STATISTICAL ANALYSIS

original reports abstract

Would the Recommended Dose Have Been Different Using Novel Dose-Finding Designs? Comparing Dose-Finding Designs in Published Trials

Rebecca B. Silva, BA[1]; Christina Yap, PhD[2]; Richard Carvajal, MD[3]; and Shing M. Lee, PhD[1,3]

PURPOSE Simulation studies have shown that novel designs such as the continual reassessment method and the Bayesian optimal interval (BOIN) design outperform the 3 + 3 design by recommending the maximum tolerated dose (MTD) more often, using less patients, and allotting more patients to the MTD. However, it is not clear whether these novel designs would have yielded different results in the context of real-world dose-finding trials. This is a commonly mentioned reason for the continuous use of 3 + 3 designs for oncology trials, with investigators considering simulation studies not sufficiently convincing to warrant the additional design complexity of novel designs.

METHODS We randomly sampled 60 published dose-finding trials to obtain 22 that used the 3 + 3 design, identified an MTD, published toxicity data, and had more than two dose levels. We compared the published MTD with the estimated MTD using the continual reassessment method and BOIN using target toxicity rates of 25% and 30% and toxicity data from the trial. Moreover, we compared patient allocation and sample size assuming that these novel designs had been implemented.

RESULTS Model-based designs chose dose levels higher than the published MTD in about 40% of the trials, with estimated and observed toxicity rates closer to the target toxicity rates of 25% and 30%. They also assigned less patients to suboptimal doses and permitted faster dose escalation.

CONCLUSION This study using published dose-finding trials shows that novel designs would recommend different MTDs and confirms the advantages of these designs compared with the 3 + 3 design, which were demonstrated by simulation studies.

JCO Precis Oncol 5:1024-1034. © 2021 by American Society of Clinical Oncology

INTRODUCTION

Phase I oncology clinical trials are designed to find the recommended dose of a treatment. For most therapeutic agents, the assumption is that a higher dose will be more efficacious although the risk of toxicity increases with higher dose. Hence, it is common to use the maximum tolerated dose (MTD) to define the recommended dose. An optimal design should maximize the frequency in the recommendation and assignment of patients to the true MTD and minimize the recommendation and assignment of patients to subtherapeutic or highly toxic doses. The challenge in dose finding is to balance simplicity in implementation, effectiveness in optimizing patient outcomes, and safety in minimizing risk.

Traditional rule-based methods such as the 3 + 3 design[1] continue to be the most commonly used[2,3] method for finding the MTD in phase I clinical trials because of their simplicity. Criticisms of the 3 + 3 design include slow dose escalation and a rigid structure that does not allow for flexibility in defining the target toxicity rate (TTR) or incorporation of previous data.[4,5] Moreover, multiple simulation studies have shown that novel designs such as the continual reassessment method (CRM)[6] and the Bayesian optimal interval (BOIN) design[7] outperform the 3 + 3 design in terms of identifying the true MTD more often while using less patients, allotting more patients to the MTD, and accommodating the design to suit specific drug characteristics.[7-11]

Simulation studies have been the primary tool used to compare these designs since they allow researchers to assume a true MTD and evaluate the frequency with which various designs select the correct dose under various trial scenarios. Using this approach, the 3 + 3 design gives biased estimates of the MTD by selecting a dose whose toxicity rate is lower than the TTR[12] and the CRM exposes fewer patients to highly toxic doses compared with the 3 + 3 design.[13] Simulation studies have also demonstrated comparable average performance between the CRM and BOIN designs.[14] Additionally, Conaway and Petroni[15] found that using the

ASSOCIATED CONTENT

Appendix

Author affiliations and support information (if applicable) appear at the end of this article.

Accepted on May 6, 2021 and published at ascopubs.org/journal/po on June 15, 2021: DOI https://doi.org/10.1200/PO.21.00136

CONTEXT

Key Objective

Novel dose-finding designs have been shown in previous simulation studies to outperform the 3 + 3 design in terms of identifying the maximum tolerated dose (MTD). However, it is not clear whether these novel designs would assign or recommend different doses in the context of real dose-finding trials applications. This study applied the continual reassessment method and the Bayesian optimal interval to 22 randomly selected published dose-finding studies that used the 3 + 3 design.

Knowledge Generated

In approximately 40% of trials, the recommended MTD using novel designs was higher than the published MTD using 3 + 3 design, with estimated and observed toxicity rates closer to the target toxicity rates of 25% and 30%. The difference was more pronounced with the increasing number of dose levels.

Relevance

The use of novel dose-finding methods could lead to different dose assignments and recommendations compared with the 3 + 3 design in the context of real dose-finding trials.

CRM and BOIN designs in phase I yields a substantially higher proportion of effective agents in successful phase III trials compared with the 3 + 3 design.

Despite the advantages of these novel designs in simulation studies, the CRM and BOIN are yet to be fully accepted and adopted in clinical research.[2,16,17] This slow uptake could be due in part to the designs' statistical complexity or clinicians' lack of familiarity with the designs.[18,19] However, these barriers have been largely addressed through the addition of tutorial papers on implementation, modifications of designs,[19-22] new software, and acknowledgment of adaptive designs in guidance documents[23] with limited success. Another common reason mentioned by clinical investigators is that simulation studies are not sufficiently convincing to warrant the additional design complexity given that these novel designs have not been compared in the context of real dose-finding clinical trials to evaluate the trial-specific level differences in dose assignments and selected MTD. In this paper, we apply the CRM and BOIN design to published phase I oncology trials that use the 3 + 3 design and compare the estimated MTD on the basis of the observed data as well as if we had applied the novel designs for each dose assignment. This allows us to evaluate the trial-specific level benefits of using novel designs.

METHODS

To obtain a representative sample of at least 20 dose-finding trials, we randomly sampled 60 rule-based studies published from 2008 to 2014 from Chiuzan et al.[3] We first sampled 30 and obtained 11 trials after excluding designs that did not reach an MTD or include toxicity data, those that did not follow the standard 3 + 3 design in practice, and those that had less than three dose levels. We then sampled another 30, which yielded 11 more eligible trials (Fig 1). Thus, a total of 22 trials were systematically reviewed to obtain the number of dose-limiting toxicities (DLTs) and patients at each dose, which we refer to as the published data for each trial. We estimated the MTD using two approaches, first applying the CRM and BOIN to the published data and second applying each design for each dose assignment and implementing design-specific characteristics, such as cohort size and starting dose on the basis of the recommendations for these designs. The MTDs were then compared with the original published MTD.

Using Published Data

We estimated the MTD with the published toxicity data using the CRM and BOIN designs assuming the TTR of 0.25 and 0.30 without using any simulations. Rates of 0.25 and 0.30 are common targets used to compare the 3 + 3 design with the CRM and BOIN, respectively.[10,24] For the CRM, we used the Bayesian estimation in the R package "dfcrm"[21] and selected the skeleton using the approach by Lee and Cheung,[25] specifying the prior guess of MTD as the median or greatest dose level below the median number of total doses. Since the CRM requires the order of dose assignment and most publications only list the number of patients assigned to each dose level, if the order was not specified, we obtained the most plausible order of dose escalation assuming a 3 + 3 design. For example, given six patients on dose level one, followed by three on dose level two, we assumed that there was one DLT in the first cohort of three and zero in the subsequent cohort on dose level one. For the BOIN, we used the R package "BOIN"[24] and the function select.mtd, which fit an isotonic regression to select the dose at which the fitted DLT rate is closest to the TTR. We did not include the expansion cohort data here because these cohorts were not considered for dose escalation in the 3 + 3 designs. We compared the selected MTD using the CRM and BOIN designs with the one in the original published paper.

Using Novel Methods for Actual Dose Assignments

We applied the CRM and BOIN from the start of the trial and assigned doses on the basis of the recommendations for

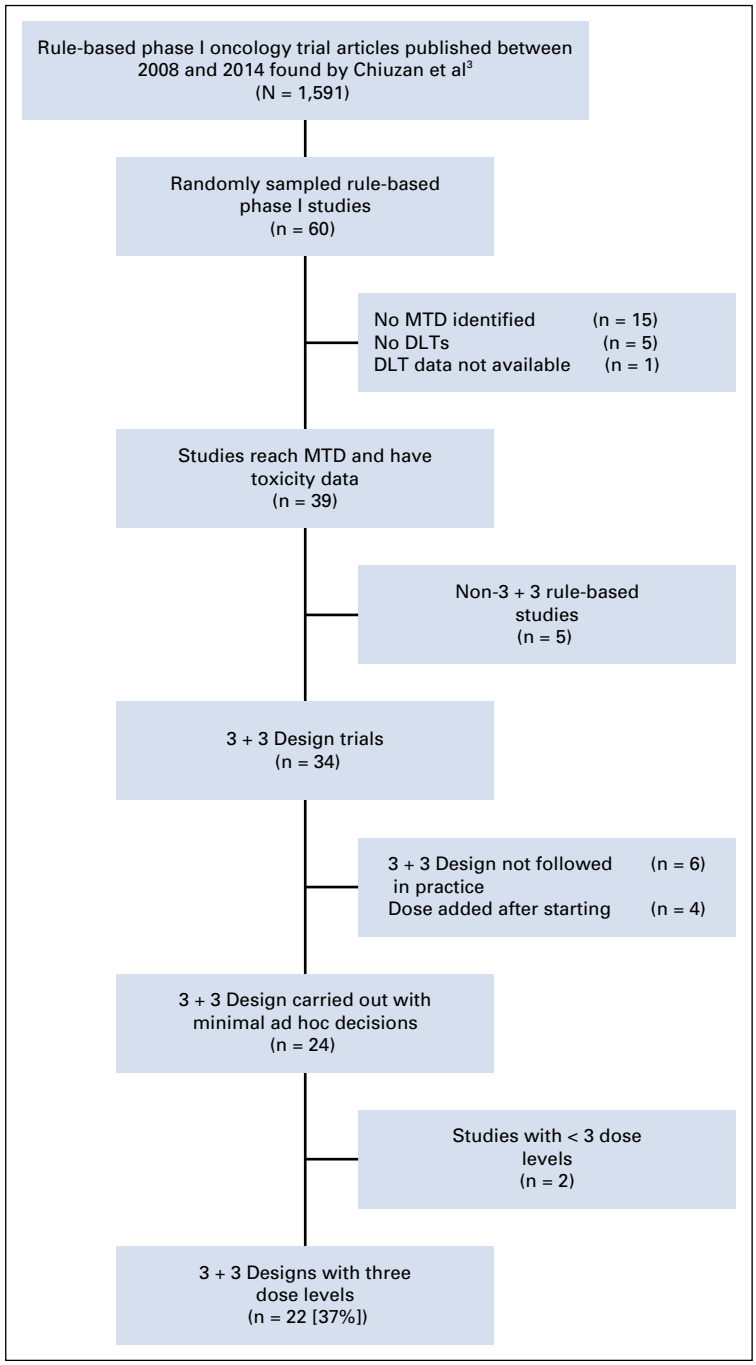

FIG 1. Selection of dose-finding trials. DLT, dose-limiting toxicity; MTD, maximum tolerated dose.

these methods. If the dose assignment was to a dose level for which we had published data and the design used a cohort size of 3, the DLT information for the entire cohort was taken. If the design used a cohort size of one, we sampled the DLT information without replacement from the cohort. If published data were not available or already used, we generated data with a probability of toxicity equal to the observed rate of DLT at the given dose level or the isotonized DLT rate to ensure monotonicity in the dose-toxicity relationship.

For the CRM, we used the crm function from the "dfcrm" package with each new cohort and getprior function for our initial skeleton. We started each trial at the median dose level, did not allow for dose skipping, and considered cohort sizes of one, CRM(1), or three, CRM(3). The sample size was calculated on the basis of the approach by Cheung[26] with a probability of correction selection of 0.60 and an odds ratio of two. We chose the sample size for all designs as the closest number larger than or equal to the recommended size that was a multiple of three.

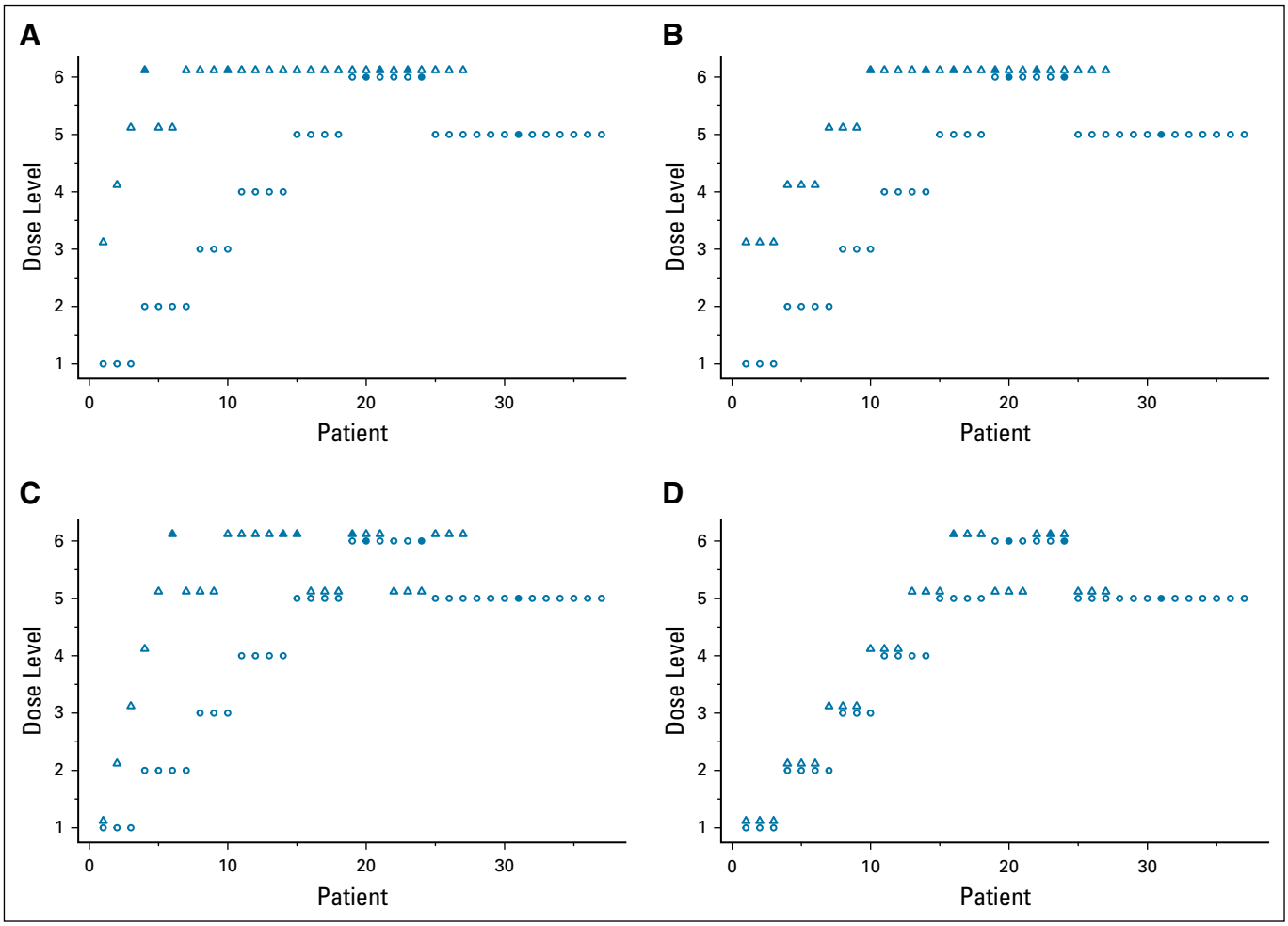

FIG 2. Dose assignment for one retraced trial by Mita et al[28] compared with original data. (A) CRM(1), (B) CRM(3), (C) BOIN(A), and (D) BOIN(3). The original data are denoted by circles, and the compared method is denoted by triangles. Solid indicates a dose-limiting toxicity. BOIN, Bayesian optimal interval; CRM, continual reassessment method.

For the BOIN, we used the get.boundary function from the "BOIN" package, with 0.6 × TTR and 1.4 × TTR as the highest toxicity probability deemed subtherapeutic and overly toxic, respectively, a cutoff of 0.95 to eliminate an overly toxic dose, and an offset of 0.05[24] to obtain the escalation and de-escalation parameters. The select.mtd function was applied after the maximum sample size recommended for the CRM was reached. We considered an accelerated BOIN using cohorts of one until the first DLT is observed,[7] BOIN(A), and cohorts of three throughout, BOIN(3). We started both designs at the first dose level.[27] For both CRM and BOIN, we repeated this process 1,000 times for each trial and obtained the following measures of comparison from each trial, design, and TTR: the proportion that each dose was selected as the MTD, the proportion and number of patients assigned to each dose, the proportion of patients with toxicities at each dose, and the proportion of patients generated.

To contrast the designs, dose assignment and DLT using the trial by Mita et al[28] are displayed in Figure 2 along with dose assignments had the CRM and BOIN been implemented with a sample size of 27 and a TTR of 0.25. The 3 + 3 design had six dose levels with 24 patients and recommended dose level 5 as the MTD with zero DLTs of four patients and two of six at the above dose. It also had an expansion cohort where one of 13 patients experienced a DLT. Both the CRM and BOIN designs recommended dose level 6 as the MTD instead. Starting at dose level 3, the CRMs homed in on the eventual MTD faster and assigned more patients to dose level 6 compared with the BOIN methods, starting at dose level 1.

RESULTS

Table 1 summarizes the observed DLT data from each trial of the 22 trials ordered by the number of dose levels in the trial. The median number of dose levels was four, and the median sample size was 18. Eight trials (36%) included an expansion cohort. The median sample size including expansion cohorts was 21.5.

Using Published Data

The recommended MTD level in the published paper and the one obtained from applying the CRM and BOIN to the

TABLE 1. No. of Dose-Limiting Toxicities and Patients Assigned to Each Dose Level in the Published Study

	Dose Levels									
Trial	**d1**	**d2**	**d3**	**d4**	**d5**	**d6**	**d7**	**d8**	**d9**	**d10**
Berenson et al[29]	0/3	**0/3**	3/6							
Frost et al[30]	1/6	**0/6 (1/6)**	2/2							
Kunz et al[31,a]	0/3	**1/9**	1/6							
Ghobrial et al[32]	0/3	0/3	0/3	**1/6**						
Kim et al[33]	0/3	0/6	**1/6**	3/6						
Ma et al[34,a]	0/1	0/1	**0/5 (0/7)**	2/5						
Oki et al[35]	0/3	0/3	**0/3 (0/19)**	2/2						
Pollyea et al[36]	0/5	0/3	0/4	**1/6**						
Sadahiro et al[37]	0/3	0/3	**0/3**	2/6						
Sanborn et al[38]	0/3	0/3	**1/6**	2/4						
Simonelli et al[39]	0/3	0/3	1/6	**1/6 (0/9)**						
Tevaarwerk et al[40]	0/3	1/6	**0/3**	3/3						
Gerecitano et al[41]	0/3	0/3	0/3	**0/7 (0/4)**	2/2					
Jakacki et al[42]	0/3	0/3	0/3	**1/6 (2/17)**	2/4					
Kurzrock et al[43]	0/3	0/3	0/3	**0/3**	2/6					
Wood et al[44]	0/3	0/6	0/6	**1/6**	2/3					
Mita et al[28]	0/3	0/4	0/3	0/4	**0/4 (1/13)**	2/6				
Garcia et al[45]	0/4	0/4	0/3	0/4	0/5	**0/5**	2/6			
Kantarjian et al[46]	0/6	0/3	0/4	0/3	0/3	**1/6**	2/4			
Harada and Omura[47]	0/3	0/3	0/3	0/3	0/3	0/6	**0/3**	2/6		
Younes et al[48]	0/3	0/4	0/3	0/3	0/3	0/4	**1/12**	3/12	1/1	
van Laarhoven et al[49,b]	0/3	0/3	0/5	1/7	0/3	0/3	0/3	0/3	0/3	1/6

NOTE. Maximum tolerated dose identified is in bold; the numbers in parentheses are additional patients enrolled in an expansion cohort.
[a]Maximum tolerated dose selection was not based on 3 + 3 design.
[b]van Laarhoven et al have 17 dose levels for d11-d15: 0/3, d16: **0/6**, and d17: 2/3. Expansion cohort for d16: 0/4.

published data are shown in Table 2. The recommended MTD by the CRM, using a TTR of 0.25, matched the published MTD in 12 trials (55%), was higher by one dose level in nine trials (41%), and was lower by one dose level in one trial (4%). The estimated DLT rate for the chosen dose ranged from 0.13 to 0.31. The recommended MTD by the BOIN, using a TTR of 0.25, matched the published MTD in 14 trials (64%) and was higher by one dose level in eight trials (36%). The estimated DLT rate for the chosen dose ranged from approximately 0.01-0.40.

For the eight trials for which both the CRM and BOIN recommended higher doses, the DLT rate in the published data was between 0.18 and 0.40, whereas the observed DLT rate of the dose below, chosen by the original design, was ≤ 0.167, with a median DLT rate of 0 (Table 1). The estimated DLT rate by the CRM and BOIN for a TTR of 0.25 was below 0.31 and 0.40, respectively. There were nine trials for the CRM and one for the BOIN in which the design with a TTR of 0.30 chose a higher MTD than that chosen with a TTR of 0.25.

Using Novel Methods for Actual Dose Assignments

Comparison of the MTD selected with a TTR of 0.25 compared with the original study is displayed in Figure 3. For both cohort sizes, the MTD selected by the CRM and BOIN was similar most of the time. In general, the designs would have recommended one dose level higher than the recommended MTD in the published paper in 8 (36%) of the 22 trials most of the time and the same dose as the published paper in 14 (64%). For trials in which the designs recommended one dose higher, the observed DLT rate of the chosen dose was between 0.17 and 0.40, which was closer to the TTR of 0.25 than the dose below. In six of eight of these trials, the 3 + 3 design chose the dose level below with a DLT rate of 0. The trials for which the MTD was the same had a DLT rate of 0.17 or less with the dose above the chosen MTD having a rate ≥ 0.50. The BOIN(3) results differed from the CRM and BOIN(A) for the trials by Kunz et al[31] and Kim et al.[33] Increasing the sample size by at least three would result in a recommended MTD that matched the other designs. The CRM(3) chose the dose level above

TABLE 2. Comparison of Dose Level Recommended as the MTD With the Original Published Data

Trial	Publication MTD	MTD Using CRM		MTD Using BOIN	
		DLT Rate		DLT Rate	
		25%	30%	25%	30%
Berenson et al[29]	2	2	**3**	2	**3**
Frost et al[30]	2	2	2	2	2
Kunz et al[31]	2[a]	**3**	**3**	**3**	**3**
Ghobrial et al[32]	4	4	4	4	4
Kim et al[33]	3	3	**4**	3	3
Ma et al[34]	3[a]	**4**	**4**	**4**	**4**
Oki et al[35]	3	3	3	3	3
Pollyea et al[36]	4	4	4	4	4
Sadahiro et al[37]	3	**4**	**4**	**4**	**4**
Sanborn et al[38]	3	3	**4**	3	3
Simonelli et al[39]	4	4	4	4	4
Tevaarwerk et al[40]	3	**2**	3	3	3
Gerecitano et al[41]	4	4	**5**	4	4
Jakacki et al[42]	4	4	**5**	4	4
Kurzrock et al[43]	4	**5**	**5**	**5**	**5**
Wood et al[44]	4	4	**5**	4	4
Mita et al[28]	5	**6**	**6**	**6**	**6**
Garcia et al[45]	6	**7**	**7**	**7**	**7**
Kantarjian et al[46]	6	**7**	**7**	6	6
Harada and Omura[47]	7	**8**	**8**	**8**	**8**
Younes et al[48]	7	**8**	**9**	**8**	**8**
van Laarhoven et al[49]	16	16	**17**	16	16

NOTE. Numbers in bold indicate different dose levels selected from MTD in the publication.

Abbreviations: BOIN, Bayesian optimal interval; CRM, continual reassessment method; DLT, dose-limiting toxicity; MTD, maximum tolerated dose.

[a]MTD selection differs from 3 + 3 design.

more frequently than the CRM(1) and BOIN designs in the trial by Tevaarwerk et al[40] and Laarhoven et al.[49]

With a TTR of 0.30, shown in Appendix Figure A1, the CRM was more aggressive for a few of the trials and the BOIN(3) more closely matched the other designs. The inconsistency between BOIN(3) with the TTR of 0.25 and 0.30 is a result of slow dose escalation with a TTR of 0.25 since it de-escalates after observing a DLT rate higher than 0.30, which is more conservative than the 3 + 3 design.

For trials with less than five dose levels, the sample sizes recommended by the CRM were typically similar to those used in the 3 + 3 design without an expansion cohort. For trials with five or more dose levels, the CRM had a smaller sample size than what was typically used in 3 + 3 designs. Figure 4 shows the distribution of the proportion of patients assigned to the respective MTD of each method, excluding the expansion cohort for the 3 + 3 design. The proportions of patients assigned to the respective MTD in the BOIN(A) and both CRM designs were higher than those in the original 3 + 3, most notably with the CRM designs. For trials without an expansion cohort, or independent of their expansion cohort, the original design tended to assign a similar if not greater proportion of patients to doses with high DLT rates (>0.50) compared with the CRM(1), CRM(3), and BOIN(3).

It should be noted that given that the CRM started at a higher dose level, approximately half of the patient DLT information was simulated; for the BOIN that started at the same dose as the original trial, about a quarter was simulated (Appendix Table A1). Finally, stopping rules can be imposed for the CRM and BOIN designs, although they were not included in this study.

DISCUSSION

Using published data from 22 dose-finding clinical trials, we demonstrated that using novel designs, the selected MTD would have been the dose above the published MTD for approximately 40% of the trials, both using the published data or if the novel design had been used for the actual dose assignments instead of the 3 + 3 design. The doses selected by the novel designs were closer to the target rate than the previous dose, which had observed rates of 0 most of the time. The CRM with a TTR of 0.30 sometimes resulted in a more aggressive MTD selection when using the published data, but matched closely to the CRM with a TTR of 0.25 when using the CRM for each dose assignment. However, we recommend using a TTR of 0.25 since a TTR of 0.30 does increase the selection of overly toxic doses. On the contrary, when applying the BOIN to published data, the recommended MTD across TTRs was similar, but when using the BOIN at each dose assignment, with a TTR of 0.25, the BOIN(3) sometimes had slow dose escalation, which allowed less patients to be evaluated at higher doses. Therefore, a TTR of 0.30 is recommended for the BOIN design or using the option of having a 3 + 3 design run with a TTR of 0.25. Applying both the CRM and BOIN to published data was for illustration purposes only, demonstrating that even fitting a model at the end using data that were produced by the 3 + 3 design would have given a different dose selection more than a third of the time.

Although some investigators may assume that the 3 + 3 design targets right below 33%, it actually overlooks any dose with an observed toxicity rate higher than 1/6, which is lower than that previous simulations have found.[50] For example, in the trial by Younes et al,[48] the 3 + 3 design selected a dose with a DLT rate of 1/12 instead of selecting the dose with a DLT rate of 0.25 (3/12). Although the 3 + 3 design is favored for its simplicity, its rule-based nature does not allow for flexibility in choosing a range of acceptable DLT rates or in making common ad hoc decisions such as altering cohort sizes or adding an intermediate dose mid-trial.

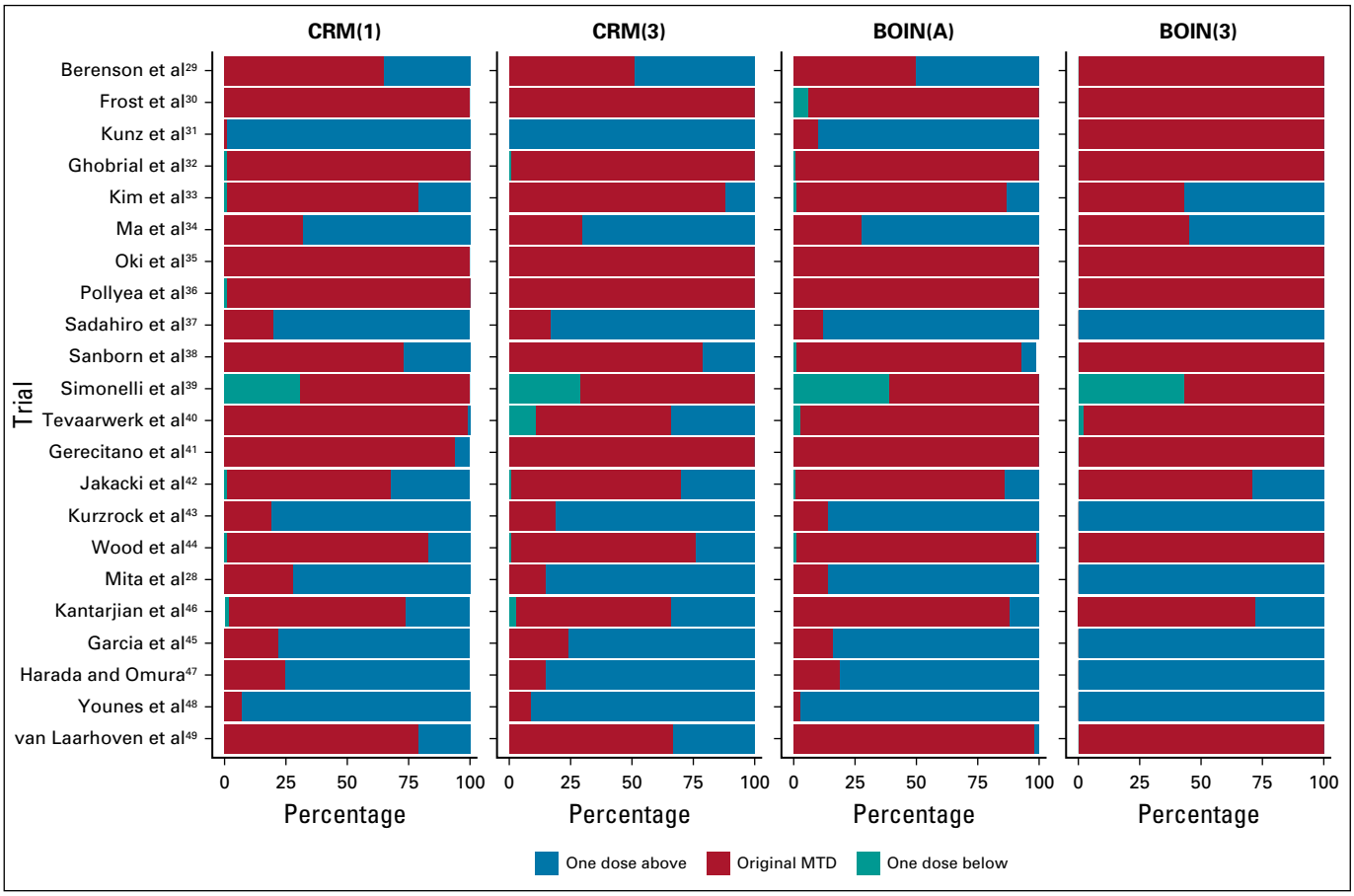

FIG 3. Selection proportion of MTD with a target toxicity rate of 0.25 compared with the original published data. Information within parentheses of column headers denotes cohort size or structure. BOIN, Bayesian optimal interval; CRM, continual reassessment method; MTD, maximum tolerated dose.

Our study allows comparison of trials with a range of dose levels. For trials with three levels, the simplicity of the 3 + 3 design is beneficial and can offer similar efficiency. As the number of dose levels increased, model-based designs in general had smaller sample sizes and treated more patients around the MTD. The 3 + 3 design is inefficient as we found 21 trials (95%) had their first DLT at a dose level greater than or equal to the median dose level and 22 trials (100%) chose an MTD at or above this dose level. The BOIN(A), CRM(1), and CRM(3) all mitigate inefficiency through faster dose escalation, starting at median dose level or both.

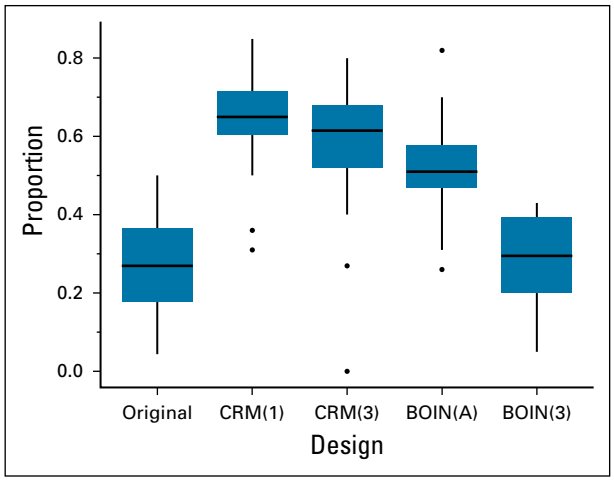

FIG 4. Distribution of the proportion of patients assigned to the chosen maximum tolerated dose of each respective design across all 22 trials, with a target toxicity rate of 0.25. Information within parentheses on x-axis denotes cohort size or structure. BOIN, Bayesian optimal interval; CRM, continual reassessment method.

Our approach for comparing the 3 + 3, CRM, and BOIN designs against each other is novel in its ability to compare them in the context of real trials. However, one limitation is that we do not know the true MTD or DLT rates like in a simulation study. Thus, we can only compare the recommended MTDs against each other and make judgements on the basis of the observed and estimated DLT rates. Moreover, we were not able to connect most trials to later phase trials or approved dosages since many did not lead to successful trials. High attrition rates pose a large problem in the success of oncology trials, which could be due in part to overly conservative MTDs that do not lead to efficacy in later phases.

In summary, our study using data from 22 dose-finding trials suggests that approximately 40% of the studies would

have recommended one dose higher if the CRM or BOIN with a TTR of 25%-30% were used. It also confirms the advantages of the CRM and BOIN that have been demonstrated through simulation studies mentioned earlier. Finally, model-based designs while more complex also offer more flexibility such as inclusion of ad hoc decisions, late-onset toxicities, dose combinations, and efficacy information, which are crucial for novel anticancer treatments.

AFFILIATIONS

[1]Department of Biostatistics, Columbia University Mailman School of Public Health, New York, NY
[2]Clinical Trials and Statistics Unit, The Institute of Cancer Research, London, United Kingdom
[3]Herbert Irving Comprehensive Cancer Center, Columbia University, New York, NY

CORRESPONDING AUTHOR

Shing M. Lee, PhD, Department of Biostatistics, Columbia University Mailman School of Public Health, 722 W. 168th Street, New York, NY 10032; e-mail: sml2114@columbia.edu.

SUPPORT

Supported by National Institute of Health IMSD Scholarship (T32) and TL1 Award and National Center for Advancing Translational Sciences, National Institutes of Health (Grant No. UL1TR001873).

AUTHOR CONTRIBUTIONS

Conception and design: Christina Yap, Shing M. Lee
Collection and assembly of data: Rebecca B. Silva
Data analysis and interpretation: All authors
Manuscript writing: All authors
Final approval of manuscript: All authors
Accountable for all aspects of the work: All authors

AUTHORS' DISCLOSURES OF POTENTIAL CONFLICTS OF INTEREST

The following represents disclosure information provided by the authors of this manuscript. All relationships are considered compensated unless otherwise noted. Relationships are self-held unless noted. I = Immediate Family Member, Inst = My Institution. Relationships may not relate to the subject matter of this manuscript. For more information about ASCO's conflict of interest policy, please refer to www.asco.org/rwc or ascopubs.org/po/author-center.

Open Payments is a public database containing information reported by companies about payments made to US-licensed physicians (Open Payments).

Christina Yap

Consulting or Advisory Role: Faron Pharmaceuticals
Travel, Accommodations, Expenses: Faron Pharmaceuticals

Richard Carvajal

Consulting or Advisory Role: Merck, Aura Biosciences, Castle Biosciences, Immunocore, PureTech, Sorrento Therapeutics, Chimeron Bio, Rgenix, InxMed, Pierre Fabre, TriSalus Life Sciences, Iovance Biotherapeutics, Oncosec, Regeneron, Genzyme, Amgen, Astellas Pharma, AstraZeneca, Bristol Myers Squibb/Medarex, Corvus Pharmaceuticals, Ideya, Mirati Therapeutics, Novartis, Pfizer, Plexxikon, Roche/Genentech
Speakers' Bureau: Bristol Myers Squibb/Medarex
Research Funding: Amgen, Astellas Pharma, AstraZeneca, Bayer, Bellicum Pharmaceuticals, Bristol Myers Squibb, Corvus Pharmaceuticals, Lilly, Immunocore, Incyte, Macrogenics, Merck, Mirati Therapeutics, Novartis, Pfizer, Plexxikon, Roche/Genentech, Array BioPharma, IDEAYA Biosciences, Regeneron

Shing M. Lee

Consulting or Advisory Role: PTC Therapeutics
Research Funding: Merck, Genentech/Roche

No other potential conflicts of interest were reported.

REFERENCES

1. Storer BE: Design and analysis of phase I clinical trials. Biometrics 45:925-937, 1989
2. van Brummelen EMJ, Huitema ADR, van Werkhoven E, et al: The performance of model-based versus rule-based phase I clinical trials in oncology: A quantitative comparison of the performance of model-based versus rule-based phase I trials with molecularly targeted anticancer drugs over the last 2 years. J Pharmacokinet Pharmacodyn 43:235-242, 2016
3. Chiuzan C, Shtaynberger J, Manji GA, et al: Dose-finding designs for trials of molecularly targeted agents and immunotherapies. J Biopharm Stat 27:477-494, 2017
4. Le Tourneau C, Lee JJ, Siu LL: Dose escalation methods in phase I cancer clinical trials. J Natl Cancer Inst 101:708-720, 2009
5. Paoletti X, Ezzalfani M, Le Tourneau C: Statistical controversies in clinical research: Requiem for the 3 + 3 design for phase I trials. Ann Oncol 26:1808-1812, 2015
6. O'Quigley J, Pepe M, Fisher L: Continual reassessment method: A practical design for phase 1 clinical trials in cancer. Biometrics 46:33-48, 1990
7. Yuan Y, Hess KR, Hilsenbeck SG, et al: Bayesian optimal interval design: A simple and well-performing design for phase I oncology trials. Clin Cancer Res 22:4291-4301, 2016
8. Thall PF, Lee S-J: Practical model-based dose-finding in phase I clinical trials: Methods based on toxicity. Int J Gynecol Cancer 13:251-261, 2003
9. Babb J, Rogatko A, Zacks S: Cancer phase I clinical trials: Efficient dose escalation with overdose control. Stat Med 17:1103-1120, 1998
10. Iasonos A, Wilton AS, Riedel ER, et al: A comprehensive comparison of the continual reassessment method to the standard 3 + 3 dose escalation scheme in phase I dose-finding studies. Clin Trials 5:465-477, 2008
11. Storer BE: An evaluation of phase I clinical trial designs in the continuous dose–response setting. Stat Med 20:2399-2408, 2001
12. He W, Liu J, Binkowitz B, et al: A model-based approach in the estimation of the maximum tolerated dose in phase I cancer clinical trials. Stat Med 25:2027-2042, 2006
13. O'Quigley J: Another look at two phase I clinical trial designs. Stat Med 18:2683-2690, 1999; discussion 2691-2692
14. Zhu Y, Hwang W-T, Li Y: Evaluating the effects of design parameters on the performances of phase I trial designs. Contemp Clin Trials Commun 15:100379, 2019

15. Conaway MR, Petroni GR: The impact of early-phase trial design in the drug development process. Clin Cancer Res 25:819-827, 2019
16. Rogatko A, Schoeneck D, Jonas W, et al: Translation of innovative designs into phase I trials. J Clin Oncol 25:4982-4986, 2007
17. Jaki T: Uptake of novel statistical methods for early-phase clinical studies in the UK public sector. Clin Trials 10:344-346, 2013
18. Love SB, Brown S, Weir CJ, et al: Embracing model-based designs for dose-finding trials. Br J Cancer 117:332-339, 2017
19. Wheeler GM, Mander AP, Bedding A, et al: How to design a dose-finding study using the continual reassessment method. BMC Med Res Methodol 19:18, 2019
20. Garrett-Mayer E: The continual reassessment method for dose-finding studies: A tutorial. Clin Trials 3:57-71, 2006
21. Cheung YK: Dose Finding by the Continual Reassessment Method. Boca Raton, FL, CRC Press, 2011
22. O'Quigley J, Iasonos A, Bornkamp B: Handbook of Methods for Designing, Monitoring, and Analyzing Dose-Finding Trials. Boca Raton, FL, CRC Press, 2017
23. FDA: Adaptive designs for clinical trials of drugs and biologics. 2019. https://www.fda.gov/regulatory-information/search-fda-guidance-documents/adaptive-design-clinical-trials-drugs-and-biologics-guidance-industry
24. Yuan Y, Liu S: Bayesian optimal INterval (BOIN) design for single-agent and drug-combination phase I clinical trials. https://cran.r-project.org/web/packages/BOIN/BOIN.pdf
25. Lee SM, Cheung YK: Model calibration in the continual reassessment method. Clin Trials 6:227-238, 2009
26. Cheung YK: Sample size formulae for the Bayesian continual reassessment method. Clin Trials 10:852-861, 2013
27. Liu S, Yuan Y: Bayesian optimal interval designs for phase I clinical trials. J R Stat Soc Ser C Appl Stat 64:507-523, 2015
28. Mita M, Kelly KR, Mita A, et al: Phase I study of E7820, an oral inhibitor of integrin α-2 expression with antiangiogenic properties, in patients with advanced malignancies. Clin Cancer Res 17:193-200, 2011
29. Berenson JR, Yellin O, Patel R, et al: A phase I study of samarium lexidronam/bortezomib combination therapy for the treatment of relapsed or refractory multiple myeloma. Clin Cancer Res 15:1069-1075, 2009
30. Frost A, Mross K, Steinbild S, et al: Phase I study of the Plk1 inhibitor BI 2536 administered intravenously on three consecutive days in advanced solid tumours. Curr Oncol 19:e28-e35, 2012
31. Kunz PL, He AR, Colevas AD, et al: Phase I trial of ixabepilone administered as three oral doses each separated by 6 hours every 3 weeks in patients with advanced solid tumors. Invest New Drugs 30:2364-2370, 2012
32. Ghobrial IM, Weller E, Vij R, et al: Weekly bortezomib in combination with temsirolimus in relapsed or relapsed and refractory multiple myeloma: A multicentre, phase 1/2, open-label, dose-escalation study. Lancet Oncol 12:263-272, 2011
33. Kim KB, Hwu W-J, Papadopoulos NE, et al: Phase I study of the combination of docetaxel, temozolomide and cisplatin in patients with metastatic melanoma. Cancer Chemother Pharmacol 64:161-167, 2009
34. Ma T, Fuld AD, Rigas JR, et al: A phase I trial and in vitro studies combining ABT-751 with carboplatin in previously treated non-small cell lung cancer patients. Chemotherapy 58:321-329, 2012
35. Oki Y, Buglio D, Fanale M, et al: Phase I study of panobinostat plus everolimus in patients with relapsed or refractory lymphoma. Clin Cancer Res 19:6882-6890, 2013
36. Pollyea DA, Zehnder J, Coutre S, et al: Sequential azacitidine plus lenalidomide combination for elderly patients with untreated acute myeloid leukemia. Haematologica 98:591-596, 2013
37. Sadahiro S, Suzuki T, Maeda Y, et al: Phase I/II study of 24-hour infusion of irinotecan combined with oral UFT for metastatic colorectal cancer. Chemotherapy 54:140-146, 2008
38. Sanborn SL, Cooney MM, Dowlati A, et al: Phase I trial of docetaxel and thalidomide: A regimen based on metronomic therapeutic principles. Invest New Drugs 26:355-362, 2008
39. Simonelli M, Zucali PA, Lorenzi E, et al: Phase I pharmacokinetic and pharmacodynamic study of lapatinib in combination with sorafenib in patients with advanced refractory solid tumors. Eur J Cancer 49:989-998, 2013
40. Tevaarwerk AJ, Holen KD, Alberti DB, et al: Phase I trial of 2-methoxyestradiol (2ME2, Panzem®) NanoCrystal® dispersion (NCD®) in advanced solid malignancies. Clin Cancer Res 15:1460-1465, 2009
41. Gerecitano JF, Stephenson JJ, Lewis NL, et al: A Phase I trial of the kinesin spindle protein (Eg5) inhibitor AZD4877 in patients with solid and lymphoid malignancies. Invest New Drugs 31:355-362, 2013
42. Jakacki RI, Hamilton M, Gilbertson RJ, et al: Pediatric phase I and pharmacokinetic study of erlotinib followed by the combination of erlotinib and temozolomide: A Children's Oncology Group Phase I Consortium Study. J Clin Oncol 26:4921-4927, 2008
43. Kurzrock R, Goel S, Wheler J, et al: Safety, pharmacokinetics, and activity of EZN-2208, a novel conjugate of polyethylene glycol and SN38, in patients with advanced malignancies. Cancer 118:6144-6151, 2012
44. Wood BJ, Poon RT, Locklin J, et al: A phase I study of heat deployed liposomal doxorubicin during radiofrequency ablation for hepatic malignancies. J Vasc Interv Radiol 23:248-255.e7, 2012
45. Garcia JA, Mekhail T, Elson P, et al: Phase I/II trial of subcutaneous interleukin-2, granulocyte-macrophage colony-stimulating factor and interferon-α in patients with metastatic renal cell carcinoma. BJU Int 109:63-69, 2012
46. Kantarjian HM, Padmanabhan S, Stock W, et al: Phase I/II multicenter study to assess the safety, tolerability, pharmacokinetics and pharmacodynamics of AZD4877 in patients with refractory acute myeloid leukemia. Invest New Drugs 30:1107-1115, 2012
47. Harada H, Omura K: Preoperative concurrent chemotherapy with S-1 and radiotherapy for locally advanced squamous cell carcinoma of the oral cavity: Phase I trial. J Exp Clin Cancer Res 29:33, 2010
48. Younes A, Bartlett NL, Leonard JP, et al: Brentuximab vedotin (SGN-35) for relapsed CD30-positive lymphomas. N Engl J Med 363:1812-1821, 2010
49. van Laarhoven HWM, Fiedler W, Desar IME, et al: Phase I clinical and magnetic resonance imaging study of the vascular agent NGR-hTNF in patients with advanced cancers (European Organization for Research and Treatment of Cancer Study 16041). Clin Cancer Res 16:1315-1323, 2010
50. Ivanova A: Escalation, group and A + B designs for dose-finding trials. Stat Med 25:3668-3678, 2006

APPENDIX

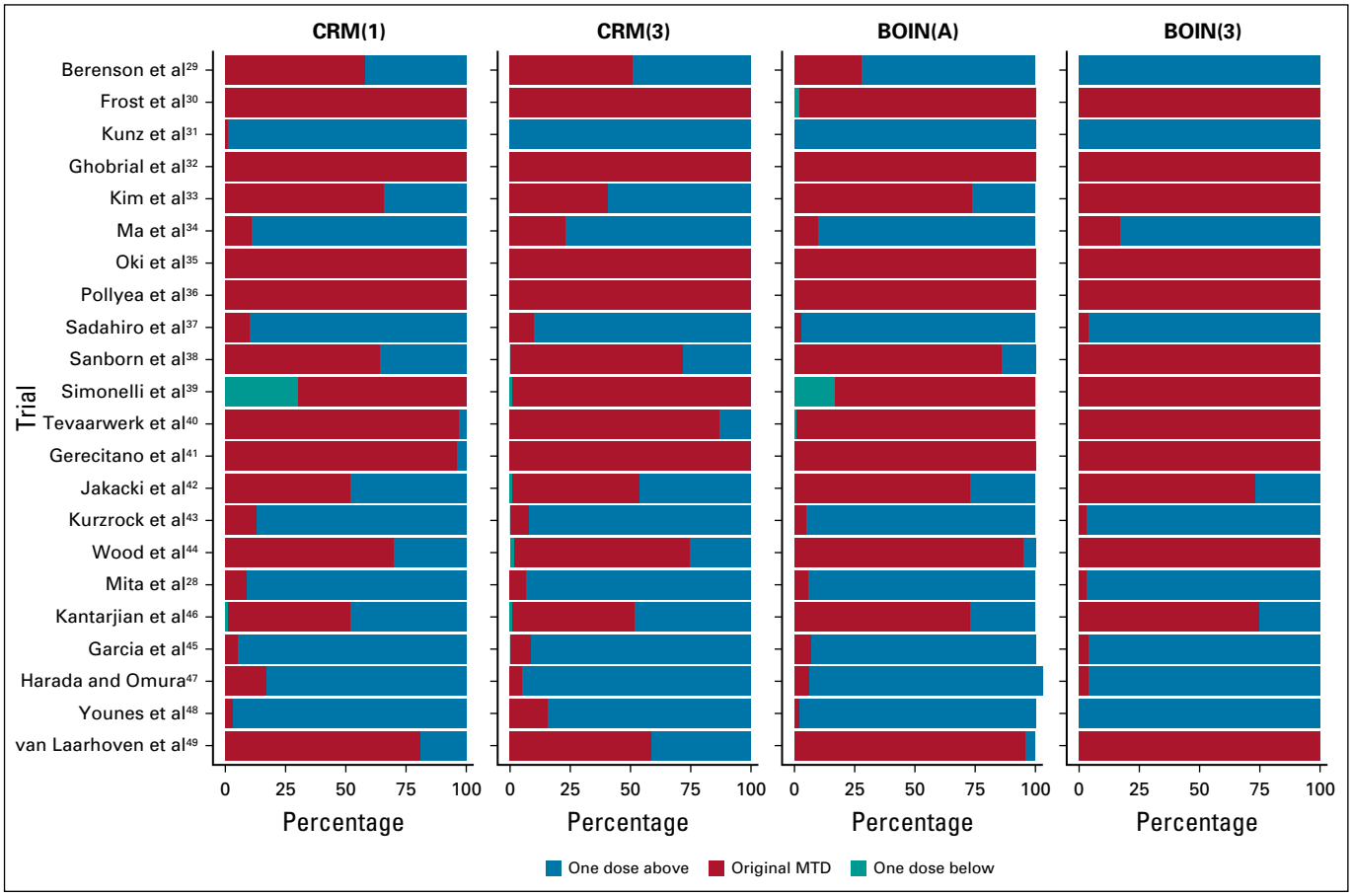

FIG A1. Selection percentage of MTD with a target toxicity rate of 0.30 compared with the original published data. Proportion is the number of times that design recommends dose as MTD of 1,000 retraced trials for each of the 22 trials. Refer to Table 1 for corresponding observed dose-limiting toxicity rates. Information within parentheses of column headers denotes cohort size or structure. BOIN, Bayesian optimal interval; CRM, continual reassessment method; MTD, maximum tolerated dose.

TABLE A1. Mean Proportion of Patient Toxicity Data Generated by Each Method With a Target Toxicity Rate of 0.25

Trial	CRM(1)	CRM(3)	BOIN(A)	BOIN(3)
Berenson et al[29]	0.40	0.40	0.19	0.20
Frost et al[30]	0.07	0.20	0.14	0.20
Kunz et al[31]	0.08	0.00	0.18	0.20
Ghobrial et al[32]	0.54	0.17	0.21	0.29
Kim et al[33]	0.36	0.38	0.16	0.14
Ma et al[34]	0.48	0.17	0.18	0.71
Oki et al[35]	0.71	0.71	0.27	0.57
Pollyea et al[36]	0.51	0.43	0.17	0.29
Sadahiro et al[37]	0.52	0.43	0.21	0.29
Sanborn et al[38]	0.47	0.47	0.20	0.29
Simonelli et al[39]	0.17	0.39	0.18	0.29
Tevaarwerk et al[40]	0.57	0.59	0.24	0.29
Gerecitano et al[41]	0.58	0.63	0.19	0.38
Jakacki et al[42]	0.49	0.50	0.21	0.25
Kurzrock et al[43]	0.58	0.50	0.21	0.25
Wood et al[44]	0.48	0.49	0.20	0.13
Mita et al[28]	0.56	0.44	0.18	0.22
Kantarjian et al[46]	0.54	0.50	0.21	0.20
Garcia et al[45]	0.58	0.50	0.18	0.20
Harada and Omura[47]	0.64	0.45	0.22	0.18
Younes et al[48]	0.24	0.26	0.08	0.09
van Laarhoven et al[49]	0.66	0.50	0.13	0.10

NOTE. Proportion of data not taken from the original data was determined in all 1,000 simulations of each trial. After original data were used on a dose level and more patients were assigned to that dose, data were generated. More data were simulated using the CRM designs since the design starts at the median dose. Information within parentheses of column headers denotes cohort size or structure.

Abbreviations: BOIN, Bayesian optimal interval; CRM, continual reassessment method.

