_C and Y_P.

2.2 | Designs

We propose three approaches based on the outcome definitions specified above. All three approaches are extensions of the CRM.[15] They differ in the way the outcomes, constraints, and target probability of toxicities are specified. Moreover, different modeling approaches are evaluated.

2.2.1 | Marginal outcomes

In this approach, we define the MTD based on clinician outcome, Y_C, as in (1) and define the MTD based on the patient's perspective as

$$d_{P,*} = \text{argmax}_{d\in D}\left\{P(Y_P = 1|d) \le \phi\right\}, \tag{2}$$

where ϕ is a preselected threshold for patient-reported DLT, and no inequality relationship is imposed between ϕ and θ. The global MTD based on the marginal method is defined as

$$d_* = \min\left\{d_{C,*}, d_{P,*}\right\}. \tag{3}$$

Since definitions (1) and (2) do not involve any joint distribution of (Y_C, Y_P), we propose to estimate them separately using two different working models. Specifically, to estimate $d_{C,*}$, we consider a working model

$$P(Y_C = 1|d) = u^\beta,\ u = H_C(d), \tag{4}$$

where $\beta > 0$ is an unknown parameter and $H_C(\cdot)$ is a known invertible function mapping $\{d_1, \ldots, d_M\}$ to $\{u_1, \ldots, u_m\}$, a dose skeleton as specified in the CRM. Similarly, to estimate $d_{P,*}$, we use the following working model:

$$P(Y_P = 1|d) = v^\gamma,\ v = H_P(d), \tag{5}$$

where $\gamma > 0$ is an unknown parameter and $H_P(\cdot)$ is a known invertible function mapping $\{d_1, \ldots, d_M\}$ to $\{v_1, \ldots, v_m\}$, another skeleton which may differ from $\{u_1, \ldots, u_m\}$. Given data on the first n patients, $\{(d(i), Y_{C,i}, Y_{P,i}), i = 1, \ldots, n\}$, the conditional likelihood function based solely on clinician outcome and conditioning on the dose assignments $\{d(i), i = 1, \ldots, n\}$ is

$$L_n(\beta) = \prod_{i=1}^{n}\{1 - u(i)^\beta\}^{I(Y_{C,i}=0)}u(i)^{I(Y_{C,i}=1)\beta},$$

where $u(i) = H_C(d(i))$. Similarly, the conditional likelihood function based solely on patient outcome is

$$L_n(\gamma) = \prod_{i=1}^{n}\{1 - v(i)^\gamma\}^{I(Y_{P,i}=0)}v(i)^{I(Y_{P,i}=1)\gamma},$$

where $v(i) = H_P(d(i))$.

Given that heterogeneity has to be achieved for both Y_C and Y_P to estimate the model parameters using the maximum likelihood estimator (MLE), a staged approach for dose selection starting at the lowest dose is used. A rule-based design is followed, for example, using a cohort of one and escalating if no toxicity is observed, until heterogeneity is achieved for either Y_C or Y_P. At this time, we proceed to the second stage and estimate the model parameter using the MLE for the outcome for which heterogeneity has been achieved and continue with the rule-based design for the other outcome.

Without loss of generality, assume that heterogeneity has been achieved for Y_P, but not for Y_C, and let $\hat{\gamma}_n$ be the MLE for γ, then the $(n+1)$th patient will be assigned to the current estimate of d_*

$$\hat{d}_{*,n} = \min\left\{\text{argmin}_{d\in D}\left(|H_P(d)^{\hat{\gamma}_n} - \phi|\right), \hat{d}_{C,n+1}\right\}, \tag{6}$$

where $\hat{d}_{C,n+1}$ is the dose to be assigned to the $(n+1)$th patient based on the rule-based design for clinician. Once heterogeneity has been achieved for both Y_C and Y_P, we switch to the third stage where the MLE is used to estimate both parameters. Let $\hat{\beta}_n$ and $\hat{\gamma}_n$ be the MLE for clinician and patient outcomes, respectively. Then, the $(n+1)$th patient will be assigned to

$$\hat{d}_{*,n} = \min\left\{\text{argmin}_{d\in D}\left(|H_C(d)^{\hat{\beta}_n} - \theta|\right), \text{argmin}_{d\in D}\left(|H_P(d)^{\hat{\gamma}_n} - \phi|\right)\right\}. \tag{7}$$

We can further prove the consistency of $\hat{d}_{*,n}$. Define

$$B = \left\{(\beta,\gamma) : d_* = \min\left(\arg\min_{d\in D}|H_C(d)^{\beta} - \theta|, \arg\min_{d\in D}|H_P(d)^{\gamma} - \phi|\right)\right\},$$
$$B_C = \left\{\beta : d_{C,*} = \arg\min_{d\in D}|H_C(d)^{\beta} - \theta|\right\},$$

and

$$B_P = \left\{\gamma : d_{P,*} = \arg\min_{d\in D}|H_P(d)^{\gamma} - \phi|\right\}.$$

Define also β_{0l} such that $P(Y_C = 1|d_l) = H_C(d_l)^{\beta_{0l}}$ for $l = 1, \ldots, m$. Similarly, define γ_{0l} such that $P(Y_P = 1|d_l) = H_P(d_l)^{\gamma_{0l}}$ for $l = 1, \ldots, m$.

Theorem 1. *Under working models (4) and (5), assume that* $\{\beta_{01}, \ldots, \beta_{0m}\} \subset B_C$ *and* $\{\gamma_{01}, \ldots, \gamma_{0m}\} \subset B_P$. *Then,* $\hat{d}_{*,n} \to d_*$ *almost surely as* $n \to \infty$.

The proof of Theorem 1 is given in Appendix A.

2.2.2 | Joint outcome with joint modeling

By redefining Y_C and Y_P, as Y_C and $\max\{Y_C, Y_P\}$, we create a joint clinician and patient outcome. In this approach, in addition to constraint (1), we propose imposing a constraint on $\max\{Y_C, Y_P\}$, that is,

$$d_{CP,*} = \text{argmax}_{d\in D}\{P(\max\{Y_C, Y_P\} = 1|d) \le \psi\} = \text{argmax}_{d\in D}\{P(Y_C = 1 \quad \text{or} \quad Y_P = 1|d) \le \psi\}, \tag{8}$$

where ψ is a prespecified threshold for either clinician or patient reporting a DLT, such that $0 < \theta < \psi < 1$. Note that we do not assume $Y_P = 0$ if $Y_C = 0$. In fact, the reason for including PROs is to accommodate cases where $Y_C = 0$ and $Y_P = 1$. In these scenarios, ignoring patient's PRO information can lead to an overly aggressive dose recommendation. We define the global MTD based on the joint model method as

$$d_{**} = \min\{d_{C,*}, d_{CP,*}\}. \tag{9}$$

The MTD d_{**} is essentially an MTD under multiple constraints as proposed by Lee et al.[16,17] To estimate d_{**}, we consider the following joint working model:

$$P(Y_C = 1 \text{ or } Y_P = 1|d) = u^{\beta_1}, \;\; P(Y_C = 1|d) = u^{\beta_1+\beta_2}, \;\; u = H_C(d), \tag{10}$$

where $\beta_1, \beta_2 > 0$ are unknown parameters, and $\{u_1, \ldots, u_m\}$ is the dose skeleton. Specifically, the maximum likelihood estimate $(\hat{\beta}_{1n}, \hat{\beta}_{2n})$ based on $\{(d(i), Y_{C,i}, Y_{P,i}), i = 1, \ldots, n\}$ of the first n subjects is the MLE of the conditional likelihood

$$L_n(\beta_1,\beta_2) = \prod_{i=1}^{n}\{1 - u(i)^{\beta_1}\}^{I(Y_{C,i}=0,Y_{P,i}=0)}\{u(i)^{\beta_1} - u(i)^{\beta_1+\beta_2}\}^{I(Y_{C,i}=0,Y_{P,i}=1)}u(i)^{(\beta_1+\beta_2)I(Y_{C,i}=1)}, \tag{11}$$

where $u(i) = H_C(d(i))$.

The dose assignment process follows a similar staged approach as previously described for marginal modeling, with the exception of having a single skeleton for the estimation of the model parameters. That is, a rule-based design is followed starting from the lowest dose. Once heterogeneity is achieved for either Y_C or Y_P, we proceed to the second stage by estimating the model parameter using the marginal CRM for the outcome for which heterogeneity has been achieved, continuing with the rule-based design for the other outcome, and assign the minimum of those two doses. Once heterogeneity has been achieved for both Y_C and Y_P, we switch to the third stage and estimate the parameters using Equation (11). Let $\hat{\beta}_{1n}$ and $\hat{\beta}_{2n}$ be the MLE, the $(n+1)$th patient is assigned to

$$\hat{d}_{**,n} = \min\left\{\text{argmin}_{d\in\mathcal{D}}\left(|H_C(d)^{\hat{\beta}_{1n}+\hat{\beta}_{2n}} - \theta|\right), \text{argmin}_{d\in\mathcal{D}}\left(|H_C(d)^{\hat{\beta}_{1n}} - \psi|\right)\right\}. \tag{12}$$

Theorem 2. *Under the joint working model (10), $\hat{d}_{**,n} \to d_{**}$ almost surely as $n \to \infty$.*

For a proof of Theorem 2, see the work of Cheng and Lee.[17]

2.2.3 | Joint outcome marginal modeling

In addition to the method based on the joint likelihood $L_n(\beta_1, \beta_2)$, we also propose a marginal method to estimate d_{**}. The method applies a marginal approach to each of the two constraints (1) and (8) separately. Specifically, in addition to working model (4), we consider a working model for the event $\{Y_C = 1 \text{ or } Y_P = 1\} = \{\max\{Y_C, Y_P\} = 1\}$ as

$$P(Y_C = 1 \text{ or } Y_P = 1|d) = P(\max\{Y_C, Y_P\} = 1|d) = w^{\eta}, \;\; w = H_{CP}(d), \tag{13}$$

where $\eta > 0$ is an unknown parameter and $H_{CP}(\cdot)$ is a known invertible function mapping $\{d_1, \ldots, d_m\}$ to a dose skeleton $\{w_1, \ldots, w_m\}$. Given the data of the first n patients $\{(d(i), Y_{C,i}, Y_{P,i}), i = 1, \ldots, n\}$, the MLE $\hat{\eta}_n$ can be obtained using the following conditional likelihood:

$$L_n(\eta) = \prod_{i=1}^{n}\{1 - w(i)^{\eta}\}^{I(\max\{Y_{C,i},Y_{P,i}\}=0)} w(i)^{I(\max\{Y_{C,i},Y_{P,i}\}=1)\eta},$$

where $w(i) = H_{CP}(d(i))$.

The dose selection process follows the same staged approach specified previously. Once heterogeneity has been achieved for both Y_C and Y_P, we can use the MLE to estimate β and η, and the recommended dose for the $(n+1)$th patient is

$$\tilde{d}_{**,n} = \min\left\{\text{argmin}_{d\in\mathcal{D}}\left(|H_C(d)^{\hat{\beta}_n} - \theta|\right), \text{argmin}_{d\in\mathcal{D}}\left(|H_{CP}(d)^{\hat{\eta}_n} - \psi|\right)\right\}. \tag{14}$$

To prove the consistency of $\tilde{d}_{**,n}$, we define

$$B_{CP} = \left\{\eta : d_{CP,*} = \arg\min_{d\in\mathcal{D}} |H_{CP}(d)^{\eta} - \psi|\right\}.$$

Also define η_{0l} such that $P(Y_C = 1 \text{ or } Y_P = 1|d_l) = H_{CP}(d_l)^{\eta_{0l}}$ for $l = 1, \ldots, m$.

Theorem 3. *Under working models (4) and (13), assume that $\{\beta_{01}, \ldots, \beta_{0m}\} \subset B_C$ and $\{\eta_{01}, \ldots, \eta_{0m}\} \subset B_{CP}$. Then, $\tilde{d}_{**,n} \to d_{**}$ almost surely as $n \to \infty$.*

The proof of Theorem 3 is similar to that for Theorem 1 and thus is omitted.

2.2.4 | Summary of designs

The three proposed approaches are compared in Table 1 in terms of their outcomes, constraints, models, and skeletons. All approaches have the constraint on Y_C. In addition, the PRO-CRM using joint outcomes, namely, PRO-CRM-joint and PRO-CRM-joint/marginal, imposes a constraint on either the clinician or patient reporting a toxicity, whereas the PRO-CRM-marginal imposes a constraint directly on Y_P. The PRO-CRM-joint and PRO-CRM-joint/marginal use the same outcomes and constraints but are modeled differently.

TABLE 1 Summary of designs

Name	Outcomes	Constraints	Model	Skeleton
PRO-CRM-marginal	Y_C; Y_P	$P(Y_C = 1\|d) \leq \theta$; $P(Y_P = 1\|d) \leq \phi$	u^{β}; v^{γ}	Different
PRO-CRM-joint	Y_C; $\max\{Y_C, Y_P\}$	$P(Y_C = 1\|d) \leq \theta$; $P(\max\{Y_C, Y_P\} = 1\|d) \leq \psi$	u^{β_1}; $u^{\beta_1+\beta_2}$	Same
PRO-CRM-joint/marginal	Y_C; $\max\{Y_C, Y_P\}$	$P(Y_C = 1\|d) \leq \theta$; $P(\max\{Y_C, Y_P\} = 1\|d) \leq \psi$	u^{β}; w^{η}	Different

Abbreviations: PRO-CRM, patient-reported outcomes continual reassessment method.

3 | APPLICATION TO THE DESIGN OF A DOSE-FINDING TRIAL

3.1 | The study design for the Bortezomib trial

We illustrate the value of this approach in the context of a dose-finding study with the goal of determining the MTD of Bortezomib when administered in combination with CHOP + Rituximab (CHOP-R).[18] The clinician DLT was defined as grade 3 or higher neurologic toxicity, symptomatic neuropathy, platelet count, nonneurologic or nonhematologic toxicity in the first cycle of treatment. The target probability of clinician defined DLT was 0.25. Five doses of Bortezomib were evaluated with a total of 18 patients using the CRM. Thus, during the design of this study only clinician-reported outcomes were considered.

A subsequent publication has shown that Bortezomib is not well tolerated at the currently approved dose.[19] While the PRO-CTCAE was not available at the time of the design of this trial, physicians postulated that it is possible that having the PRO-CTCAE would have revealed issues with tolerability from the patient's perspective and led to the recommendation of a different dose. In this section, we use the general setting of this Bortezomib trial for our numerical studies and assume the availability of patient-reported DLT.

3.2 | Simulation setup

Assuming the setting of the Bortezomib study, we consider five dose levels $d_1 < \cdots < d_5$. To incorporate patient-reported DLT, we apply the three proposed methods. The constraints for the PRO-CRM marginal are $P(Y_C = 1|d) \leq \theta = 0.25$ and $P(Y_P = 1|d) \leq \phi = 0.35$ to impose an additional constraint on patient-reported DLT. For the methods with joint outcome, the same constraint for clinician-reported DLT $P(Y_C = 1|d) \leq \theta = 0.25$ and in addition we impose a constraint on either the clinician or patient reporting a DLT, that is, $P(Y_C = 1 \text{ or } Y_P = 1|d) \leq \psi = 0.50$. The sample sizes of 18 and 40 are considered. Forty is chosen to evaluate the effect of sample size on the operating characteristics of the proposed methods. The best guess of the MTD at the start of the trial is dose level 3. The skeleton for each approach is calibrated using the method by Lee and Cheung, which provides a systematic approach for selecting a skeleton with good operating characteristics.[20] For the clinician constraint alone, given five dose levels, a target DLT rate of 0.25, and assuming the MTD is dose level 3, the skeleton $(u_1, \ldots, u_5)$ is $(0.02, 0.10, 0.25, 0.44, 0.62)$ for $N = 18$ and $(0.06, 0.14, 0.25, 0.38, 0.50)$ for $N = 40$, respectively. For the patient constraint alone, given five dose levels, a target DLT rate of 0.35, and assuming the MTD is dose level 3, the skeleton $(v_1, \ldots, v_5)$ is $(0.06, 0.18, 0.35, 0.53, 0.68)$ for $N = 18$ and $(0.10, 0.21, 0.35, 0.49, 0.61)$ for $N = 40$, respectively. For the clinician-patient joint outcome based on target DLT rate of 0.50, the skeleton $(w_1, \ldots, w_5)$ is $(0.17, 0.33, 0.50, 0.65, 0.76)$ for both $N = 18$ and $N = 40$. As noted previously, different skeletons are chosen when implementing marginal modeling, while one skeleton is used for the joint modeling approach. The starting dose is dose level 1 for all methods.

For all methods, the rule-based approach in the first stage uses a cohort size of one. If no DLT is observed, the dose is escalated. However, if a DLT is observed the same dose is assigned. Dose skipping is not allowed nor dose escalation immediately after a DLT is observed.

Seven scenarios are considered. For scenarios 1, 2, 3, and 4, the MTD d_* which incorporates both clinician- and patient-reported outcomes is the same as $d_{C,*}$, the MTD based solely on clinician outcome. These scenarios evaluate the performance when the additional constraint for patient-reported outcomes is not needed to ensure comparable performance to the CRM. In scenarios 5, 6, and 7, d_* is less than $d_{C,*}$. Thus, the objective of the CRM is different than that of the PRO-CRM. Consequently, we would not expect the CRM to select a dose level that incorporates patient-reported outcomes. These scenarios were chosen solely to evaluate the performance of the PRO-CRM at targeting a dose d_* that is one or two dose levels below $d_{C,*}$. To be able to compare the methods using marginal outcomes and joint outcomes, these seven scenarios are chosen so that $d_* = d_{**}$. That is, the MTD under the marginal model is the same as that under the

TABLE 2 Percentage of recommendations of each dose level using the marginal, joint, and joint/marginal methods with N=18

	Dose level				
	1	**2**	**3**	**4**	**5**
Scenario 1					
Probability of DLT by clinician	0.05	0.05	**0.25**	0.40	0.55
Probability of DLT by patient	0.17	0.18	**0.35**	0.50	0.65
Probability of DLT by clinician/patient	0.20	0.20	**0.50**	0.65	0.80
% recommended by CRM	1	13	**58**	25	3
% recommended by PRO-CRM-marginal	6	30	**57**	7	0
% recommended by PRO-CRM-joint	1	19	**64**	15	1
% recommended by PRO-CRM-joint/marginal	1	19	**64**	15	1
Scenario 2					
Probability of DLT by clinician	0.05	**0.25**	0.40	0.55	0.70
Probability of DLT by patient	0.10	0.15	**0.35**	0.50	0.65
Probability of DLT by clinician/patient	0.10	0.30	**0.50**	0.65	0.80
% recommended by CRM	11	**60**	26	3	0
% recommended by PRO-CRM-marginal	11	**62**	25	2	0
% recommended by PRO-CRM-joint	11	**60**	26	3	0
% recommended by PRO-CRM-joint/marginal	11	**60**	26	3	0
Scenario 3					
Probability of DLT by clinician	0.01	0.02	0.05	0.10	**0.25**
Probability of DLT by patient	0.04	0.09	0.17	0.20	**0.35**
Probability of DLT by clinician/patient	0.05	0.10	0.20	0.25	**0.50**
% recommended by CRM	0	0	3	22	**75**
% recommended by PRO-CRM-marginal	0	2	13	38	**47**
% recommended by PRO-CRM-joint	0	0	5	29	**66**
% recommended by PRO-CRM-joint/marginal	0	0	4	28	**68**
Scenario 4					
Probability of DLT by clinician	0.02	0.05	0.10	**0.25**	0.40
Probability of DLT by patient	0.09	0.17	0.20	**0.35**	0.50
Probability of DLT by patient	0.10	0.20	0.25	**0.50**	0.65
% recommended by CRM	0	2	21	**53**	24
% recommended by PRO-CRM-marginal	1	11	37	**45**	6
% recommended by PRO-CRM-joint	0	4	29	**53**	14
% recommended by PRO-CRM-joint/marginal	0	3	28	**55**	14

Abbreviations: CRM, continual reassessment method; DLT, dose-limiting toxicity; PRO-CRM, patient-reported outcomes CRM.

joint model. If d_* is not the same as d_{**}, the objectives for the two methods are different and consequently they would target different dose levels.

At each dose d, true probabilities $P(Y_C = 1|d)$ (probability of DLT by clinician), $P(Y_P = 1|d)$ (probability of DLT by patient), and $P(\max\{Y_C, Y_P\} = 1|d)$ (probability of DLT by clinician or patient) are specified so that they are nondecreasing functions of d (see Tables 2 and 3). Data (Y_C, Y_P) were generated from a multinomial distribution based on the three probabilities provided in Tables 2 and 3 such that $P(Y_C = 0, Y_P = 0|d) = 1 - P(\max\{Y_C, Y_P\} = 1|d)$, $P(Y_C = 1, Y_P = 0|d) = P(\max\{Y_C, Y_P\} = 1|d) - P(Y_P = 1|d)$, and $P(Y_C = 0, Y_P = 1|d) = P(\max\{Y_C, Y_P\} = 1|d) - P(Y_C = 1|d)$. The same seed was used for all methods. Ten thousand simulations were performed under each scenario.

4 | SIMULATION RESULTS

4.1 | A redesign of the Bortezomib trial with patient-reported DLT

Based on the simulation setup above, we illustrate and contrast the proposed methods in the context of a single Bortezomib trial. Toxicities are generated based on scenario 5 where the true MTD d_* and d_{**} are dose level 3, and $d_{C,*}$ is

TABLE 3 Percentage of recommendations of each dose level using the marginal, joint, and joint/marginal methods with N=18

	Dose Level				
	1	**2**	**3**	**4**	**5**
Scenario 5					
Probability of DLT by clinician	0.05	0.10	0.16	**0.25**	0.40
Probability of DLT by patient	0.05	0.20	**0.35**	0.50	0.65
Probability of DLT by clinician/patient	0.10	0.30	**0.50**	0.65	0.80
% recommended by CRM	1	11	**28**	39	21
% recommended by PRO-CRM-marginal	2	31	**55**	12	0
% recommended by PRO-CRM-joint	1	21	**53**	23	2
% recommended by PRO-CRM-joint/marginal	1	21	**53**	23	2
Scenario 6					
Probability of DLT by clinician	0.05	0.18	0.20	**0.25**	0.40
Probability of DLT by patient	0.17	**0.35**	0.50	0.65	0.80
Probability of DLT by clinician/patient	0.20	**0.50**	0.65	0.80	0.90
% recommended by CRM	3	**23**	24	32	18
% recommended by PRO-CRM-marginal	21	**62**	16	1	0
% recommended by PRO-CRM-joint	10	**61**	27	2	0
% recommended by PRO-CRM-joint/marginal	11	**61**	26	2	0
Scenario 7					
Probability of DLT by clinician	0.01	0.05	0.10	0.16	**0.25**
Probability of DLT by patient	0.04	0.05	0.20	**0.35**	0.50
Probability of DLT by clinician/patient	0.05	0.10	0.30	**0.50**	0.65
% recommended by CRM	0	2	11	**28**	59
% recommended by PRO-CRM-marginal	0	3	33	**52**	12
% recommended by PRO-CRM-joint	0	2	21	**53**	24
% recommended by PRO-CRM-joint/marginal	0	2	21	**52**	25

Abbreviations: CRM, continual reassessment method; DLT, dose-limiting toxicity; PRO-CRM, patient-reported outcomes CRM.

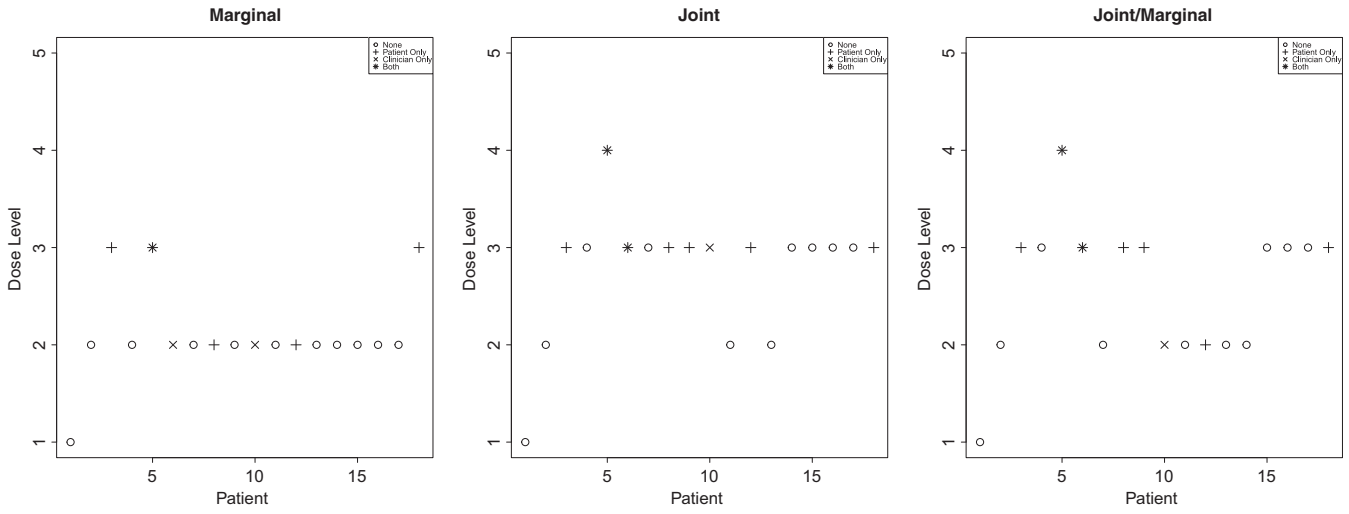

FIGURE 1 Dose assignment flow chart

dose level 4. Figure 1 displays the dose assignment for all 18 patients in the trial. The legend on the Figures specifies the presence of no DLT, patient-reported DLT only, clinician-reported DLT only, or both reporting DLT. The first panel shows the dose assignments for a simulated trial using the marginal outcome. Simulated results using the same seed for the joint outcome with the joint modeling are displayed in the second panel and with the marginally modeling in the third panel. We can see that given the same seed the assignments all follow the same algorithm until occurrence of the first patient-reported DLT (patient 3). In general, the joint and marginal modeling using joint outcome yield

similar dose recommendations, with the joint modeling assigning four more patients to the true MTD. Differences in the estimated probabilities of DLT at each dose level are more pronounced at the beginning of the trial and decrease with increase in sample size. These differences can lead to different dose recommendations, which can occur even with small differences in estimated probability of DLT. The results from the marginal approach differed from those of the joint outcome once a patient-reported DLT was observed given the different objectives. For example, after the first 3 patients, one patient-reported DLT was observed out of 1 patient at dose level 3. Thus, the marginal outcome recommended dose level 2 based on a target of 0.35 for patient-reported DLT. However, using a joint outcome with marginal modeling, the dose recommendation remained at dose level 3 given that the target for either patient or clinician-reported DLT is 0.50. The recommended dose at the end of the study was dose level three for all three approaches.

4.2 | Simulation study results

The percentage of recommendations for each dose level are summarized in Tables 2 and 3, and the percentage of patients assigned to each dose level are summarized in Tables 4 and 5, for the various scenarios with $N = 18$. Tables B1 through B4 in Appendix B are for $N = 40$. When the MTD d_* which incorporates both clinician- and patient-reported outcomes coincides with $d_{C,*}$, the MTD based solely on clinician outcome (scenarios 1, 2, 3, and 4; Tables 2 and 4), having an additional constraint tends to be more conservative and reduces the percentage of selecting a dose above the MTD. Furthermore, the joint modeling method (PRO-CRM-joint) and the marginal modeling with a joint outcome (PRO-CRM-joint/marginal) have similar or better percentage of correct selection compared to the CRM, except when the true MTD is the highest dose level. However, the difference in the probability of correct selection at the highest dose level diminishes from around 8% to 4% or less when the sample size increases from $N = 18$ to $N = 40$. The marginal modeling method (PRO-CRM-marginal) tends to be more conservative and generally performs worse than the other two proposed methods when the MTDs based on the constraints coincide, particularly when the true MTD is the median dose level or higher. The percentage of patients assigned to each dose level follows a similar trend and is presented in Tables 4 and 5.

When d_* does not coincide with $d_{C,*}$ (scenarios 5, 6, and 7; Tables 3 and 5), the CRM as expected targets $d_{C,*}$ instead of d_*. Thus, if d_* is greater than $d_{C,*}$, it recommends dose levels higher than the true MTD d_* with considerably higher probabilities. The three proposed methods, however, are all able to account for both constraints and recommend the true MTD with high probability. Percentages of correct selection higher than 55% are achieved with a sample size of $N = 18$ and 5 dose levels. The marginal modeling method (PRO-CRM-marginal) tends to recommend less frequently doses above the MTD with similar percentage of correct selection. The performance of the joint and marginal modeling approaches using the joint outcome are similar. The percentage of patients assigned to each dose level follows a similar trend. The conclusions are similar for $N = 40$.

5 | DISCUSSION

This paper proposes three novel methods, called the PRO-CRMs, to estimate the MTD using both clinician- and patient-reported toxicity data in the context of dose-finding clinical trials. While traditionally, only clinician outcomes have been considered, it is clear that patients provide a unique perspective in terms of tolerability of drugs. With the new and ongoing efforts to incorporate PROs in regulatory decision-making and to use PROs to assess treatment tolerability,[11,12] these methods will allow for the selection of doses that are well tolerated from both clinician's and patient's perspectives. These methods should only be applied with the availability of rigorous PRO measurements. In the absence of rigorous PRO measurements, it is possible that the application of the proposed methods can result in the selection of conservative doses that do not truly reflect treatment tolerability. This paper also provides a justification of running separate CRM for each toxicity component when the MTD is defined via componentwise constraints. This can be done as long as the componentwise constraints are the same as those in the joint modeling approach. Thus, the method can be applied whenever there is the desire to select a dose that satisfies two constraints. Another example is the setting where one is interested in having different toxicity thresholds for different types of toxicities (eg, a threshold of 0.10 for unusual toxicities and a threshold for 0.25 for expected toxicities).

Generally, the methods based on joint outcome perform better than modeling clinician and patient outcome separately when the constraints coincide. However, their performance was similar when the constraints led to different true MTDs. The advantage of modeling clinician and patient outcome separately is its simplicity and familiarity for clinicians in the

specification of target probabilities of DLT. With the joint outcome, which requires some knowledge regarding the correlation between clinician- and patient-reported toxicities, the performance of the joint modeling and marginal modeling where similar. However, the latter is much easier to implement and can be implemented using existing software. Thus, in practice, if investigators are comfortable specifying a constraint for the joint outcome, we would recommend the joint outcome with marginal modeling.

While the method performs well at selecting a dose that satisfies both clinician and patient constraints, further research should evaluate the appropriate target probability of toxicity based on the patient's perspective. Target probabilities of DLT in the range of 0.20 to 0.33 have been used for clinician-reported DLT. It is not clear if the same threshold should be used for patients or if higher target probabilities should be used to account for the potential greater variability in PROs, and rigor and reproducibility of the PRO instruments used. This should be assessed in conjunction with the definition of DLT used for the PRO-CTCAE and the symptom items that are selected. With more experience in the collection and analysis of the PRO-CTCAE, this may be reevaluated.

With the increase of interest in including patient-reported outcomes in clinical trials and the advent of more standardized and rigorous instruments for the collection of patient-reported outcomes, the development of clinical trial designs and analysis methods that incorporate both clinician- and patient-reported outcomes becomes important and even imperative.

TABLE 4 Percentage of patients assigned to each dose level using the marginal, joint, and joint/marginal methods with N=18

	Dose Level				
	1	**2**	**3**	**4**	**5**
Scenario 1					
Probability of DLT by clinician	0.05	0.05	**0.25**	0.40	0.55
Probability of DLT by patient	0.17	0.18	**0.35**	0.50	0.65
Probability of DLT by clinician/patient	0.20	0.20	**0.50**	0.65	0.80
% assigned by CRM	10	20	**39**	22	9
% assigned by PRO-CRM-marginal	20	32	**38**	9	1
% assigned by PRO-CRM-joint	13	26	**43**	15	3
% assigned by PRO-CRM-joint/marginal	14	25	**43**	15	3
Scenario 2					
Probability of DLT by clinician	0.05	**0.25**	0.40	0.55	0.70
Probability of DLT by patient	0.10	0.15	**0.35**	0.50	0.65
Probability of DLT by clinician/patient	0.10	0.30	**0.50**	0.65	0.80
% assigned by CRM	23	**42**	25	8	2
% assigned by PRO-CRM-marginal	25	**45**	23	6	1
% assigned by PRO-CRM-joint	24	**43**	25	7	1
% assigned by PRO-CRM-joint/marginal	24	**43**	25	7	1
Scenario 3					
Probability of DLT by clinician	0.01	0.02	0.05	0.10	**0.25**
Probability of DLT by patient	0.04	0.09	0.17	0.20	**0.35**
Probability of DLT by clinician/patient	0.05	0.10	0.20	0.25	**0.50**
% assigned by CRM	6	7	10	22	**55**
% assigned by PRO-CRM-marginal	8	11	19	29	**33**
% assigned by PRO-CRM-joint	7	8	14	28	**43**
% assigned by PRO-CRM-joint/marginal	7	8	14	27	**44**
Scenario 4					
Probability of DLT by clinician	0.02	0.05	0.10	**0.25**	0.40
Probability of DLT by patient	0.09	0.17	0.20	**0.35**	0.50
Probability of DLT by clinician/patient	0.10	0.20	0.25	**0.50**	0.65
% assigned by CRM	8	11	22	**34**	25
% assigned by PRO-CRM-marginal	13	19	31	**29**	8
% assigned by PRO-CRM-joint	9	14	28	**35**	14
% assigned by PRO-CRM-joint/marginal	9	13	29	**35**	14

Abbreviations: CRM, continual reassessment method; DLT, dose-limiting toxicity; PRO-CRM, patient-reported outcomes CRM.

TABLE 5 Percentage of patients assigned to each dose level using the marginal, joint, and joint/marginal methods with N=18

	Dose Level				
	1	2	3	4	5
Scenario 5					
Probability of DLT by clinician	0.05	0.10	0.16	**0.25**	0.40
Probability of DLT by patient	0.05	0.20	**0.35**	0.50	0.65
Probability of DLT by clinician/patient	0.10	0.30	**0.50**	0.65	0.80
% assigned by CRM	11	16	**24**	27	22
% assigned by PRO-CRM-marginal	16	34	**36**	12	2
% assigned by PRO-CRM-joint	13	27	**37**	19	4
% assigned by PRO-CRM-joint/marginal	13	26	**37**	19	5
Scenario 6					
Probability of DLT by clinician	0.05	0.18	0.20	**0.25**	0.40
Probability of DLT by patient	0.17	**0.35**	0.50	0.65	0.80
Probability of DLT by clinician/patient	0.20	**0.50**	0.65	0.80	0.90
% assigned by CRM	15	**23**	21	22	19
% assigned by PRO-CRM-marginal	36	**45**	16	3	0
% assigned by PRO-CRM-joint	25	**45**	23	6	1
% assigned by PRO-CRM-joint/marginal	26	**43**	24	6	1
Scenario 7					
Probability of DLT by clinician	0.01	0.05	0.10	0.16	**0.25**
Probability of DLT by patient	0.04	0.05	0.20	**0.35**	0.50
Probability of DLT by clinician/patient	0.05	0.10	0.30	**0.50**	0.65
% assigned by CRM	7	10	16	**23**	44
% assigned by PRO-CRM-marginal	8	14	32	**34**	12
% assigned by PRO-CRM-joint	8	11	25	**35**	21
% assigned by PRO-CRM-joint/marginal	8	11	25	**35**	21

Abbreviations: CRM, continual reassessment method; DLT, dose-limiting toxicity; PRO-CRM, patient-reported outcomes CRM.

Otherwise, there is the risk that these measures will be incorporated in a more ad hoc manner. This paper represents a first effort to address this issue in the context of dose-finding clinical trials.

DATA AVAILABILITY STATEMENT

The programs used to generate the data that support the findings of this study are available from the corresponding author upon reasonable request.

ORCID

Shing M. Lee https://orcid.org/0000-0001-8413-6869
Bin Cheng https://orcid.org/0000-0001-8295-2467

REFERENCES

1. National Cancer Institute. Common terminology criteria for adverse events (CTCAE) Version 4.0. http://ctep.cancer.gov. 2003.
2. Basch E, Iasonos A, McDonough T, et al. Patient-reported outcomes in cancer drug development and US regulatory review: perspectives from industry, the food and drug administration, and the patient. *J Am Med Assoc Oncol*. 2015;1:375-379.
3. Basch E. Toward a patient-centered value framework in oncology. *J Am Med Assoc*. 2016;315:2073-2074.
4. El-Khoueiry AB, Sangro B, Yau T, et al. Nivolumab in patients with advanced hepatocellular carcinoma (CheckMate 040): an open-label, non-comparative, phase 1/2 dose escalation and expansion trial. *Lancet*. 2017;389:2492-2502.
5. Basch E, Iasonos A, McDonough T, et al. Patient versus clinician symptom reporting using the national cancer institute common terminology criteria for adverse events: results of a questionnaire-based study. *Lancet Oncol*. 2006;7:903-909.
6. Fromme EK, Eilers KM, Mori M, Hsieh YC, Beer TM. How accurate is clinician reporting of chemotherapy adverse effects? A comparison with patient-reported symptoms from the quality-of-life questionnaire C30. *J Clin Oncol*. 2004;22:3485-3490.

7. Di Maio M, Gallo C, Leighl NB, et al. Symptomatic toxicities experienced during anticancer treatment: agreement between patient and physician reporting in three randomized trials. *J Clin Oncol*. 2015;33:910-915.
8. Falchook AD, Green R, Knowles ME, et al. Comparison of patient- and practitioner-reported toxic effects associated with chemoradiotherapy for head and neck cancer. *J Am Med Assoc Otolaryngol Head Neck Surg*. 2016;142:517-523.
9. Dueck AC, Mendoza TR, Mitchell SA, et al. Validity and reliability of the US National Cancer Institute's Patient-Reported Outcomes Version Of The Common Terminology Criteria for Adverse Events (PRO-CTCAE). *J Am Med Assoc Oncol*. 2015;1:1051-1059.
10. Basch E, Rogak LJ, Dueck AC. Methods for implementing and reporting patient-reported outcome (PRO) measures of symptomatic adverse events in cancer clinical trials. *Clinical Therapeutics*. 2016;38:821-830.
11. Kluetz PG, Kanapuru B, Lemery S, et al. Informing the tolerability of cancer treatments using patient-reported outcome measures: summary of an FDA and critical path institute workshop. *Value Health*. 2018;21:742-747.
12. Kluetz PG, O'Connor DJ, Soltys K. Incorporating the patient experience into regulatory decision making in the USA, Europe, and Canada. *Lancet Oncology*. 2018;19:267-274.
13. Qin R, Dueck AC, Satele D, et al. A pilot study of the patient-reported outcomes version of the common terminology criteria for adverse events (PRO-CTCAE) in a phase I clinical trial. *J Clin Oncol*. 2013;31:6587-6587.
14. Chiuzan C, Shtaynberger J, Manji G, et al. Dose-finding designs for trials of molecularly targeted agents and immunotherapies. *J Biopharm Stat*. 2017;27:477-494.
15. O'Quigley J, Pepe M, Fisher L. Continual reassessment method: a practical design for phase I clinical trials in cancer. *Biometrics*. 1990;46:33-48.
16. Lee SM, Cheng B, Cheung YK. Continual reassessment method with multiple toxicity constraints. *Biostatistics*. 2011;12:386-398.
17. Cheng B, Lee SM. On the consistency of the continual reassessment method with multiple constraints. *J Stat Plan Inference*. 2015;164:1-9.
18. Leonard JP, Furman RR, Cheung YKK, et al. Phase I/II trial of botezomib plus CHOP-Rituximab in diffuse large B cell (DLBCL) and mantle cell lymphona (MCL): phase I results. *Blood*. 2005;106:491.
19. Lee SM, Backenroth D, Cheung YK, et al. Case example of dose optimization using data from Bortezomib dose-finding clinical trials. *J Clin Oncol*. 2016;34:1395-1401.
20. Lee SM, Cheung YK. Model calibration in the continual reassessment method. *Clinical Trials*. 2009;6:227-238.
21. Shen LZ, O'Quigley J. Consistency of continual reassessment method under model misspecification. *Biometrika*. 1996;83:395-405.

How to cite this article: Lee SM, Lu X, Cheng B. Incorporating patient-reported outcomes in dose-finding clinical trials. *Statistics in Medicine*. 2019;1–16. https://doi.org/10.1002/sim.8402

APPENDIX A

Proof of Theorem 1. We first prove the following result:

$$B_C \times B_P \subset B. \tag{A1}$$

When $d_{C,*} = d_{P,*}$, we have $d_* = d_{C,*} = d_{P,*}$ and it is easy to see that $B_C \times B_P = B$. When $d_{C,*} \neq d_{P,*}$, any $(\beta, \gamma) \in B_C \times B_P$ satisfies

$$\arg\min_{d\in D} |F_C(d)^\beta - \theta| = d_{C,*}, \arg\min_{d\in D} |F_P(d)^\gamma - \phi| = d_{P,*},$$

and thus

$$\min\left\{\arg\min_{d\in D}|F_C(d)^\beta - \theta|, \arg\min_{d\in D} |F_P(d)^\gamma - \phi|\right\} = \min\left\{d_{C,*}, d_{P,*}\right\} = d_*,$$

implying that $(\beta, \gamma) \in B$, and we conclude that $B_C \times B_P \subset B$.

Let $S_n(\beta)$ denote the average marginal score function based on $\{(d(i), Y_{Ci}), i = 1, \ldots, n\}$, and the conditional expected score function given dose assignments denoted as $\tilde{s}_n(\beta) = E[s_n(\beta)|(d(1), \ldots, d(n))]$. Similarly, let $S_n(\gamma)$

denote the average marginal score function based on $\{(d(i), Y_{Pi}), i = 1, \ldots, n\}$ and $\tilde{s}_n(\gamma) = E[s_n(\gamma)|(d(1), \ldots, d(n))]$. Let β_n^* and γ_n^* be the unique solution to $\tilde{s}_n(\beta) = 0$ and $s_n(\gamma) = 0$, respectively. Then, $\{\beta_{01}, \ldots, \beta_{0m}\} \subset B_C$ implies $\beta_n^* \in B_C$ and $\{\gamma_{01}, \ldots, \gamma_{0m}\} \subset B_P$ implies $\gamma_n^* \in B_P$, and by (A1), $(\beta_n^*, \gamma_n^*) \in B_C \times B_P \subset B$. Let B_0 be a convex null of $\{\beta_{01}, \ldots, \beta_{0m}\} \times \{\gamma_{01}, \ldots, \gamma_{0m}\}$, then B_0 is a compact set, and $B_0 \subset B$ by the convexity of B_C and B_P and (A1). By Shen and O'Quigley,[21] we conclude that

$$\sup_{(\beta,\gamma)\in B_0} (|s_n(\beta) - \tilde{s}_n(\beta)| + |s_n(\gamma) - \tilde{s}_n(\gamma)|) \to 0$$

almost surely. This implies that when n large enough $(\hat{\beta}_n, \hat{\gamma}_n) \in B_0 \subset B$, hence $\hat{d}_{*,n} \to d_*$, almost surely. □

APPENDIX B

TABLE B1 Percentage of recommendations of each dose level using the marginal, joint, and joint/marginal methods with N=40

	Dose Level				
	1	**2**	**3**	**4**	**5**
Scenario 1					
Probability of DLT by clinician	0.05	0.05	**0.25**	0.40	0.55
Probability of DLT by patient	0.17	0.18	**0.35**	0.50	0.65
Probability of DLT by clinician/patient	0.20	0.20	**0.50**	0.65	0.80
% recommended by CRM	0	7	**74**	19	0
% recommended by PRO-CRM-marginal	1	18	**76**	5	0
% recommended by PRO-CRM-joint	0	8	**81**	11	0
% recommended by PRO-CRM-joint/marginal	0	9	**82**	9	0
Scenario 2					
Probability of DLT by clinician	0.05	**0.25**	0.40	0.55	0.70
Probability of DLT by patient	0.10	0.15	**0.35**	0.50	0.65
Probability of DLT by clinician/patient	0.10	0.30	**0.50**	0.65	0.80
% recommended by CRM	7	**76**	17	0	0
% recommended by PRO-CRM-marginal	6	**75**	19	0	0
% recommended by PRO-CRM-joint	7	**73**	20	0	0
% recommended by PRO-CRM-joint/marginal	7	**74**	19	0	0
Scenario 3					
Probability of DLT by clinician	0.01	0.02	0.05	0.10	**0.25**
Probability of DLT by patient	0.04	0.09	0.17	0.20	**0.35**
Probability of DLT by clinician/patient	0.05	0.10	0.20	0.25	**0.50**
% recommended by CRM	0	0	0	12	**88**
% recommended by PRO-CRM-marginal	0	0	2	27	**71**
% recommended by PRO-CRM-joint	0	0	0	14	**86**
% recommended by PRO-CRM-joint/marginal	0	0	0	16	**84**
Scenario 4					
Probability of DLT by clinician	0.02	0.05	0.10	**0.25**	0.40
Probability of DLT by patient	0.09	0.17	0.20	**0.35**	0.50
Probability of DLT by clinician/patient	0.10	0.20	0.25	**0.50**	0.65
% recommended by CRM	0	0	13	**69**	18
% recommended by PRO-CRM-marginal	0	2	26	**68**	4
% recommended by PRO-CRM-joint	0	0	14	**75**	11
% recommended by PRO-CRM-joint/marginal	0	0	14	**77**	9

Abbreviations: CRM, continual reassessment method; DLT, dose-limiting toxicity; PRO-CRM, patient-reported outcomes CRM.

TABLE B2 Percentage of recommendations of each dose level using the marginal, joint, and joint/marginal methods with N=40

	Dose Level				
	1	**2**	**3**	**4**	**5**
Scenario 5					
Probability of DLT by clinician	0.05	0.10	0.16	**0.25**	0.40
Probability of DLT by patient	0.05	0.20	**0.35**	0.50	0.65
Probability of DLT by clinician/patient	0.10	0.30	**0.50**	0.65	0.80
% recommended by CRM	0	2	**27**	56	15
% recommended by PRO-CRM-marginal	0	15	**73**	12	0
% recommended by PRO-CRM-joint	0	10	**68**	22	0
% recommended by PRO-CRM-joint/marginal	0	9	**72**	19	0
Scenario 6					
Probability of DLT by clinician	0.05	0.18	0.20	**0.25**	0.40
Probability of DLT by patient	0.17	**0.35**	0.50	0.65	0.80
Probability of DLT by clinician/patient	0.20	**0.50**	0.65	0.80	0.90
% recommended by CRM	1	**13**	28	45	13
% recommended by PRO-CRM-marginal	10	**76**	14	0	0
% recommended by PRO-CRM-joint	6	**71**	23	0	0
% recommended by PRO-CRM-joint/marginal	4	**74**	21	0	0
Scenario 7					
Probability of DLT by clinician	0.01	0.05	0.10	0.16	**0.25**
Probability of DLT by patient	0.04	0.05	0.20	**0.35**	0.50
Probability of DLT by clinician/patient	0.05	0.10	0.30	**0.50**	0.65
% recommended by CRM	0	0	2	**25**	73
% recommended by PRO-CRM-marginal	0	0	15	**73**	12
% recommended by PRO-CRM-joint	0	0	10	**67**	23
% recommended by PRO-CRM-joint/marginal	0	0	9	**72**	19

Abbreviations: CRM, continual reassessment method; DLT, dose-limiting toxicity; PRO-CRM, patient-reported outcomes CRM.

TABLE B3 Percentage of patients assigned to each dose level using the marginal, joint, and joint/marginal methods with N=40

	Dose Level				
	1	**2**	**3**	**4**	**5**
Scenario 1					
Probability of DLT by clinician	0.05	0.05	**0.25**	0.40	0.55
Probability of DLT by patient	0.17	0.18	**0.35**	0.50	0.65
Probability of DLT by clinician/patient	0.20	0.20	**0.50**	0.65	0.80
% assigned by CRM	4	15	**53**	22	6
% assigned by PRO-CRM-marginal	11	26	**53**	9	1
% assigned by PRO-CRM-joint	6	18	**58**	16	2
% assigned by PRO-CRM-joint/marginal	6	19	**59**	14	2
Scenario 2					
Probability of DLT by clinician	0.05	**0.25**	0.40	0.55	0.70
Probability of DLT by patient	0.10	0.15	**0.35**	0.50	0.65
Probability of DLT by clinician/patient	0.10	0.30	**0.50**	0.65	0.80
% assigned by CRM	17	**54**	24	4	1
% assigned by PRO-CRM-marginal	17	**55**	24	3	1
% assigned by PRO-CRM-joint	17	**54**	24	4	1
% assigned by PRO-CRM-joint/marginal	17	**54**	24	4	1
Scenario 3					
Probability of DLT by clinician	0.01	0.02	0.05	0.10	**0.25**
Probability of DLT by patient	0.04	0.09	0.17	0.20	**0.35**
Probability of DLT by clinician/patient	0.05	0.10	0.20	0.25	**0.50**
% assigned by CRM	3	3	5	18	**71**
% assigned by PRO-CRM-marginal	4	5	10	30	**51**
% assigned by PRO-CRM-joint	3	3	6	22	**66**
% assigned by PRO-CRM-joint/marginal	3	4	7	23	**63**
Scenario 4					
Probability of DLT by clinician	0.02	0.05	0.10	**0.25**	0.40
Probability of DLT by patient	0.09	0.17	0.20	**0.35**	0.50
Probability of DLT by clinician/patient	0.10	0.20	0.25	**0.50**	0.65
% assigned by CRM	3	5	19	**47**	26
% assigned by PRO-CRM-marginal	6	9	30	**47**	8
% assigned by PRO-CRM-joint	4	6	22	**52**	16
% assigned by PRO-CRM-joint/marginal	4	6	24	**52**	14

Abbreviations: CRM, continual reassessment method; DLT, dose-limiting toxicity; PRO-CRM, patient-reported outcomes CRM.

TABLE B4 Percentage of patients assigned to each dose level using the marginal, joint, and joint/marginal methods with N=40

	Dose Level				
	1	**2**	**3**	**4**	**5**
Scenario 5					
Probability of DLT by clinician	0.05	0.10	0.16	**0.25**	0.40
Probability of DLT by patient	0.05	0.20	**0.35**	0.50	0.65
Probability of DLT by clinician/patient	0.10	0.30	**0.50**	0.65	0.80
% assigned by CRM	5	9	**26**	38	22
% assigned by PRO-CRM-marginal	8	25	**52**	14	1
% assigned by PRO-CRM-joint	6	18	**49**	23	4
% assigned by PRO-CRM-joint/marginal	6	18	**51**	22	3
Scenario 6					
Probability of DLT by clinician	0.05	0.18	0.20	**0.25**	0.40
Probability of DLT by patient	0.17	**0.35**	0.50	0.65	0.80
Probability of DLT by clinician/patient	0.20	**0.50**	0.65	0.80	0.90
% assigned by CRM	8	**18**	25	31	18
% assigned by PRO-CRM-marginal	24	**57**	17	2	0
% assigned by PRO-CRM-joint	16	**53**	26	4	1
% assigned by PRO-CRM-joint/marginal	16	**56**	25	3	0
Scenario 7					
Probability of DLT by clinician	0.01	0.05	0.10	0.16	**0.25**
Probability of DLT by patient	0.04	0.05	0.20	**0.35**	0.50
Probability of DLT by clinician/patient	0.05	0.10	0.30	**0.50**	0.65
% assigned by CRM	3	4	9	**24**	60
% assigned by PRO-CRM-marginal	4	6	24	**51**	15
% assigned by PRO-CRM-joint	3	5	17	**47**	28
% assigned by PRO-CRM-joint/marginal	3	5	18	**50**	24

Abbreviations: CRM, continual reassessment method; DLT, dose-limiting toxicity; PRO-CRM, patient-reported outcomes CRM.