Received: 13 May 2018 | Revised: 19 January 2019 | Accepted: 7 May 2019

DOI: 10.1002/pst.1955

MAIN PAPER

WILEY

Improvements to the Escalation with Overdose Control design and a comparison with the restricted Continual Reassessment Method

Lingyun Ji[1] | Juan Pablo Lewinger[1] | Mark Krailo[1] | Susan Groshen[1] | David V. Conti[1] | Shahab Asgharzadeh[2] | Richard Sposto[1]

[1]Department of Preventive Medicine, Keck School of Medicine, The University of Southern California, Los Angeles, CA, USA

[2]Department of Pediatrics and Pathology, Keck School of Medicine, The University of Southern California, Los Angeles, CA, USA

Correspondence
Lingyun Ji, PhD, Assistant Professor of Clinical Preventive Medicine, Department of Preventive Medicine, Keck School of Medicine, University of Southern California, Children's Oncology Group, 222 E. Huntington Dr. Suite 100, Monrovia, CA 91016, USA.
Email: lji@usc.edu

The Escalation with Overdose Control (EWOC) design for cancer dose finding clinical trials is a variation of the Continual Reassessment Method (CRM) that was proposed to overcome the limitation of the original CRM of exposing patients to high toxic doses. The properties of EWOC have been studied to some extent, but some aspects of the design are not well studied, and its performance is not fully understood. Comparisons of the EWOC design to the most commonly used modified CRM designs have not yet been performed, and the advantages of EWOC over the modified CRM designs are unclear.

In this paper, we assess the properties of the EWOC design and of the restricted CRM and some variations of these designs. We show that EWOC has several weaknesses that CRM does not have that make it impractical to use in its original formulation. We propose modified EWOC designs that address some of the weaknesses and that have some desirable statistical properties compared with the original EWOC design, the restricted CRM design, and the 3 + 3 design. However, their statistical properties are sensitive to correct specification of the prior distribution of their parameters and hence nevertheless will need to be used with some caution. The restricted CRM design is shown to have more stable performance across a wider family of dose-toxicity curves than EWOC and therefore may be a preferable general choice in cancer clinical research.

KEYWORDS
Bayesian, CRM, dose finding, EWOC, Phase I

1 | INTRODUCTION

Most cancer Phase I clinical trials still use rule-based designs (eg, 3 + 3) despite much methodological research on model-based designs,[1,2] likely because model-based designs are less familiar and more complicated.[3] Also, the differences and similarities of various model-based designs are not well-understood, which may hinder researchers from using them. In this paper, we assess the properties of the Escalation with Overdose Control (EWOC) design[4] and compare it to a modified Continual Reassessment Method (CRM),[5] to provide a clear guidance in their usages.

Abbreviations: CRM, The Continual Reassessment Method; EWOC, The Escalation with Overdose Control design

The CRM[5] was the first model-based, Bayesian method in Phase I trial design. With CRM, one starts with a parametric model that describes the dose-toxicity relationship and a prior distribution for the parameters. As data on dose-limiting toxicity (DLT) accumulate, the posterior distribution of the parameters and investigators' knowledge of the dose-toxicity relationship get updated. "The rationale of the CRM is to concentrate as many patients as we can on doses at, or close to, the MTD (maximum tolerated dose)."[6] Despite its strengths, the original CRM raised many concerns and was considered to be impractical[7-12] and led to subsequent recommended modifications.[6-12]

One of the major criticisms of the original CRM was that it can lead to exposure of patients to toxic doses and is aggressive in assigning high dose levels when little knowledge is available regarding the dose-toxicity relationship. To overcome this limitation, Babb et al proposed the EWOC design.[4] The goal of EWOC is to identify the MTD at the end of the study while limiting the number of patients exposed to toxic doses during the study. It has additional safety measures put in place to avoid exposing patients to doses that are potentially too toxic.[1,13] In contrast to CRM recommending a dose estimated to be closest to the MTD, EWOC recommends a dose for which the probability of it being too toxic is minimized.

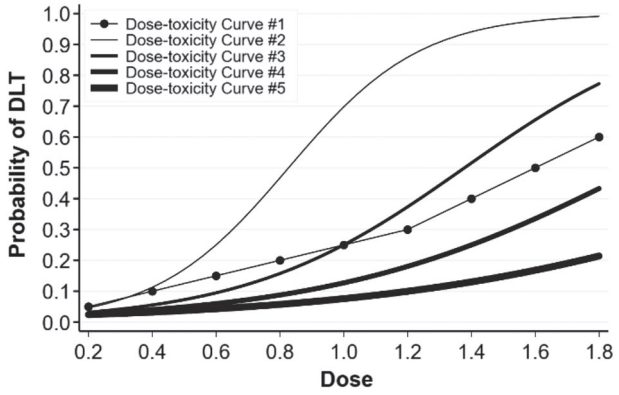
The properties of EWOC have been studied to some extent, but some aspects of it are not well studied and its performance is not fully understood. In addition, although EWOC has been compared with the original CRM, comparison of it to the most commonly used modified CRM designs has not yet been performed. The advantages of EWOC over the modified CRM designs are unclear. Hence, there is a lack of guidance in the use of CRM, EWOC, and their variations.

This paper presents some weaknesses with the EWOC design that have not been reported previously and demonstrates that EWOC in its original formulation is impractical to be used in clinical trials. We examine the reasons that cause these weaknesses and propose modified EWOC designs to remedy the EWOC formulation. Using a simulation study, we assess the properties of the original EWOC and the proposed modified EWOC designs, and compare them to the restricted CRM[10] and the 3 + 3[14,15] design. We then make recommendations regarding their usages in cancer clinical research.

2 | MATERIAL AND METHODS

A simulation study was conducted to examine and compare the properties of the EWOC design and five other design methods that are described below.

Five dose-toxicity curves are used in the simulation study, as presented in Figure 1. The simulation pre-specifies nine dose levels ranging from 0.2 to 1.8, with an increment of 0.2 between any two dose levels. The target toxicity rate is set at 0.25. For Curve #1, four scenarios are considered, with starting dose being (a) dose 0.2, (b) dose 0.4, (c) dose 0.6, and (d) dose 0.8, respectively. For Curves #2, #3, #4, and #5, three scenarios are considered, with starting dose being (a) dose



Curve	Dose								
	0.2	**0.4**	**0.6**	**0.8**	**1**	**1.2**	**1.4**	**1.6**	**1.8**
#1	0.05	0.10	0.15	0.20	**0.25**	0.30	0.40	0.50	0.60
#2	0.046	0.112	**0.25**	0.467	0.697	0.859	0.941	0.977	0.991
#3	0.032	0.055	0.095	0.157	**0.25**	0.373	0.516	0.655	0.773
#4	0.027	0.040	0.060	0.088	0.127	0.180	**0.25**	0.335	0.433
#5	0.024	0.032	0.043	0.057	0.076	0.100	0.130	0.168	0.214

FIGURE 1 Dose-toxicity relationship assumed in the simulation study

0.2, (b) dose 0.6, and (c) dose 1.0, respectively. A total of 1000 trials were simulated for each scenario with each of the design methods.

2.1 | Design method 1—The originally proposed EWOC design with α fixed at 0.25 (noted as EWOC)

This is the initially proposed EWOC design.[4] Let $\{x_i : i = 1, \ldots, k, x_1 < x_2 < \ldots < x_k\}$ be a sample space of pre-specified doses. As in the original paper, a two-parameter logit function $F(\beta_0+\beta_1 x_j)$ is used to model the dose-toxicity relationship. EWOC re-parameterizes the dose-toxicity model and introduces the random variable γ that denotes the dose giving the target toxicity rate θ, and the parameter ρ_0 that denotes the probability of a DLT at the minimum dose X_{min}. It is assumed that ρ_0 and γ take priors $Uniform(0,\theta)$ and $Uniform(x_1,x_k)$, respectively. So $\gamma = \frac{F^{-1}(\theta)-\beta_0}{\beta_1}$, and $\rho_0 = F(\beta_0+\beta_1 X_{min})$. For the j^{th} $(j = 1, \ldots, n-1)$ patient, let $\pi(\gamma|\mathcal{D}_j) = \int f(\rho_0,\gamma|\mathcal{D}_j)d\rho_0$ with $f(\rho_0,\gamma|\mathcal{D}_j)$ being the joint distribution of ρ_0 and γ given data $\mathcal{D}_j$ (ie, data on the first j patients), and define the cumulative distribution function $\Pi_j(z|\mathcal{D}_j) = \int_{X_{\min}}^{z} \pi(\gamma|\mathcal{D}_j)d\gamma$. Except for the first patient, after j patients are treated, the recommended dose for the $(j+1)^{th}$ patient is the dose whose probability of exceeding the true MTD is closest to but less than the feasibility bound parameter α, ie, $\Pi_j^{-1}(\alpha)$. As in the original EWOC paper, α is set at 0.25 in this simulation study. At the completion of a trial, the recommended dose or estimated MTD is defined in two ways: the dose that is the closest to but not above $\Pi_n^{-1}(\alpha)$, or the dose that is the closest to but not above the posterior median of γ, $\Pi_n^{-1}(0.5)$.

2.2 | Design method 2—The EWOC with varying α design (noted as EWOC-varying α)

The EWOC with varying α design is a variation of the original EWOC design.[13,16] Instead of fixing α at a certain value, EWOC-varying α design increases α as the trial keeps enrolling patients. Generally, there is a lack of guidance in how α can be varied in a trial. Here, α is set to start at 0.25 and increase by 0.05 for every two patients enrolled, until it reaches 0.5. Other than the difference in α, the setup of EWOC-varying α is the same as the EWOC design. At the completion of a trial, the recommended dose of the trial is $\Pi_n^{-1}(0.5)$.

2.3 | Design method 3—The 3 + 3 design (noted as 3 + 3)

The 3 + 3 design used in this simulation study follows the most common 3 + 3 algorithm.[14,15]

2.4 | Design method 4—The restricted CRM design (noted as R-CRM)

In the simulation study, we did not consider the original CRM[5] because it is seldom used in clinical trials in its original form. Instead, we considered the restricted CRM design (R-CRM) proposed by Moller,[10] which prevents dose escalation from ever being greater than a single step. As for EWOC, the two-parameter logit function is used for R-CRM. We assumed that β_0 takes the prior $Uniform(-10,0)$ and β_1 takes the prior $Uniform(0,10)$.[13]

In addition to the above four methods, we propose two modified EWOC designs intended to remedy weaknesses of the original EWOC formulation that will be presented in the results section. We would like to assess whether these modified designs possess any advantages over other dose-finding methods.

2.5 | Design method 5—A modified EWOC design (noted as EWOC-2-stage)

In this proposed modified EWOC design, enrollments are separated into two stages. The first stage fixes α at 0.25. After a good amount of information has been gathered on the dose-toxicity curve, α will be changed to 0.5 in the second stage of the trial so that patients can be treated at dose levels that will more likely be the recommended dose. With this design, the first patient is treated at the chosen starting dose level, and the second patient is treated per the original EWOC formulation. If the $(j-1)^{th}$ $(1 < j < n)$ patient and the j^{th} patient have received the same dose level x_i with no DLT, and if

the EWOC formulation recommends dose x_i again to the $(j+1)^{th}$ patient, apply the following "look-ahead rule" before determining the recommended dose for the $(j+1)^{th}$ patient:

> *"Look ahead for the next patient, ie, the $(j+2)^{th}$ patient. If the $(j+1)^{th}$ patient does not have a DLT and the $(j+2)^{th}$ patient is still recommended the same dose level x_i by the EWOC formulation, then recommend to the $(j+1)^{th}$ patient the dose level one step higher than x_i, ie, x_{i+1}."*

To ensure coherence[8,17] of the design, we apply the additional restriction that no dose de-escalation will be allowed if the previous patient did not experience a DLT. In the simulation study, with total sample size of 25, the first stage is defined as the first 12 patients. At the completion of the trial, the recommended dose is $\Pi_n^{-1}(0.5)$.

2.6 | Design method 6—The restricted EWOC-mean design (noted as R-EWOC-mean)

This proposed modified design is similar to the original EWOC in that it uses the same parameters γ and ρ_0 but is similar to CRM in that it uses posterior mean of γ as the recommended dose. As in R-CRM, a restriction is placed such that at any time dose escalation cannot be greater than a single step.

3 | RESULTS

We first present three weaknesses of the original EWOC design that were revealed by our simulation study that have not been reported previously. The first weakness is that when the first or second patient enrolled on a trial has a DLT, the design can get stuck and recommend the same low and possibly ineffective dose level to all future patients regardless of their DLT status. Hence, the trial will fail. In our simulation study, this issue occurred in a non-ignorable proportion of trials. With dose-toxicity Curve #1, with the starting dose of 0.2, 0.4, 0.6, and 0.8, the proportion of simulated trials that failed due to this reason was 14.4%, 8.9%, 14.2%, and 19.6%, respectively.

Four example failed trials are given in Table 1. For example, for trial 1A, the first patient was treated at dose 0.2 with no DLT, the second patient was recommended dose 0.6 and experienced a DLT, and EWOC recommended dose 0.2 for all remaining patients regardless of their DLT status. This design characteristic would not be acceptable to clinical investigators. Examination of the change in posterior dose-toxicity function after each enrolled patient illustrates the cause of this problem. Figure 2 presents the dose-toxicity function before enrolling the first patient and after enrolling the first through fifth patients using an example trial that got stuck. The first patient was treated at dose 0.4 and experienced a DLT, and the posterior dose-toxicity function increased steeply between dose 0.2 and dose 0.4 and remained essentially the same after each successive patient, causing EWOC to repeatedly recommend dose 0.2 to all patients.

The second weakness of the EWOC design is that patients can be treated very inefficiently even if the trial does not get stuck at a single low dose level, in that long series of patients can be assigned to a low dose level before a higher dose level is recommended. This can be demonstrated by two example trials shown in Table 2.

The third weakness with the original EWOC design, when α is fixed at 0.25, is that depending on how one selects the MTD estimate at completion of the trial it can be either biased or at a dose where no patient has been treated. Although EWOC has the flexibility in choosing from various MTD estimators (eg, $\Pi_n^{-1}(\alpha)$ which is the dose that will be recommended to the $(n+1)^{th}$ patient per EWOC, mean of the posterior distribution of γ, $\Pi_n^{-1}(0.5)$ which is the median of the posterior distribution of γ, among others[4]), these estimators can all have unfavorable properties as illustrated by Figure 3. Figure 3 gives the distribution of the estimated MTD among the simulated trials for dose-toxicity Curve #1 using 0.2 as starting dose, with MTD estimated either as $\Pi_n^{-1}(\alpha)$, $\alpha = 0.25$ (left panel), or as $\Pi_n^{-1}(0.5)$ (right panel). As shown in Figure 3, choosing $\Pi_n^{-1}(\alpha)$ as the MTD estimator was very conservative and biased, recommending doses much lower than the true MTD. Choosing $\Pi_n^{-1}(0.5)$ as the estimator leads to much smaller bias, but here it could be the case that no patients in the trial have yet been treated at the selected recommended dose. This is demonstrated by an example trial in Table 3. For this trial, the recommended dose was dose 1.0, but no patients in the trial were treated at dose 1.0. In our simulation study of EWOC with this median estimator, more than 50% of the simulated trials recommended a dose that was never given to any enrolled patients. Clinical investigators will not likely agree to a recommended dose at which no patients have yet been treated, as extrapolation of dose-toxicity relationship beyond the

TABLE 1 Four example EWOC trials (α fixed at 0.25) that failed after observation of a DLT in the first or second patient enrolled on the trial

Patients	α	Dose Calculated	Dose Recommended	P (DLT) at Dose Given	DLT
1	0.25	--	0.2	0.05	0
2	0.25	0.60	0.6	0.15	1
3	0.25	0.32	0.2	0.05	0
4	0.25	0.32	0.2	0.05	0
5-24	0.25	0.30-0.32	0.2	0.05	0 or 1
25	0.25	0.32	0.2	0.05	0
Patients	α	**Dose Calculated**	**Dose Recommended**	**P (DLT) at Dose Given**	**DLT**
1	0.25	--	0.4	0.1	1
2	0.25	0.32	0.2	0.05	0
3	0.25	0.32	0.2	0.05	0
4	0.25	0.30	0.2	0.05	1
5-24	0.25	0.28-0.36	0.2	0.05	0
25	0.25	0.28	0.2	0.05	0
Patients	α	**Dose Calculated**	**Dose Recommended**	**P (DLT) at Dose Given**	**DLT**
1	0.25	--	0.6	0.15	1
2	0.25	0.34	0.2	0.05	0
3	0.25	0.32	0.2	0.05	0
4	0.25	0.32	0.2	0.05	0
5-24	0.25	0.30-0.32	0.2	0.05	0
25	0.25	0.30	0.2	0.05	0
Patients	α	**Dose Calculated**	**Dose Recommended**	**P (DLT) at Dose Given**	**DLT**
1	0.25	--	0.8	0.2	1
2	0.25	0.34	0.2	0.05	0
3	0.25	0.34	0.2	0.05	0
4	0.25	0.34	0.2	0.05	0
5-24	0.25	0.32-0.34	0.2	0.05	0 or 1
25	0.25	0.32	0.2	0.05	0

Trial 1A: The first patient was treated at dose 0.2 and did not have a DLT. The second patient was treated at dose 0.6 and had a DLT. The dose recommended for all the remaining patients was dose 0.2.

Trial 1B: The first patient was treated at dose 0.4 and had a DLT. The dose level recommended for all the remaining patients was dose 0.2.

Trial 1C: The first patient was treated at dose 0.6 and had a DLT. The dose level recommended for all the remaining patients was dose 0.2.

Trial 1D: The first patient was treated at dose 0.8 and had a DLT. The dose level recommended for all the remaining patients was dose 0.2.

data is always risky. Although we did not assess all possible MTD estimators, eg, mean of the posterior distribution of γ, we think that the above issue will exist with the other estimators. Hence, if a trial adopts the original EWOC design, recommending $\Pi_n^{-1}(\alpha)$ with $\alpha = 0.25$ at the end of the trial may be the only feasible choice.

In an intention to remedy the above three weaknesses of EWOC, we proposed two modified EWOC designs, EWOC-2-stage, and R-EWOC-mean (ie, design methods 5 and 6). In the following, we present a comparison of the design characteristics among design methods 1 to 6. Using the simulated trials that assumed the dose-toxicity Curve #1 with the starting dose of 0.2, Figure 4 shows a comparison of the distribution of the recommended dose among the six designs. Varying starting dose for Curve #1 gave similar performance, and results are hence not presented. Similar patterns are observed on Curves #2, #3, and #4, and results are presented in Supplemental Figures S1, S2, and S3.

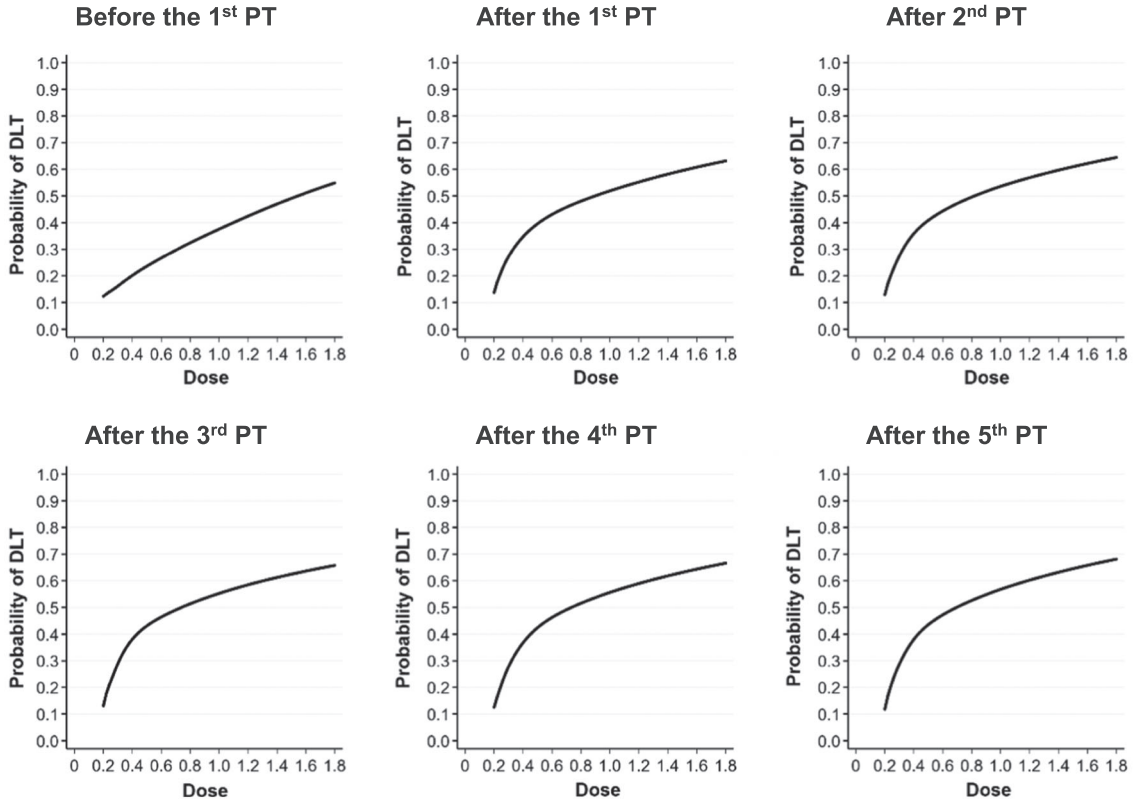


FIGURE 2 Posterior dose-toxicity functions for a trial in which the first patient was treated at dose 0.4, and all remaining patients were recommended dose 0.2 when none of them had a DLT

TABLE 2 Two example EWOC trials showing that patients can be treated inefficiently

Patients	α	Dose Calculated	Dose Recommended	P (DLT) at Dose Given	DLT
1	0.25	--	0.6	0.15	0
2	0.25	0.72	0.6	0.15	1
3	0.25	0.54	0.4	0.1	1
4-20	0.25	0.46-0.58	0.4	0.1	2 1 s, 15 0 s
21	0.25	0.60	0.6	0.15	0
22	0.25	0.64	0.6	0.15	0
23	0.25	0.66	0.6	0.15	0
24	0.25	0.68	0.6	0.15	1
25	0.25	0.66	0.6	0.15	1
Patients	α	**Dose Calculated**	**Dose Recommended**	**P (DLT) at Dose Given**	**DLT**
1	0.25	--	0.4	0.05	0
2	0.25	0.60	0.6	0.15	0
3	0.25	0.76	0.6	0.15	0
4	0.25	0.82	0.8	0.2	1
5	0.25	0.60	0.6	0.15	1
6-25	0.25	0.46-0.52	0.4	0.1	3 1 s, 17 0 s

Trial 2A: Patients #3 to #20 are all recommended dose 0.4.

Trial 2B: Patients #6 to #25 are all recommended dose 0.4.

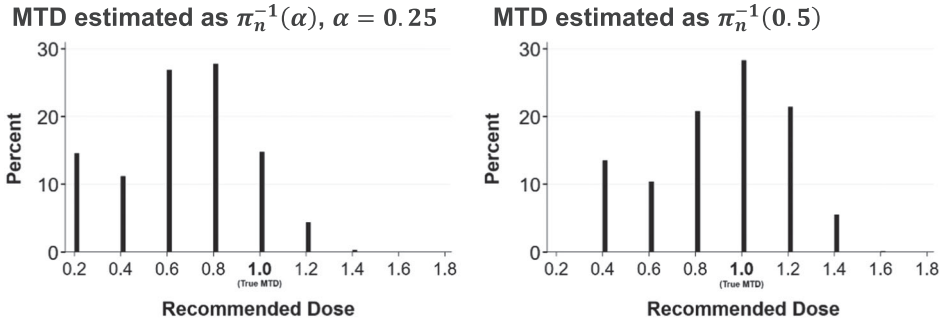


FIGURE 3 The distribution of the recommended dose using Curve #1 with starting dose of 0.2

TABLE 3 One example trial showing that no patients on the trial have been treated at the estimated MTD of $\pi_n^{-1}(0.5)$ =1.0

Patients	α	Dose Calculated	Dose Recommended	P (DLT) at Dose Given	DLT
1	0.25	--	0.2	0.05	0
2	0.25	0.60	0.6	0.15	0
3-20	0.25	0.6-0.8	0.6-0.8	0.15-0.2	0
21	0.25	0.78	0.6	0.15	0
22	0.25	0.80	0.8	0.2	0
23	0.25	0.82	0.8	0.2	0
24	0.25	0.88	0.8	0.2	0
25	0.25	0.90	0.8	0.2	0

As shown in Figure 4, with the original EWOC with α fixed at 0.25, a non-ignorable proportion of trials recommended the lowest dose level at the completion of the trial due to the getting stuck issue. Also, both the 3 + 3 design and the original EWOC on average underestimated the true MTD and showed large variability in the recommended dose. EWOC-varying α and R-CRM gave similar distributions of the final recommended dose and displayed the good property of estimating the true MTD correctly in a large proportion of trials with small or ignorable bias. EWOC-2-stage and R-EWOC-mean showed even better characteristics than EWOC-varying α and R-CRM. They not only showed ignorable bias in the distribution of the estimated MTD but also had better precision in the estimated MTD than the other designs. With these two modified designs, the issue for the EWOC algorithm getting stuck was alleviated. Specifically, among all designs, R-EWOC-mean demonstrated the highest likelihood of selecting true MTD and the smallest variability in the estimated MTD.

Figure 5 shows a comparison of the proportion of patients treated at each dose level among the six designs using dose-toxicity Curve #1. Results for Curves #2, #3, and #4 are presented in Supplemental Figures S4, S5, and S6. The original EWOC design on average treated <5% of patients at dose levels above the true MTD and demonstrated a good property of overdose control. Despite this good feature, many patients were treated at dose levels below the true MTD, with 10% to 20% treated at the lowest dose level. Similar to EWOC, the 3 + 3 design also treated few patients at doses above the true MTD and a large proportion of patients at the low dose levels. EWOC-varying α and R-CRM showed similar patterns, and the advantage of EWOC-varying α in overdose control was not impressively better compared with R-CRM. With EWOC-2-stage and R-EWOC-mean, the issue for an EWOC trial getting stuck at a low dose level was alleviated, and the proportion of patients treated at the lowest dose level was very small and comparable to R-CRM. Compared with the original EWOC, although EWOC-2-stage treated slightly more patients at doses higher than the true MTD (~15% vs <5%), this design maintains to a large extent the good property of overdose control. R-CRM and R-EWOC-mean expose approximately 30% and 35% patients at dose levels above the true MTD, respectively. R-EWOC-mean seemed to be able to avoid extremely toxic dose levels to patients. Of all designs, R-EWOC-mean treated patients the most efficiently, giving the most patients the true MTD and only a very small proportion of patients the very low dose levels.

In short, for drugs with a dose-toxicity relationship similar to Curve #1, the two proposed modified designs R-EWOC-mean and EWOC-2-stage have the best properties among the six designs. R-EWOC-mean appears to be most

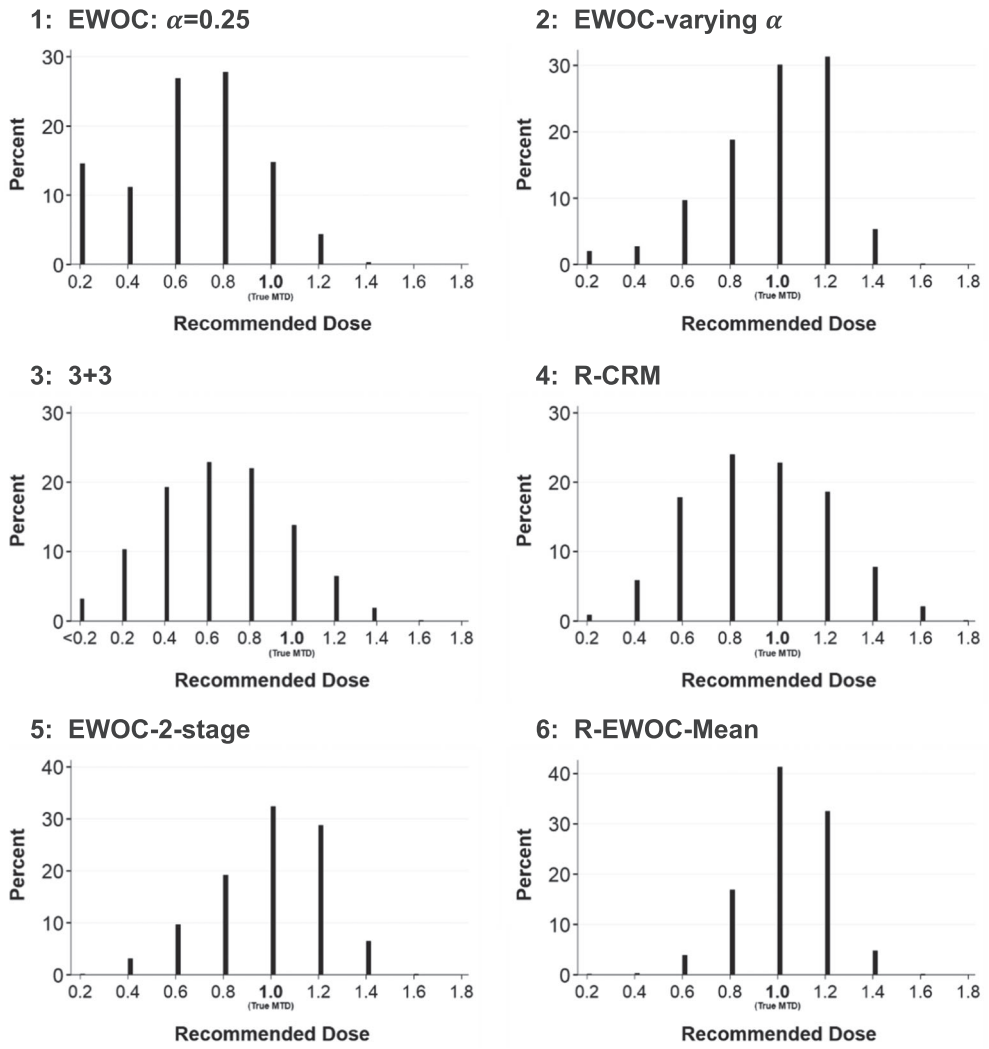


FIGURE 4 Comparison of the distribution of the recommended dose using Curve #1 with starting dose of 0.2

preferable among all the six designs, having the best precision in the distribution of the recommended dose and treating patients most efficiently. EWOC-2-stage exposes fewer patients to dose levels above the true MTD than R-EWOC-mean and can be more desirable than the other designs when an investigator would like to limit exposing patients to toxic dose levels.

Unfortunately, the above good properties of the two modified EWOC designs do not generalize to Curve #5, which revealed a new issue with all the EWOC designs. Different from Curves #1, #2, #3, and #4 where the range of the pre-specified dose levels covers the true MTD, for Curve #5 the drug is generally not very toxic and the pre-specified dose levels are all below the true MTD. The distribution of the recommended dose for Curve #5 with starting dose of 1.0 is presented in Figure 6. (The proportion of patients treated at each dose level is presented in Supplemental Figure S7). All four EWOC variations had the problem of not being able to recommend the highest pre-specified dose level even though that dose did not surpass the true MTD. The reason for this feature of all EWOC designs is that the prior for γ was chosen as $Uniform(x_1, x_k)$ as recommended in the original paper, which limits γ to the range of the pre-specified dose levels between X_{min} (or x_1) and X_{max} (or x_k). In fact, even when the true MTD is smaller but approaches X_{max}, as is the case for Curve #4, this issue manifests itself with EWOC designs tending not to recommend doses in the higher range (Supplemental Figure S3).

Had the prior for γ been selected differently, the properties of the EWOC designs could change significantly. Figure 7 shows the distribution of the recommended dose of EWOC-2-stage and R-EWOC-mean when the prior for γ was chosen as truncated normal distribution: $N(1,1),\gamma \geq 0.2$. As shown by Figure 7.1 and Figure 7.2, for Curve #5, with the prior for γ changed to truncated normal distribution, the two modified EWOC designs were able to recommend the highest dose at the completion of the trial. However, this change of prior led to a dispersed distribution of the recommended dose for

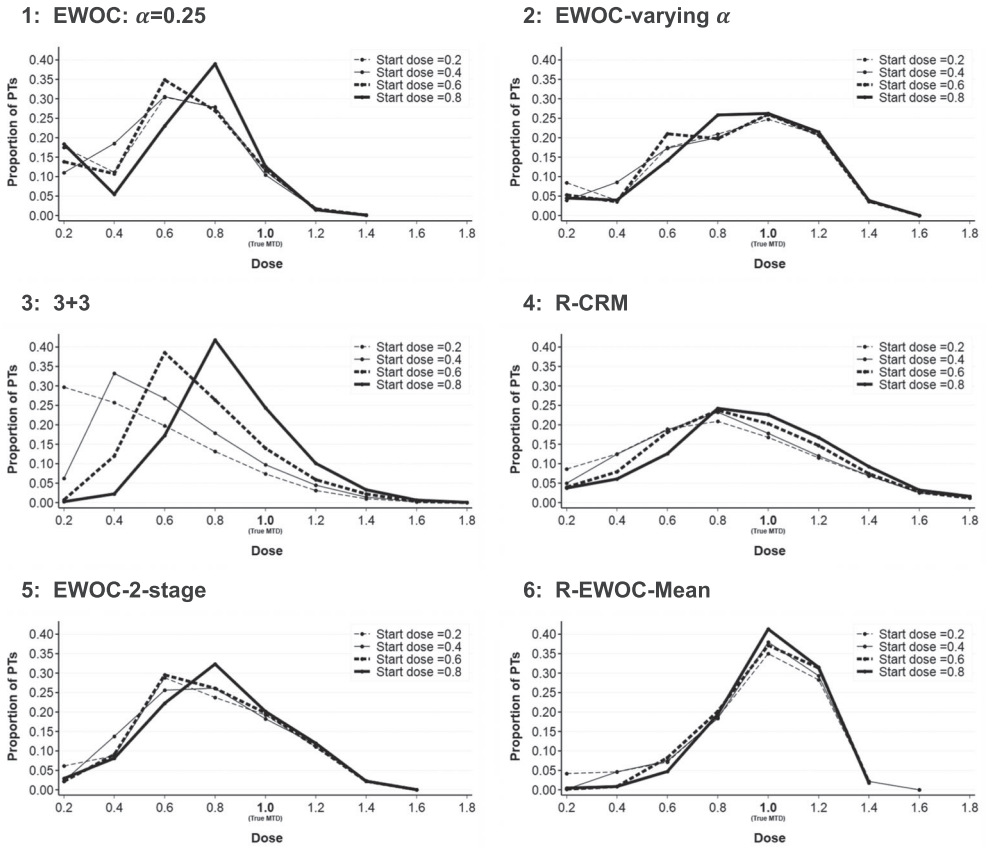


FIGURE 5 Comparison of the proportion of patients treated at each dose using Curve #1

Curve #1 for the two modified EWOC designs, demonstrated by the comparison between Figures 7.3 and 7.4 vs Figures 4.5 and 4.6, which diminished their advantage over R-CRM in precision.

4 | DISCUSSION

We have illustrated three weaknesses with the original EWOC design as presented in the results section. Given these weaknesses, the original EWOC design should not be used without modifications when we design Phase I clinical trials.

A comparison between EWOC vs the other designs demonstrates that the original EWOC has some good properties that the other designs do not have, particularly the property of overdose control and directly estimating the posterior distribution of γ. Considering the two good features of EWOC, it would be ideal if we can have a modified EWOC design that is free from the abovementioned weaknesses. We should note that in our paper we only considered designs where dose-limiting toxicities are defined in binary form. When the raw toxicities are given in continuous or ordinal instead of binary form, a proposed novel method[18] that extends the CRM and accounts for multiple toxicity constraints (eg, toxicity constraints to higher-grade toxicities) can be used to achieve overdose control.

For the existing modified EWOC design, EWOC-varying α, there is a lack of guidance in how α should be changed during trial conduct. The behavior of EWOC-varying α may be very different depending on how α varies. If α is increased fast, EWOC-varying α will treat more patients at doses above the true MTD. If α is increased slowly, EWOC-varying α will treat more patients at doses below the true MTD, and the design can have the issue of getting stuck at a low dose level and the issue of using and treating patients inefficiently, at least in the early phase of the trial. If using this design, how to vary α during the trial should be assessed and its influence on performance should be carefully examined when designing a trial.

Given the above, this paper proposed two modified EWOC designs: EWOC-2-stage and R-EWOC-mean. We demonstrated that EWOC-2-stage and R-EWOC-mean have alleviated the three issues with the original EWOC and can work well with trials for which the range of the pre-specified dose levels covers the true MTD, and that EWOC-2-stage keeps to a good extent the overdose control feature of EWOC.

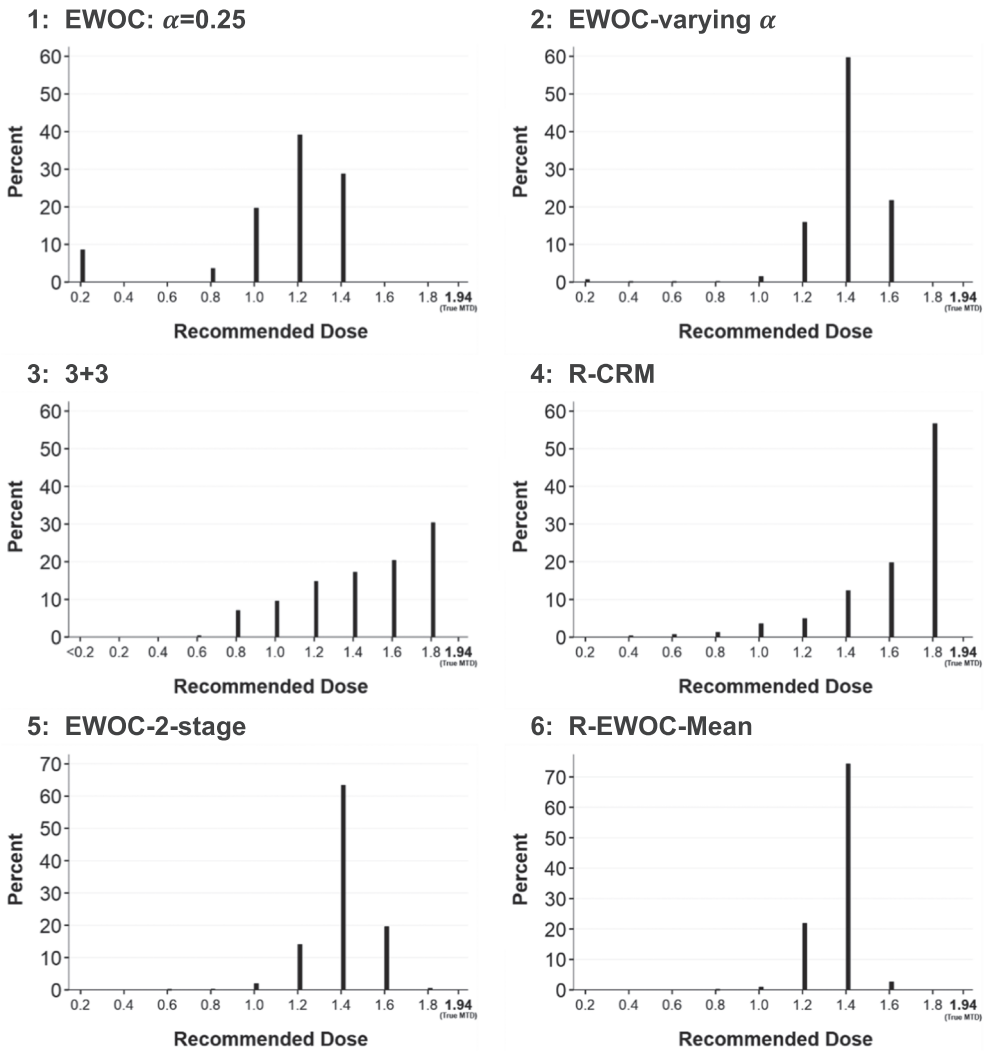


FIGURE 6 Comparison of the distribution of the recommended dose using Curve #5 with starting dose of 1.0

However, we also showed that if the prior $Uniform(X_{min}, X_{max})$ is used for γ, all EWOC designs can suffer from critical problems with drugs that are generally not toxic and for which MTD is not reached with the pre-specified dose levels. These EWOC designs will fail to recommend the highest pre-specified dose. This problem also manifests itself when the true MTD is smaller but approaches X_{max}. With carefully chosen prior for γ, such as the truncated normal distribution, log normal distribution, uniform distribution with an upper bound considerably larger than X_{max}, etc., this problem can be alleviated, but the change of the prior can also decrease the performance of the EWOC designs, making them behave worse than R-CRM. Hence, the EWOC designs need to be used with some caution. If we have an understanding of the range of doses that will cover the true MTD—eg, when an established MTD in adults can inform the design of a pediatric study—then using EWOC-2-stage or R-EWOC-mean with uniform prior results in efficient designs with good statistical properties. However, these types of scenarios may be rare, and very often it is the case that we do not have a good understanding of the doses that will cover the true MTD.

When designing Bayesian clinical trials, a misunderstanding is that uniform priors are minimally informative, where in fact they can contain a large amount of information. For EWOC, uniform priors limit the range of doses that can be selected as MTD. O'Quigley and Conaway commented that "there is no fundamental difference between these approaches [EWOC] and the CRM, aside from the making use of a particular distance measure."[6] We show that the EWOC formulation can be more fragile to the specification of priors than the CRM. The choice of a prior can have a profound influence on the statistical properties of EWOC designs, leading to either very good or very poor designs depending on a drug's underlying dose-toxicity relationship. R-CRM, by contrast, shows more stable performance. Hence, R-CRM is a better general choice than the EWOC designs.

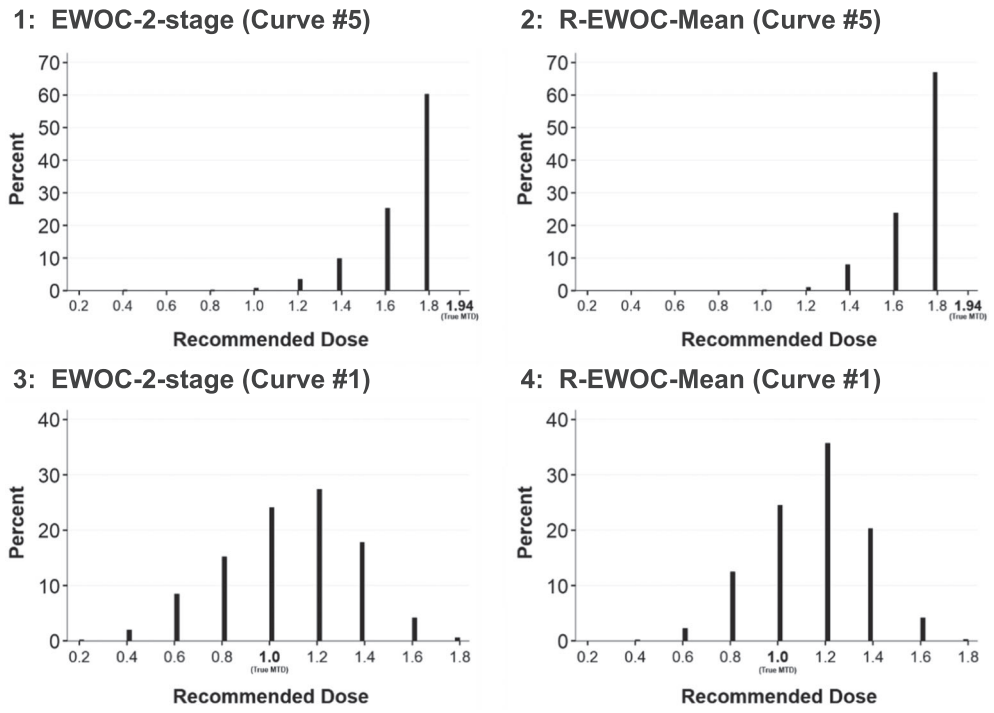


FIGURE 7 Distribution of the recommended dose, with prior of γ changed to truncated normal distribution: $N(1,1),\gamma \geq 0.2$

5 | CONCLUSIONS

The original EWOC design has weaknesses and should not be used without modifications when we design Phase I clinical trials. The two modified EWOC designs EWOC-2-stage and R-EWOC-mean fixed some issues with the original EWOC formulation and can have good statistical properties when the range of the pre-specified dose levels covers the true MTD. However, they are not uniformly better than R-CRM, and all EWOC designs can be fragile to specification of priors. If we do not have any knowledge in the dose-toxicity relationship, which is very often the case when we conduct dose-finding trials, R-CRM may be preferred than EWOC designs given that its performance is more stable across a wider family of dose-toxicity curves.

ACKNOWLEDGEMENTS

We thank the reviewers and editors for their detailed review of the manuscript and helpful comments.

CONFLICT OF INTEREST

None.

DATA AVAILABILITY STATEMENT

Data sharing is not applicable to this article as no new data were created or analyzed in this study.

ORCID

Lingyun Ji https://orcid.org/0000-0003-2500-1993
Mark Krailo https://orcid.org/0000-0002-7608-0698
Susan Groshen https://orcid.org/0000-0001-9705-1012
David V. Conti https://orcid.org/0000-0001-8930-2483
Shahab Asgharzadeh https://orcid.org/0000-0002-0510-3848
Richard Sposto https://orcid.org/0000-0002-2729-2360

REFERENCES

1. Le Tourneau C, Lee JJ, Siu LL. Dose escalation methods in phase I cancer clinical trials. *J Natl Cancer Inst*. 2009;101:708-720.
2. Rogatko A, Schoeneck D, Jonas W, Tighiouart M, Khuri FR, Porter A. Translation of innovative designs into phase I trials. *J Clin Oncol*. 2007;25:4982-4986.
3. Garrett-Mayer E. The continual reassessment method for dose-finding studies: a tutorial. *Clin Trials*. 2006;3(1):57-71.
4. Babb J, Rogatko A, Zacks S. Cancer phase I clinical trials: efficient dose escalation with overdose control. *Stat Med*. 1998 May 30;17(10):1103-1120.
5. O'Quigley J, Pepe M, Fisher L. Continual reassessment method: a practical design for phase 1 clinical trials in cancer. *Biometrics*. 1990 Mar;46(1):33-48.
6. O'Quigley J, Conaway M. Continual reassessment and related dose-finding designs. *Stat Sci*. 2010;25(2):202-216.
7. Goodman SN, Zahurak ML, Piantadosi S. Some practical improvements in the continual reassessment method for phase I studies. *Stat Med*. 1995 Jun 15;14(11):1149-1161.
8. Faries D. Practical modifications of the continual reassessment method for phase I cancer clinical trials. *J Biopharm Stat*. 1994 Jul;4(2):147-164.
9. Korn EL, Midthune D, Chen TT, Rubinstein LV, Christian MC, Simon RM. A comparison of two phase I trial designs. *Stat Med*. 1994 Sep 30;13(18):1799-1806.
10. Moller S. An extension of the continual reassessment methods using a preliminary up-and-down design in a dose finding study in cancer patients, in order to investigate a greater range of doses. *Stat Med*. 1995 May 15–30;14(9–10):911-922. *discussion 923*
11. Piantadosi S, Fisher JD, Grossman S. Practical implementation of a modified continual reassessment method for dose-finding trials. *Cancer Chemother Pharmacol*. 1998;41(6):429-436.
12. Heyd JM, Carlin BP. Adaptive design improvements in the continual reassessment method for phase I studies. *Stat Med*. 1999 Jun 15;18(11):1307-1321.
13. Chu P-L, Lin Y, Shih WJ. Unifying CRM and EWOC designs for phase I cancer clinical trials. *Journal of Statistical Planning and Inference*. 2009;139:1146-1163.
14. Lin Y, Shih WJ. Statistical properties of the traditional algorithm-based designs for phase I cancer clinical trials. *Biostatistics*. 2001 Jun;2(2):203-215.
15. Storer BE. Design and analysis of phase I clinical trials. *Biometrics*. 1989 Sep;45(3):925-937.
16. Tighiouart M, Rogatko A. Dose finding with escalation with overdose control (EWOC) in cancer clinical trials. *Stat Sci*. 2010;25(2):217-226.
17. Cheung YK. Coherence principles in dose-finding studies. *Biometrika*. 2005;92:863-873.
18. Lee SM, Cheng B, Cheung YK. Continual reassessment method with multiple toxicity constraints. *Biostatistics*. 2011 Apr;12(2):386-398.

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of the article.

How to cite this article: Ji L, Lewinger JP, Krailo M, et al. Improvements to the Escalation with Overdose Control design and a comparison with the restricted Continual Reassessment Method. *Pharmaceutical Statistics*. 2019;18:659–670. https://doi.org/10.1002/pst.1955