ESMO GOOD SCIENCE BETTER MEDICINE BEST PRACTICE

ESMO OPEN SCIENCE FOR OPTIMAL CANCER CARE

ORIGINAL RESEARCH

U-PRO-CRM: designing patient-centred dose-finding trials with patient-reported outcomes

E. Alger[1], S. M. Lee[2], Y. K. Cheung[2] & C. Yap[1]*

[1]Clinical Trial and Statistics Unit, Institute of Cancer Research, London, UK; [2]Department of Biostatistics, Mailman School of Public Health, Columbia University, New York, USA

Available online xxx

Background: Determining the maximum tolerated dose (MTD) remains the primary objective for the majority of dose-finding oncology trials. Whilst MTD determination often relies upon clinicians to identify dose-limiting toxicities (DLTs) experienced by patients during the trial, research suggests that clinicians may underreport patient's adverse events. Therefore, contemporary practice may be exposed to recommending intolerable doses to patients for further investigation in subsequent trials. There is increasing interest in patients self-assessing their own symptoms using patient-reported outcomes (PROs) in dose-finding trials.

Design: We present Utility-PRO-Continual Reassessment Method (U-PRO-CRM), a novel trial design which simultaneously uses clinician-rated and patient-rated DLTs (Clinician-DLTs and Patient-DLTs, respectively) to make dose (de-)escalation decisions and to recommend an MTD. U-PRO-CRM contains the published PRO-CRM as a special case and provides greater flexibility to trade-off the rate of Patient-DLTs and Clinician-DLTs to find an optimal dose. We present simulation results for U-PRO-CRM.

Results: For specified trade-offs between Clinician-DLT and Patient-DLT rate, U-PRO-CRM outperforms the PRO-CRM design by identifying the true MTD more often. In the special case where U-PRO-CRM generalises to PRO-CRM, U-PRO-CRM performs as well as its published counterpart. U-PRO-CRM minimises the number of patients overdosed whilst maintaining a similar proportion of patients allocated to the true MTD.

Conclusions: By using a utility-based dose selection approach, U-PRO-CRM offers the flexibility to define a trade-off between the risk of patient-rated and clinician-rated DLTs for an optimal dose. Patient-centric dose-finding strategies, which integrate PROs, are poised to assume an ever more pivotal role in significantly advancing our understanding of treatment tolerability. This bears significant implications in shaping the future landscape of early-phase trials.

Key words: patient-reported outcomes, dose-finding, maximum tolerated dose, continual reassessment method, phase I

INTRODUCTION

The aim of dose-finding trials is to determine the optimal dose or doses for further investigation in subsequent trials. In oncology, the conventional criterion for the optimal dose has been the maximum tolerated dose (MTD). Under the assumption that efficacy is likely to increase with dose, this approach looks to maximise treatment efficacy whilst safeguarding dose tolerability within a patient population.

For dose-finding trials, toxicities are often identified by clinicians using the National Cancer Institute Common Terminology Criteria for Adverse Events (NCI-CTCAE) and graded on severity.[1] Generally, toxicities identified as at least 'severe' are deemed a dose-limiting toxicity (DLT). The dose with a probability of DLT closest to some pre-specified, clinician elicited target is recognised as the MTD and often recommended as the phase II dose (RP2D).

When evaluating DLTs, toxicities are traditionally assessed solely by a clinician and thus do not reflect a patient's own personal assessment of toxicity.[2] Specific, subjective adverse events (AEs) such as nausea and fatigue can be difficult for a clinician to grade, and can be under-reported by clinicians.[3] Whilst the NCI-CTCAE is used by clinicians to grade life-threatening AEs, patients may also be concerned about other side-effects which impact their quality of life. Previous reviews have suggested that patients' most feared AEs often differ from the AEs which concern clinicians the most when assessing DLTs.[4]

*Correspondence to: Prof. Christina Yap, Clinical Trials and Statistics Unit, The Institute of Cancer Research, Sir Richard Doll Building, 15 Cotswold Road, London, SM2 5NG, UK. Tel: +44-20-8722-4022
E-mail: christina.yap@icr.ac.uk (C. Yap).

There is increasing interest in the leveraging of patient-reported outcomes (PROs) within dose-finding trials in order to assess treatment tolerability. Defined by the US Department of Health, a PRO is 'any report of the status of a patient's health condition that comes directly from the patient, without interpretation of the patient's response by a clinician or anyone else'.[5] Research has shown that patients often report symptomatic AEs earlier and more frequently than clinicians[6] and there is significant disparity between patient- and clinician-assessed AEs. Evidence is emerging that the number of dose discontinuations due to AEs within real world administration of a treatment may be disproportionately higher than that identified within later-phase clinical trials.[7] In light of such research, the US Food and Drug Administration's (FDA) Project Optimus initiative is encouraging the use of both clinical and non-clinical data to aid dose optimisation during drug development and investigation.[8] Whilst many dose-finding trial designs assess tolerability using clinician-assessed DLTs, there is limited research exploring the extension of these designs to incorporate PROs. Current research has indicated infrequent utilisation of PROs (5.3%) in dose-finding trials.[9] Among trials that have analysed PROs, their incorporation typically occurred solely at the conclusion of the trial, confirming the tolerability of the MTD, rather than being integrated into the interim dose escalation and de-escalation component.[10] Taking inspiration from the Continual Reassessment Method (CRM) design,[11] the PRO-CRM design[12] and the time-to-event (TiTE) extension TiTE-PRO-CRM[13] are two of only three model-based trial designs which incorporate patient and clinician toxicity information in the determination of the MTD.

In this paper we introduce the Utility-PRO-CRM (U-PRO-CRM) design, a novel model-based trial design which generalises the two-step dose selection decision of the PRO-CRM and TiTE-PRO-CRM using a utility curve.

In 'Methods' we present the U-PRO-CRM using a likelihood framework. In 'Numerical Study' we present a numerical study to assess how U-PRO-CRM performs.

METHODS

Endpoints

Like the PRO-CRM design, the U-PRO-CRM design utilises binary endpoints to define a clinician-assessed DLT (Clinician-DLT) and patient-assessed DLT (Patient-DLT). A Clinician-DLT is routinely defined within phase I trials, and often identified as at least a grade 4 haematological toxicity or grade 3 non-haematological toxicity using the NCI-CTCAE.[14] There is currently no standardised definition for Patient-DLT, although new standardised PRO measures such as the PRO-CTCAE have been constructed for patients to assess their AEs and directly complements the CTCAE used by clinicians to assess toxicities.[15] In two trials, examples of Patient-DLT definitions include a 'severe' or 'very severe' gastrointestinal toxicity as defined by the PRO-CTCAE;[16] and a 10-point rise and/or a 15-point rise in the O'Leary Interstitial Cystitis scale and AUA Symptom Score PROMs.[17]

Utility PRO-CRM (U-PRO-CRM)

Building on PRO-CRM. We will begin by introducing the PRO-CRM design, as the U-PRO-CRM design is an extension of it. U-PRO-CRM uses the same model estimation method as the PRO-CRM.[12] In this instance, two one-parameter empirical models are used to independently estimate the outcomes for patients (Patient-DLT) and clinicians (Clinician-DLT), utilising a two-stage maximum likelihood CRM approach.

In the first stage, whilst there are no differences in DLT outcomes as rated by patients and clinicians (i.e. whilst no Clinician-DLTs or Patient-DLTs have been observed), a simple rule-based design is utilised for dose-escalation. Once a DLT is observed (as rated by patients or clinicians), the model parameter for that endpoint (Patient-DLT or Clinician-DLT) is estimated using the model, whilst the rule-based design is utilised for the other endpoint. At the next interim analysis, we minimise the dose recommended by the model and rule-based design for the next cohort of patients. When DLTs have been observed for both endpoints, we solely use the models to separately estimate the rate of Patient-DLT and Clinician-DLT for each dose.

More detail for the models is provided in Section 1 of the Supplementary Materials, available at https://doi.org/10.1016/j.esmoop.2024.103626.

Dose selection using a utility-based approach. Once the Clinician-DLT and Patient-DLT rates are estimated, U-PRO-CRM recommends a dose using a utility curve (as shown in Figure 1). Inspired by the EffTox design,[18] for target Clinician-DLT ($\tilde{\pi}_C$) and Patient-DLT ($\tilde{\pi}_P$) rates the utility curve is defined as,

$$f(\pi_P) = \tilde{\pi}_C\left(1 - \left(\frac{\pi_P}{\tilde{\pi}_P}\right)^{\alpha}\right)^{\frac{1}{\alpha}} \quad (1)$$

where π_P denotes Patient-DLT rate and α denotes the trade-off between Clinician-DLT and Patient-DLT rate which remains fixed throughout the trial. At each interim analysis, U-PRO-CRM recommends the next cohort's dose such that its distance (δ_d) between the estimated Clinician-DLT and Patient-DLT rates and the utility curve is smaller than for all other doses, using accumulated outcomes. This distance δ_d is the minimal Euclidean (straight line) distance between the utility curve and dose d's estimated Clinician-DLT and Patient-DLT rates and is computed as:

$$\delta_d = \min_{\pi_P \in [0,\tilde{\pi}_P]} \sqrt{\left(\pi_P - \widehat{\pi}_{P,d}\right)^2 + \left(f(\pi_P) - \widehat{\pi}_{C,d}\right)^2} \quad (2)$$

where $\widehat{\pi}_{C,d}$ and $\widehat{\pi}_{P,d}$ are the estimated Clinician-DLT and Patient-DLT rates for dose d, respectively. To recommend the MTD at the end of the trial, the U-PRO-CRM design recommends the MTD similarly to how it recommends the

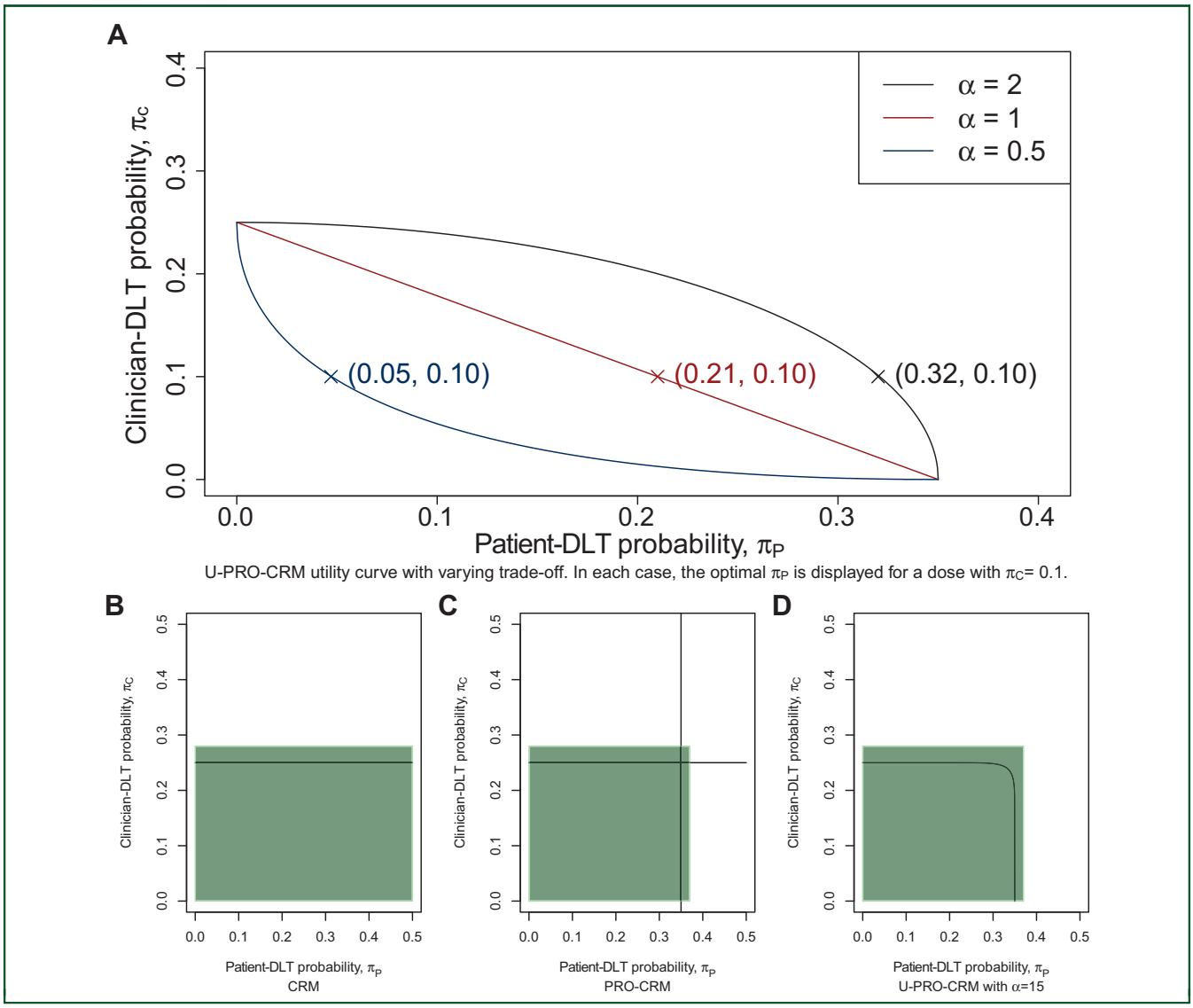

Figure 1. Dose selection curves for the CRM, PRO-CRM, and U-PRO-CRM trial designs where α denotes the trade-off between Clinician-DLT and Patient-DLT for the U-PRO-CRM design. Target patient and clinician DLT rates are set to 0.35 and 0.25, respectively. Regions of admissible doses are shaded in green. CRM, Continual Reassessment Method; DLT, dose-limiting toxicity; PRO, patient-reported outcome; PRO-CRM, PRO-Continual Reassessment Method; U-PRO-CRM, Utility-PRO-Continual Reassessment Method.

next dose in interim analyses—by selecting the dose with estimated Patient-DLT and Clinician-DLT rates that are closest to the utility curve. A flow diagram illustrating the U-PRO-CRM trial design is presented in Supplementary Figure S1 of the Supplementary Materials, available at https://doi.org/10.1016/j.esmoop.2024.103626.

Optimising Clinician-DLT and Patient-DLT trade-off. The target Clinician-DLT rate, as elicited by a clinician when Patient-DLT is assumed zero, is defined as $\tilde{\pi}_C$. Similarly, the target Patient-DLT rate as elicited by a patient when Clinician-DLT rate is assumed to be zero, is defined as $\tilde{\pi}_P$. The parameter α controls the shape of the utility curve between the points ($\tilde{\pi}_P$,0) and (0, $\tilde{\pi}_C$), and can intuitively be considered as the trade-off of Clinician-DLT rate as Patient-DLT rate increases.

The target Clinician-DLT rate $\tilde{\pi}_C$ occurs when the Patient-DLT rate is zero. As the rate of Patient-DLT increases, the optimal Clinician-DLT rate decays away to compromise the rate of both a patient and clinician DLT occurring. When $\alpha = 1$, this decay is linear. The decay is more rapid when $\alpha < 1$ and the decay is less rapid when $\alpha > 1$. This is presented graphically in Figure 1A where $\tilde{\pi}_C = 0.25, \tilde{\pi}_P = 0.35$. In this example, for an estimated Clinician-DLT rate of 0.1, the optimal Patient-DLT rate is largest (0.32) for $\alpha = 2$ and smallest (0.05) for $\alpha = 0.5$.

Comparison to PRO-CRM. PRO-CRM first minimises the distance between doses' estimated Clinician-DLT and Patient-DLT rates to their respective targets. After a clinician dose (recommended using the clinician target) and patient dose (recommended using the patient target) are identified, the smaller of the two doses is then administered to the next cohort of patients.

For α large, U-PRO-CRM can be considered a generalisation of the PRO-CRM design, with the key distinction that

it compresses the two-step decision rule into a single step. A comparison of the decision rules (and utility shapes) associated with the CRM, PRO-CRM, and U-PRO-CRM are presented in Figure 1B-D. In Figure 1B-D, the region of admissible doses is highlighted—graphically showing how the PRO-CRM and U-PRO-CRM adds an additional restriction on admissible doses compared with the CRM design. In the simulation study to follow in 'Numerical Study', we compare the U-PRO-CRM to the PRO-CRM when α is chosen to be 15.

Whilst PRO-CRM supposes that trialists want to identify the dose which maximises the DLT rate with respect to either the clinician or patient target, U-PRO-CRM allows for increased flexibility. U-PRO-CRM serves PRO-CRM's objective as a special case, and allows trialists the flexibility to choose an MTD with respect to any other predefined trade-off they deem optimal.

Elicitation of trade-off between Patient-DLT and Clinician-DLT

The parameter α determines the curvature of the utility curve and is fixed throughout the trial and at every interim analysis. This, in turn, guides the decision on the optimal dose for the next cohort of patients, and influences the number of patients who may experience a DLT during the trial. For a dose with some given rate of Clinician-DLT, the optimal Patient-DLT rate required to ensure this dose lies on the utility curve depends on α. Under a simulation scenario where there is a moderate correlation between Patient-DLT and Clinician-DLT endpoints (see 'Simulating correlated Clinical-DLT and Patient-DLT binary outcomes'), suppose one dose has an estimated Clinician-DLT rate of 0.20. We can then estimate what proportion of patients experience at least one DLT (Clinician-DLT and/or Patient-DLT) for varying α values. The value of α can then be chosen to ensure that the total proportion of patients who experience a Clinician-DLT and/or a Patient-DLT is kept below a certain level. One such example is given in Table 1. For target Clinician-DLT rate 0.20 and by varying α, the rate that a patient experiences at least one DLT can be estimated using simulations. Under the scenario where a dose has a Clinician-DLT rate of 0.20, a threshold may be set such that no more than 35% of patients should be unable to tolerate treatment due to a Clinician-DLT, Patient-DLT or both. In this specific scenario, $\alpha = 2$ should be used to define the utility curve.

Table 1. Estimated probability a patient experiences a Patient-DLT or a Clinician-DLT and/or Patient-DLT when the target Clinician-DLT rate is 0.20 under a simulation study of 100 000 patients

Trade-off parameter, α	Probability that a patient experiences Clinician-DLT	Probability that a patient experiences Patient-DLT	Probability that a patient experiences Patient-DLT, and/or Clinician-DLT
0.5	0.2	0.004	0.226
1.0	0.2	0.070	0.244
2.0	0.2	0.210	0.337
15.0	0.2	0.350	0.438

Moderate correlation is induced using a Clayton model with $\varphi = 0.9$, see Section 3 of the Supplementary Materials, available at https://doi.org/10.1016/j.esmoop.2024.103626. DLT, dose-limiting toxicity.

NUMERICAL STUDY

We use computer simulations to compare U-PRO-CRM with different α (2 and 15) to PRO-CRM. The trials are simulated over nine simulation scenarios, each with 5000 simulations. Seven scenarios match those of the original PRO-CRM simulation study which was based on a phase I study of bortezomib.[19] Two additional scenarios are added to evaluate the effect of the stopping rule: one where the lowest dose is the MTD, and another where no dose is tolerable.

As well as comparing U-PRO-CRM performance with the PRO-CRM design, we also compare the U-PRO-CRM design to the simulation-based benchmark proposed by Cheung.[20] Thorough details of this benchmark approach are presented in Section 4 of the Supplementary Materials, available at https://doi.org/10.1016/j.esmoop.2024.103626.

Fixed trial parameters

We simulate trials with the aim of identifying an MTD for a single-agent treatment across five increasing doses. Patients are enrolled sequentially in cohorts of size three, and assessed for DLTs within a 1 month window. All DLTs are observed before the next cohort of patients are enrolled in the trial. The maximum sample size for the trial is 39 patients. Similar to the simulation study for PRO-CRM,[12] the target Clinician-DLT rate is set to 0.25 and the target Patient-DLT rate is set to 0.35.

Model skeletons. For the U-PRO-CRM design, we assume that dose 3 is the *a priori* guess of where the true MTD is. The same model skeletons (displayed in Table 2) are used as those from the PRO-CRM simulation study.[12]

Simulating correlated Clinical-DLT and Patient-DLT binary outcomes. To generate correlated Clinician-DLT and Patient-DLT outcomes, we use a Clayton model with Cox-exponential proportional hazard survival models.[21] Correlation is induced using parameter φ, with $\varphi \to 0$ inducing a strong positive correlation, and $\varphi \to \infty$ inducing no correlation.

The value of $\varphi = 0.9$ is chosen to induce a moderate positive correlation between Clinician-DLT and Patient-DLT rate.[21] Additional details of the Clayton model are presented in Section 3 of the Supplementary Materials, available at https://doi.org/10.1016/j.esmoop.2024.103626.

Safety stopping rule. A beta-binomial stopping rule is introduced to stop a trial early if the composite endpoint that the risk of any DLT (Clinician-DLT and/or Patient-DLT) at the lowest dose is too toxic (above a pre-specified level).

Under simulations where $\varphi = 0.9$, ~90% of trials are stopped prematurely when no dose is safe and ~30% of trials are stopped prematurely when dose 1 is safe. Further details are provided in Section 5 of the Supplementary

Table 2. Proportion of MTD recommendations for each dose under nine scenarios with $\alpha = 2$ and $\varphi = 0.9$ over 5000 simulations for U-PRO-CRM, PRO-CRM, and benchmark

		Dose level				
		1	2	3	4	5
Skeletons	Model skeleton for Clinician-DLT	0.06	0.14	0.25	0.38	0.50
	Model skeleton for Patient-DLT	0.10	0.21	0.35	0.49	0.61
Scenario 1	True probability of Clinician-DLT	0.05	0.05	**0.25**	0.40	0.55
	True probability of Patient-DLT	0.17	0.18	**0.35**	0.50	0.65
	U-PRO-CRM dose recommendation (%)	3	48	**47**	0	0
	PRO-CRM dose recommendation (%)	1	20	**72**	5	0
	Benchmark dose recommendation (%)	34	13	**52**	0	0
Scenario 2	True probability of Clinician-DLT	0.05	**0.25**	0.40	0.55	0.70
	True probability of Patient-DLT	0.10	**0.15**	0.35	0.50	0.65
	U-PRO-CRM dose recommendation (%)	12	**82**	6	0	0
	PRO-CRM dose recommendation (%)	6	**72**	21	0	0
	Benchmark dose recommendation (%)	6	**91**	3	0	0
Scenario 3	True probability of Clinician-DLT	0.01	0.02	0.05	**0.10**	0.25
	True probability of Patient-DLT	0.04	0.09	0.17	**0.20**	0.35
	U-PRO-CRM dose recommendation (%)	0	0	10	**58**	32
	PRO-CRM dose recommendation (%)	0	0	3	**30**	66
	Benchmark dose recommendation (%)	0	0	3	**59**	37
Scenario 4	True probability of Clinician-DLT	0.02	0.05	**0.10**	0.25	0.40
	True probability of Patient-DLT	0.09	0.17	**0.20**	0.35	0.50
	U-PRO-CRM dose recommendation (%)	0	7	**61**	32	1
	PRO-CRM dose recommendation (%)	0	2	**29**	63	5
	Benchmark dose recommendation (%)	0	4	**58**	37	0
Scenario 5	True probability of Clinician-DLT	0.05	0.10	**0.16**	0.25	0.40
	True probability of Patient-DLT	0.05	0.20	**0.35**	0.50	0.65
	U-PRO-CRM dose recommendation (%)	0	40	**56**	3	0
	PRO-CRM dose recommendation (%)	0	14	**69**	17	0
	Benchmark dose recommendation (%)	0	42	**56**	2	0
Scenario 6	True probability of Clinician-DLT	0.05	**0.18**	0.20	0.25	0.40
	True probability of Patient-DLT	0.17	**0.35**	0.50	0.65	0.80
	U-PRO-CRM dose recommendation (%)	28	**68**	4	0	0
	PRO-CRM dose recommendation (%)	9	**72**	18	0	0
	Benchmark dose recommendation (%)	26	**71**	2	0	0
Scenario 7	True probability of Clinician-DLT	0.01	0.05	0.10	**0.16**	0.25
	True probability of Patient-DLT	0.04	0.05	0.20	**0.35**	0.50
	U-PRO-CRM dose recommendation (%)	0	0	41	**55**	3
	PRO-CRM dose recommendation (%)	0	0	15	**67**	18
	Benchmark dose recommendation (%)	0	0	42	**56**	2
Scenario 8	True probability of Clinician-DLT	**0.25**	0.40	0.55	0.70	0.80
	True probability of Patient-DLT	**0.35**	0.50	0.65	0.80	0.85
	U-PRO-CRM dose recommendation (%)	**66**	0	0	0	0
	PRO-CRM dose recommendation (%)	**62**	5	0	0	0
	Benchmark dose recommendation (%)	**100**	0	0	0	0
Scenario 9	True probability of Clinician-DLT	0.45	0.50	0.55	0.70	0.80
	True probability of Patient-DLT	0.55	0.60	0.65	0.80	0.85
	U-PRO-CRM dose recommendation (%)	6	0	0	0	0
	PRO-CRM dose recommendation (%)	6	0	0	0	0
	Benchmark dose recommendation (%)	NA	NA	NA	NA	NA

The MTD under each scenario is presented in bold. Admissible doses within a distance of 0.15 from the utility curve are presented in yellow. CRM, Continual Reassessment Method; DLT, dose-limiting toxicity; MTD, maximum tolerated dose; PRO, patient-reported outcome; PRO-CRM, PRO-Continual Reassessment Method; U-PRO-CRM, Utility-PRO-Continual Reassessment Method.

Materials, available at https://doi.org/10.1016/j.esmoop.2024.103626.

The trial ends once 39 patients have been assessed for DLTs or the stopping rule is executed, whatever occurs earlier.

Hypothetical trial example

To illustrate how U-PRO-CRM recommends different doses under distinct α values, we detail the conduct of a hypothetical phase I trial in Figure 2. The trial is simulated as presented in 'Fixed trial parameters'. Patient DLT data are simulated as per Cheung[20] to generate a complete outcome profile for each patient in the trial.

The true Clinician-DLT and Patient-DLT rates over five increasing doses follow those of scenario 4 as detailed in Table 2 with the following true Clinician-DLT rates (0.02, 0.05, 0.10, 0.25, 0.40) and Patient-DLT rates (0.09, 0.17, 0.20, 0.35, 0.50) for doses 1 to 5. Dose 3 and dose 4 are the true MTDs for the U-PRO-CRM when α is equal to 2 and 15, respectively.

The trial commences by treating the first cohort of three patients at dose 1, where no patient experiences a DLT. The first stage of U-PRO-CRM continues with a rule-based approach, with the next cohort of three patients given dose 2, where a Patient-DLT is experienced. Thus Patient-DLT rates are estimated using a likelihood CRM from cohort 2 onwards, whereas the straightforward rule-based

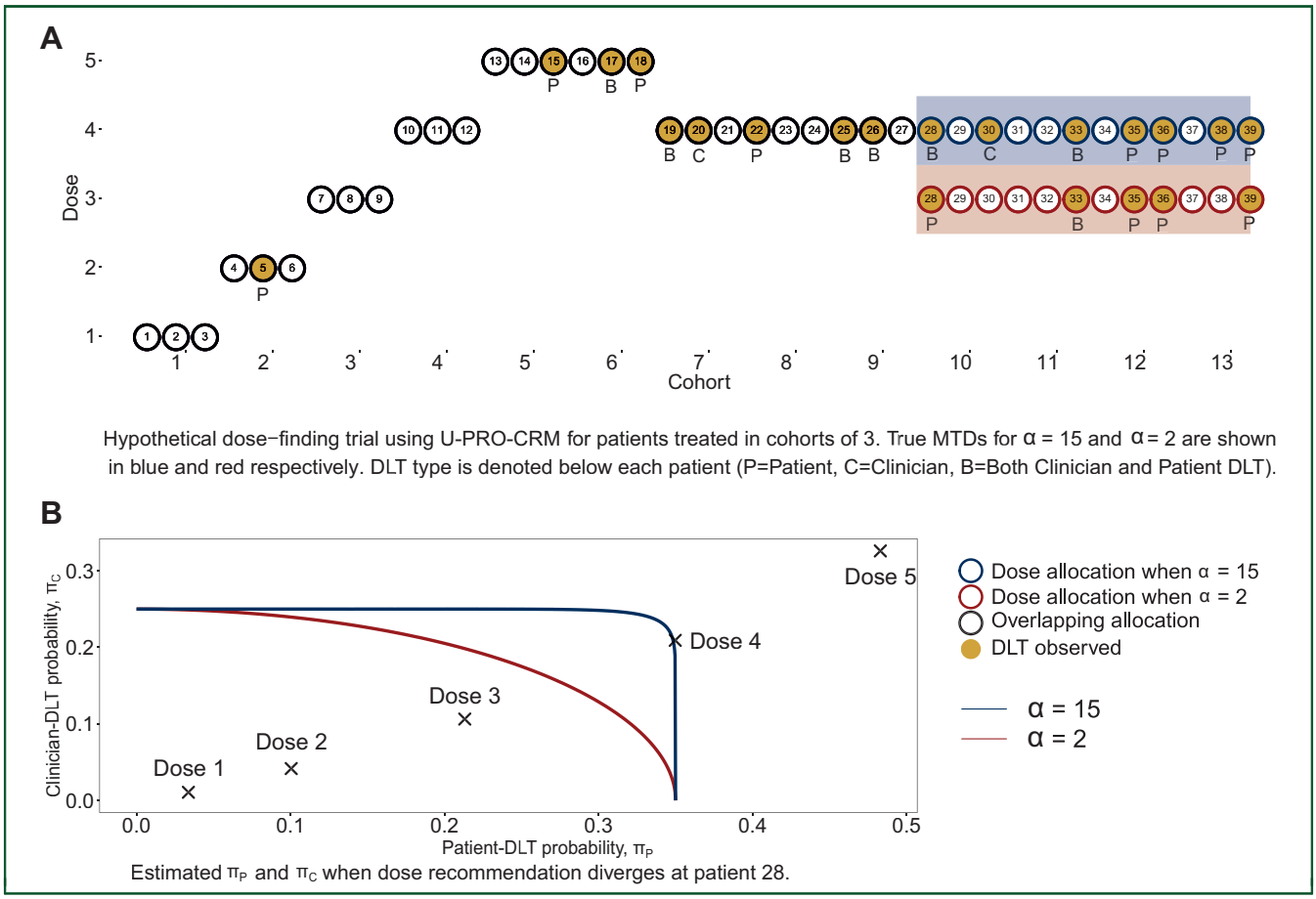

Figure 2. Hypothetical case study using the U-PRO-CRM design within a simulated clinical trial where α denotes the trade-off between Clinician-DLT and Patient-DLT for the U-PRO-CRM design.
DLT, dose-limiting toxicity; MTD, maximum tolerated dose; PRO, patient-reported outcome; U-PRO-CRM, Utility-PRO-Continual Reassessment Method.

escalation approach for Clinician-DLT continues. Patient 17 in cohort 6 experiences a Clinician-DLT, thus from cohort 7 onwards, U-PRO-CRM estimates both Clinician-DLTs and Patient-DLTs using a likelihood CRM and utilises the utility curve to guide dose escalation. Figure 2A graphically displays the escalation and de-escalation of doses within the hypothetical trial under two decision rules—when α is equal to 15 or 2. Until patient 27, U-PRO-CRM under both α values recommends patients be allocated to the same dose. Following the DLT observations of the ninth cohort (which includes two DLTs), the dose selection decision diverges depending on which α is used to define the trade-off between Clinician-DLT and Patient-DLT. Figure 2B presents the estimated Clinician-DLT and Patient-DLT rate for each dose up until the ninth cohort, as well as the utility curves defined for α equal to 2 and 15. For a fixed Patient-DLT rate, when $\alpha = 2$, U-PRO-CRM will trade off a greater rate of Clinician-DLT compared with when $\alpha = 15$. From the 10th cohort onwards, U-PRO-CRM recommends different doses for patients depending on the value of α. U-PRO-CRM recommends doses 4 and 3 consistently (including the eventual MTD) for α equals 15 or 2, respectively. Here, we can observe that as we reduce α we further constrain the selection of the final MTD.

RESULTS

Flexible utility design

As well as generalising PRO-CRM (as shown in 'Comparison to PRO-CRM design'), U-PRO-CRM has the supplementary benefit of providing additional flexibility when it comes to dose selection. The choice of α allows for a more flexible compromise between the target Clinician-DLT and Patient-DLT rate. In this section, we consider the operating characteristics of the U-PRO-CRM model when α is chosen to be 2.

In comparison to PRO-CRM (and the special U-PRO-CRM case when $\alpha = 15$), setting $\alpha = 2$ considerably reduces the acceptable Clinician-DLT rate as the estimated Patient-DLT rate increases. Thus, for $\alpha = 2$, the utility curve will always choose a dose which is at most that recommended by the utility curve for $\alpha = 15$. In simulation scenarios 1, 3, 4, 5, 6, and 7, more than one dose can be considered optimal as multiple doses' Clinician-DLT and Patient-DLT rates are approximately equidistant from the utility curve.

Thus, we define an equivalence region ϵ around the utility function. Within this region, multiple doses could be considered optimal. In Table 2, we present operating characteristics of the U-PRO-CRM method when the utility

curve is specified as $\alpha = 2$. An equivalence region around the utility curve of $\epsilon = 0.15$ is defined to allow for the selection of multiple optimal doses within a distance of 0.15 from the utility curve. Supplementary Figure S2 in the Supplementary Materials, available at https://doi.org/10.1016/j.esmoop.2024.103626 highlights the shape of the utility curve and equivalence region of optimal doses under this scenario.

When we compare simulation results to those of the PRO-CRM, under scenarios 2 and 8 (where only one dose is optimal under the utility curve), the U-PRO-CRM improves the probability of correct selection by 4%-10% compared with PRO-CRM.

The proportion of times a dose is recommended depends on how close other doses are to the utility curve. For example, under scenario 5, doses 2 and 3 are approximately equidistant from the utility curve. Therefore, the proportion of times each dose is recommended as the MTD is almost equal. Supplementary Table S2 in the Supplementary Materials, available at https://doi.org/10.1016/j.esmoop.2024.103626 presents the mean proportion of patients allocated to each dose across the nine simulation scenarios. In simulation scenarios where there are multiple admissible doses, U-PRO-CRM assigns between 79% and 95% of patients to a tolerable dose. Under scenarios 2 and 8, where there is only one admissible dose, setting α equal to two improved the mean proportion of patients assigned to the true MTD by 6%-7%, respectively.

For U-PRO-CRM, trials are prematurely stopped 34% of the time under scenario 8 and are correctly stopped 94% of the time under scenario 9.

The benchmark generally performs similarly to U-PRO-CRM when $\alpha = 2$. In scenario 1, the benchmark selects dose 1 more often than dose 2 unlike U-PRO-CRM. This may be because under scenario 1, doses 1 and 2 have very similar true Clinician-DLT and Patient-DLT rates.

Comparison to PRO-CRM design

Simulation results for the U-PRO-CRM with $\alpha = 15$ are presented in Table 3. When $\alpha = 15$ this represents the special case where U-PRO-CRM imitates the PRO-CRM design. Hence, we would expect U-PRO-CRM to imitate PRO-CRM's performance in this instance.

In terms of the proportion of correct MTD recommendations, U-PRO-CRM performs similarly to the PRO-CRM design—with performance of each design within two percentage points of each other among all scenarios. U-PRO-CRM and PRO-CRM perform similarly with respect to the stopping rule. For both designs, trials are prematurely stopped 33% of the time under scenario 8 where dose 1 is the MTD. Under scenario 9 (where no dose is tolerable), both U-PRO-CRM and PRO-CRM correctly stopped trials nearly 94% of the time.

Supplementary Table S3 in the Supplementary Materials, available at https://doi.org/10.1016/j.esmoop.2024.103626 shows the mean proportion of patients allocated to the MTD and overdosed for each design across all scenarios. Once again, U-PRO-CRM performs similarly to PRO-CRM in terms of the proportion of patients allocated to the MTD, but we can note that U-PRO-CRM overdoses a smaller proportion of patients under each scenario.

As expected, the benchmark outperforms both PRO-CRM and U-PRO-CRM designs by 3%-13% in the scenarios which do not execute the stopping rule (scenarios 1-7). This is generally competitive when compared with other trial designs.[22] Under scenario 8 and compared with the benchmark, the performances of the U-PRO-CRM and PRO-CRM are affected by the inclusion of the stopping rule.

DISCUSSION

Unlike trial designs which rely solely on Clinician-DLTs to recommend an MTD, U-PRO-CRM aims to make use of a more informed assessment of treatment tolerability by using PROs before escalating treatment dose.

As well as imitating PRO-CRM in a special case, U-PRO-CRM also provides trialists with additional flexibility to define an MTD with respect to both clinician and patient endpoints. Whilst PRO-CRM constrains dose selection to maximising either Clinician-DLT or Patient-DLT rate, U-PRO-CRM has greater adaptability. U-PRO-CRM gives more control to trialists who want to personally tailor the trade-off between target Clinician-DLT and Patient-DLT rate by using a utility curve.

Under the special case where U-PRO-CRM imitates the PRO-CRM decision rule, U-PRO-CRM performs just as well. Compared with PRO-CRM, U-PRO-CRM has slight advantageous operating characteristics—both in terms of the proportion of correct MTD recommendations, and the mean number of patients allocated to the optimal MTD and overdosed.

The U-PRO-CRM design makes a final MTD selection by choosing the dose with estimated DLT rate closest to some utility curve. In such cases, the U-PRO-CRM design critically relies on the selection of an admissible region (particularly important if the recommended MTD's probability of Patient-DLT and Clinician-DLT lies above the utility curve). This admissible region could be developed to allow for asymmetry—permitting a smaller equivalence region above the curve to prevent overdosing of patients. Alternatively, this design could be extended to include an overdose control, similar to designs such as EWOC.[22] During U-PRO-CRM implementation within a host trial, one must also consider how best to define and elicit the equivalence region describing an indifference between doses. This region, and thus implicitly ϵ, should be decided within the trial protocol.

Whilst this paper has presented an empirical CRM likelihood framework to estimate the Clinician-DLT and Patient-DLT rate, U-PRO-CRM could be extended to a Bayesian framework or other additional working models.

To accurately define the MTD using either the PRO-CRM or U-PRO-CRM design, care must be taken to elicit the target Patient-DLT rate. To accurately define the MTD using either the PRO-CRM or U-PRO-CRM design, it is essential to

Table 3. Proportion of MTD recommendations for each dose level under nine scenarios with $\alpha = 15$ and $\varphi = 0.9$ over 5000 simulations for U-PRO-CRM, PRO-CRM and benchmark[20]

		Dose level				
		1	2	3	4	5
Scenario 1	True probability of Clinician-DLT	0.05	0.05	**0.25**	0.40	0.55
	True probability of Patient-DLT	0.17	0.18	**0.35**	0.50	0.65
	U-PRO-CRM dose recommendation (%)	1	22	**73**	4	0
	PRO-CRM dose recommendation (%)	1	20	**72**	5	0
	Benchmark dose recommendation (%)	9	6	**83**	3	0
Scenario 2	True probability of Clinician-DLT	0.05	**0.25**	0.40	0.55	0.70
	True probability of Patient-DLT	0.10	**0.15**	0.35	0.50	0.65
	U-PRO-CRM dose recommendation (%)	6	**74**	20	0	0
	PRO-CRM dose recommendation (%)	6	**72**	21	0	0
	Benchmark dose recommendation (%)	2	**87**	11	0	0
Scenario 3	True probability of Clinician-DLT	0.01	0.02	0.05	0.10	**0.25**
	True probability of Patient-DLT	0.04	0.09	0.17	0.20	**0.35**
	U-PRO-CRM dose recommendation (%)	0	0	3	33	**64**
	PRO-CRM dose recommendation (%)	0	0	3	30	**66**
	Benchmark dose recommendation (%)	0	0	2	22	**76**
Scenario 4	True probability of Clinician-DLT	0.02	0.05	0.10	**0.25**	0.40
	True probability of Patient-DLT	0.09	0.17	0.20	**0.35**	0.50
	U-PRO-CRM dose recommendation (%)	0	2	32	**61**	4
	PRO-CRM dose recommendation (%)	0	2	29	**63**	5
	Benchmark dose recommendation (%)	0	3	22	**73**	2
Scenario 5	True probability of Clinician-DLT	0.05	0.10	**0.16**	0.25	0.40
	True probability of Patient-DLT	0.05	0.20	**0.35**	0.50	0.65
	U-PRO-CRM dose recommendation (%)	0	15	**71**	15	0
	PRO-CRM dose recommendation (%)	0	14	**69**	17	0
	Benchmark dose recommendation (%)	0	14	**74**	12	0
Scenario 6	True probability of Clinician-DLT	0.05	**0.18**	0.20	0.25	0.40
	True probability of Patient-DLT	0.17	**0.35**	0.50	0.65	0.80
	U-PRO-CRM dose recommendation (%)	9	**74**	16	0	0
	PRO-CRM dose recommendation (%)	9	**72**	18	0	0
	Benchmark dose recommendation (%)	7	**80**	12	0	0
Scenario 7	True probability of Clinician-DLT	0.01	0.05	0.10	**0.16**	0.25
	True probability of Patient-DLT	0.04	0.05	0.20	**0.35**	0.50
	U-PRO-CRM dose recommendation (%)	0	0	16	**69**	15
	PRO-CRM dose recommendation (%)	0	0	15	**67**	18
	Benchmark dose recommendation (%)	0	0	15	**73**	12
Scenario 8	True probability of Clinician-DLT	**0.25**	0.40	0.55	0.70	0.80
	True probability of Patient-DLT	**0.35**	0.50	0.65	0.80	0.85
	U-PRO-CRM dose recommendation (%)	**63**	4	0	0	0
	PRO-CRM dose recommendation (%)	**62**	5	0	0	0
	Benchmark dose recommendation (%)	**98**	2	0	0	0
Scenario 9	True probability of Clinician-DLT	0.45	0.50	0.55	0.70	0.80
	True probability of Patient-DLT	0.55	0.60	0.65	0.80	0.85
	U-PRO-CRM dose recommendation (%)	6	0	0	0	0
	PRO-CRM dose recommendation (%)	6	0	0	0	0
	Benchmark dose recommendation (%)	NA	NA	NA	NA	NA

The MTD under each scenario is presented in bold. PRO-CRM results follow those of Table 2 and are replicated here for your convenience. CRM, Continual Reassessment Method; DLT, dose-limiting toxicity; MTD, maximum tolerated dose; PRO, patient-reported outcome; PRO-CRM, PRO-Continual Reassessment Method; U-PRO-CRM, Utility-PRO-Continual Reassessment Method.

carefully elicit the target Patient-DLT rate. For the U-PRO-CRM, additional care should be taken to ensure the utility curve accurately reflects the trade-off between patient-reported and clinician-reported toxicity. It is noteworthy that the elicited utility curve may vary depending on the patient stakeholders involved. Patient representatives involved in trial design (who do not have lived experience of phase I trials) and phase I trial participants may have differing views of what is considered tolerable. We encourage trialists to develop the utility curve in collaboration with patient partners, including, if possible, those with phase I trial experience, to capture their perspectives. Engaging with patient partners and clinicians, and explaining the concepts clearly to both groups, is crucial for gathering the necessary information to elicit the utility curve. Continued efforts to involve and engage patients in developing such novel intricate designs can assist trialists in refining this trade-off. Continued efforts to involve and engage patients in developing such novel intricate designs can assist trialists in refining this trade-off.

PRO-CRM and U-PRO-CRM would also benefit from an extension which looks to assess efficacy as a ternary outcome within a seamless phase I/II design. Such an advanced design would allow for an optimal biological dose to be identified which looks to maximise efficacy whilst monitoring whether a treatment is tolerable as assessed by the clinicians and the patients. This integration of PRO and efficacy endpoints will help identify optimal doses that

more comprehensively capture the patients' perspective on drug tolerability whilst assessing treatment efficacy during an early-phase dose-finding trial.

Patient-centric dose-finding approaches, incorporating PROs, are likely to play an increasingly pivotal role in advancing our understanding of treatment tolerability, holding substantial implications for the future of early-phase trials.

FUNDING

EA has been supported to undertake this work as part of a PhD studentship from the Institute of Cancer Research within the MRC/NIHR Trials Methodology Research Partnership (no grant number).

DISCLOSURE

The authors have declared no conflicts of interest.

DATA SHARING

The code used to simulate results presented for this paper is publicly available in the following GitHub repository: https://github.com/alemily100/u-pro-crm-simulations.

REFERENCES

1. National Cancer Institute. Common terminology criteria for adverse events (CTCAE). 2023. Available at http://ctep.cancer.gov. Accessed December 8, 2023.
2. Fiteni F, Ray IL, Ousmen A, Isambert N, Anota A, Bonnetain F. Health-related quality of life as an endpoint in oncology phase I trials: a systematic review. *BMC Cancer*. 2019;19(1):361.
3. 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(11):903-909.
4. Henon C, Lissa D, Paoletti X, et al. Patient-reported tolerability of adverse events in phase 1 trials. *ESMO Open*. 2017;2(2):e000148.
5. U.S. Department of Health and Human Services FDA Center for Drug Evaluation and Research. U.S. Department of Health and Human Services FDA Center for Biologics Evaluation and Research; U.S. Department of Health and Human Services FDA Center for Devices and Radiological Health. Guidance for industry: patient-reported outcome measures: use in medical product development to support labeling claims: draft guidance. *Health Qual Life Outcomes*. 2006;4(1):79.
6. Basch E, Jia X, Heller G, et al. Adverse symptom event reporting by patients vs clinicians: relationships with clinical outcomes. *J Natl Cancer Inst*. 2009;101(23):1624-1632.
7. Hou JZ, Ryan K, Du S, et al. Real-world ibrutinib dose reductions, holds and discontinuations in chronic lymphocytic leukemia. *Future Oncol*. 2021;17(35):4959-4969.
8. Murphy R, Halford S, Symeonides SN. Project Optimus, an FDA initiative: considerations for cancer drug development internationally, from an academic perspective. *Front Oncol*. 2023;13:1144056.
9. Lai-Kwon J, Yin Z, Minchom A, Yap C. Trends in patient-reported outcome use in early phase dose-finding oncology trials — an analysis of ClinicalTrials.gov. *Cancer Med*. 2021;10(22):7943-7957.
10. Alger E, Minchom A, Lee Aiyegbusi O, Schipper M, Yap C. Statistical methods and data visualisation of patient-reported outcomes in early phase dose-finding oncology trials: a methodological review. *EClinicalMedicine*. 2023;64:102228.
11. O'Quigley J, Pepe M, Fisher L. Continual reassessment method: a practical design for Phase 1 clinical trials in cancer. *Biometrics*. 1990;46(1):33-48.
12. Lee SM, Lu X, Cheng B. Incorporating patient-reported outcomes in dose-finding clinical trials. *Stat Med*. 2020;39(3):310-325.
13. Andrillon A, Biard L, Lee SM. Incorporating patient-reported outcomes in dose-finding clinical trials with continuous patient enrollment. *J Biopharm Stat*; 2023. p. 1-12. https://doi.org/10.1080/10543406.2023.2236216.
14. Le Tourneau C, Razak AR, Gan HK, et al. Heterogeneity in the definition of dose-limiting toxicity in phase I cancer clinical trials of molecularly targeted agents: a review of the literature. *Eur J Cancer*. 2011;47(10):1468-1475.
15. Basch E, Reeve BB, Mitchell SA, et al. Development of the National Cancer Institute's patient-reported outcomes version of the common terminology criteria for adverse events (PRO-CTCAE). *J Natl Cancer Inst*. 2014;106(9):dju244.
16. Wages NA, Nelson B, Kharofa J, Meier T. Application of the patient-reported outcomes continual reassessment method to a phase I study of radiotherapy in endometrial cancer. *Int J Biostat*. 2022;19(1):163-176.
17. Moe A, Liow E, Redfern A, et al. A phase I open label dose-escalation study to evaluate the tolerability, safety and immunological efficacy of sub-urothelial durvalumab injection in adults with muscle-invasive or high-risk non-muscle-invasive bladder cancer (SUBDUE-1, SUBurothelial DUrvalumab injection-1 study): clinical trial protocol. *BJU Int*. 2021;128(suppl 1):9-17.
18. Thall PF, Cook JD. Dose-finding based on efficacy—toxicity trade-offs. *Biometrics*. 2004;60(3):684-693.
19. Leonard JP, Furman RR, Cheung YKK, et al. Phase I/II trial of bortezomib + CHOP-rituximab in diffuse large B cell (DLBCL) and mantle cell lymphoma (MCL): phase I results. *Blood*. 2005;106(11):491.
20. Cheung YK. Simple benchmark for complex dose finding studies. *Biometrics*. 2014;70(2):389-397.
21. Yuan Y, Yin G. Bayesian dose finding by jointly modelling toxicity and efficacy as time-to-event outcomes. *J R Stat Soc C*. 2009;58(5):719-736.
22. Babb J, Rogatko A, Zacks S. Cancer phase I clinical trials: efficient dose escalation with overdose control. *Stat Med*. 1998;17(10):1103-1120.