

Shortening the Timeline of Pediatric Phase I Trials: The Rolling Six Design

Jeffrey M. Skolnik, Jeffrey S. Barrett, Bhuvana Jayaraman, Dimple Patel, and Peter C. Adamson

ABSTRACT

Purpose

To shorten the study conduct timeline of pediatric phase I oncology trials by employing a novel trial design.

Methods

A comparison of the traditional 3 + 3 patients per cohort, phase I trial design with a novel, rolling six design was performed by using discrete event simulation. The rolling six design allows for accrual of two to six patients concurrently onto a dose level based on the number of patients currently enrolled and evaluable, the number experiencing dose-limiting toxicity (DLT), and the number still at risk of developing a DLT. Clinical trial simulations (n = 1,000) were based on historical data and were performed using SAS 9.1.3 (SAS Institute, Cary, NC). Study timelines and patient numbers were determined for each design, and safety was assessed as a function of the number of DLTs observed.

Results

In twelve completed historical studies, the median time to study completion was 452 days (range, 220 to 606 days); number of evaluable participants enrolled was 22 (range, 11 to 33), and DLTs occurring per study was three (range, 0 to 5). In 1,000 study simulations, in which the average time to new patient accrual was 10 days, the average \pm standard deviation (SD) time to study completion was 294 ± 75 days for the rolling six design versus 350 ± 84 days for the 3 + 3 design, whereas the number of DLTs per study was the same (average \pm SD, 3.3 ± 1.1 v 3.2 ± 1.1 for the rolling six and 3 + 3 designs, respectively).

Conclusion

The rolling six design may significantly decrease the duration of pediatric phase I studies without increasing the risk of toxicity. The design will be tested prospectively in upcoming Children's Oncology Group phase I trials.

J Clin Oncol 26:190-195. © 2008 by American Society of Clinical Oncology

INTRODUCTION

The development timeline of new agents for childhood cancer is inherently longer than for adult cancer. Multiple factors contribute to this, including the need to begin pediatric evaluation of new agents usually only after phase I adult trials are completed, the limited number of pediatric patients eligible for study, and the lack of pediatric cancer drug development initiatives undertaken by the pharmaceutical industry.¹ Efforts to shorten the overall timeline have increased during the past 5 years and have been impacted most notably by legislative initiatives, including the Best Pharmaceuticals for Children Act. There have not been changes, however, in the timeline associated with actual phase I trial conduct for the last 40 years.

Phase I studies in oncology are a critical first step in the evaluation of novel anticancer agents.²

Adult and pediatric phase I studies are similar, in that both determine a recommended phase II dose, often the maximum-tolerated dose (MTD), as their primary objective and both use dose-limiting toxicity (DLT) as the primary end point. Many adult, and almost all pediatric, phase I oncology trials use a modified version of the up-and-down method created in 1948 by Dixon and Mood.³ In the traditional 3 + 3, phase I cancer trial design, a minimum of three participants are studied at each dose level. If none of these three participants experience a DLT, a subsequent three participants are enrolled onto the next highest dose level. If one of three participants at a dose level experiences a DLT, up to three more participants are enrolled. When a DLT is observed in at least two participants in a cohort of three to six, the MTD is exceeded and an additional three participants (up to a total of six) are treated at the next

From the Division of Clinical Pharmacology and Therapeutics; and Division of Oncology, the Children's Hospital of Philadelphia, Philadelphia, PA.

Submitted May 23, 2007; accepted September 27, 2007.

Authors' disclosures of potential conflicts of interest and author contributions are found at the end of this article.

Corresponding author: Jeffrey M. Skolnik, MD, the Children's Hospital of Philadelphia, Abramson Research Center 916, 3615 Civic Center Blvd, Philadelphia, PA 19104-4318; e-mail: skolnik@email.chop.edu.

© 2008 by American Society of Clinical Oncology

0732-183X/08/2602-190/\$20.00

DOI: 10.1200/JCO.2007.12.7712

Table 1. Comparison of Decision Properties for the 3 + 3 v Rolling Six Design

No. Enrolled	DLT Data			Enrolling Dose Level*			
	No. DLTs	No. Without DLT	No. With Data Pending	MTD Not Exceeded		MTD Exceeded	
				3 + 3	Rolling Six	3 + 3	Rolling Six
2	0, 1	Any	Any	n	n		
2	2	0	0	n - 1	n - 1		
3	0	0, 1, 2	3, 2, 1	Suspend	n		
3	0	3	0	n + 1	n + 1		
3	1	0, 1	2, 1	Suspend	n		
3	1	2	0	n	n		
3	≥ 2	Any	Any	n - 1	n - 1		
4	0	0, 1, 2	4, 3, 2	—	n	—	n
4	0	3	1	—	n	n	n
4	0	4	0	—	n + 1	n	n
4	1	0, 1	3, 2	—	n	—	n
4	1	2	1	n	n	—	n
4	1	3	1	n	n	n	n
4	≥ 2	Any	Any	n - 1	n - 1	n - 1	n - 1
5	0	0, 1, 2	5, 4, 3	—	n	—	n
5	0	3, 4	2, 1	—	n	n	n
5	0	5	0	—	n + 1	n	n
5	1	0, 1	4, 3	—	n	—	n
5	1	2	2	n	n	—	n
5	1	3, 4	1, 0	n	n	n	n
5	≥ 2	Any	Any	n - 1	n - 1	n - 1	n - 1
6	0	0, 1, 2	6, 5, 4	—	Suspend	—	Suspend
6	0	3, 4	3, 2	—	Suspend	Suspend	Suspend
6	0	5, 6	1, 0	—	n + 1	MTD	MTD
6	1	0, 1	5, 4	—	Suspend	—	Suspend
6	1	2	3	Suspend	Suspend	—	Suspend
6	1	3, 4	2, 1	Suspend	Suspend	Suspend	Suspend
6	1	5	0	n + 1	n + 1	MTD	MTD
6	≥ 2	Any	Any	n - 1	n - 1	n - 1	n - 1

NOTE. This table does not take into account inevaluable patients.

Abbreviations: DLT, dose-limiting toxicity; MTD, maximum-tolerated dose.

*n is the current dose level of patients enrolled; n + 1 and n - 1 represent dose level escalation and de-escalation, respectively.

lower dose level. The MTD is defined as the dose level at which none or one of six participants (0% to 17%) experience a DLT, when at least two of three to six participants (33% to 67%) experience a DLT at the next highest dose. In the 3 + 3 design, accrual is suspended after enrollment of each cohort of three patients. When a participant becomes inevaluable for toxicity, most commonly because of early disease progression, the cohort is reopened to a single patient. Key factors that contribute to the overall timeline of a phase I study include delays associated with patient accrual, replacement of inevaluable patients, time to event (ie, DLT), and time associated with data submission and review.

One of the primary reasons for the original development of the 3 + 3 design was to limit the number of patients exposed to a potentially toxic or lethal dose of a new drug. Pediatric phase I trials are remarkably safe given the high-risk population being studied. In our prior review,⁴ we found that, among 1,066 pediatric patients reported from 47 studies, the toxic death rate was 0.5%, which was not different from that observed in adult phase I trials.⁵ Given the overall safety profile of pediatric phase I trials, the extended periods of time that studies are suspended to accrual, and the observation that the large majority of dose levels are ultimately expanded to accrue six patients,⁴ we propose a design modification for pediatric phase I studies, termed the rolling

six design, that may shorten the timeline of such trials. The rolling six design allows for accrual of two to six patients concurrently onto a dose level. Decisions as to which dose level to enroll a patient are based on the number of patients currently enrolled and evaluable, the number of patients experiencing DLTs, and the number of patients still at risk of developing a DLT at the time of new patient entry. We have compared the performance of the rolling six and 3 + 3 designs by using discrete event simulation.⁶⁻⁸ Our objectives were to define the characteristics of a pediatric phase I design that would reduce the overall time to study completion, to construct a simulation approach to compare the proposed design with the conventional 3 + 3 design, and to evaluate the performance of each design under typical clinical trial conditions.

METHODS

Historical Phase I Data

Retrospective data from a series of 14 completed, Phase I oncology trials conducted by the Children's Oncology Group (COG) Phase I and Pilot Consortium between 2000 and 2006 were used to examine study conduct and timelines. From this series, we selected a subset of completed trials that performed in accordance with the traditional 3 + 3 methodology. Primary end

Table 2. Twelve Completed Children's Oncology Group Studies Used for Historical Data

Study	Agent	No. of Patients		No. of DLTs	No. of Dose Levels	Study Duration (days)	Time Suspended to Accrual (days)	Average Time to Complete Level (days)
		Evaluable	Inevaluable					
ADVL0011	TMZ/CCNU	22	2	2	4	528	86	134
ADVL0015	Bortezomib (PS-341; Velcade)	15	3	2	2	281	158	95
ADVL0016	Gefitinib (ZD1839; Iressa)	21	4	2	4	477	347	89
ADVL0018	Hu14.18-IL2 fusion protein	28	1	3	7	563	430	59
ADVL0211	G3139 (Genasense)/Dox/CPM	29	5	4	5	606	378	107
ADVL0212	Depsipeptide	24	7	4	4	539	284	135
ADVL0214	Erlotinib (OSI-774; Tarceva)	22	3	3	5	344	188	78
ADVL0215	Decitabine/Dox/CPM	11	2	2	2	220	147	94
ADVL0311	Pemetrexed (LY231514; Alimta)	33	2	3	8	596	200	61
ADVL0314	Bevacizumab (Avastin)	14	2	0	3	233	87	132
ADVL0316	17-AAG	17	5	0	4	427	181	117
ADVL0415	Oxaliplatin/irinotecan	13	1	5	3	289	178	52
	Median	22	3	3	4	452	185	77
	Range	11-33	1-7	0-5	2-8	220-606	86-430	33-274

Abbreviations: DLT, dose-limiting toxicity; TMZ, temozolomide; CCNU, lomustine/nitrogen mustard; Dox, doxorubicin; CPM, cyclophosphamide; 17-AAG, 17-allylamino-17-demethoxygeldanamycin.

points for analysis included the time to study completion, the number of participants per study, the number of DLTs per study, and the MTD dose level. Secondary end points included the number of participants per dose level, the number of dose levels per study, and the time to completion of the dose level. Outcomes included the occurrence of DLTs, becoming inevaluable for toxicity (IE), or completing a cycle without a DLT or without becoming inevaluable (termed PASS); outcomes and the time of an event were summarized from individual patient event records. Additional information collected included administrative time, which represented time spent per trial not actively recruiting patients because of trial closure or suspension. Descriptive statistics were generated for the historical studies.

Simulating Patient Populations for Evaluation

For the simulation study, sets of patients were defined based on the historical phase I data. Nine potential cohorts of fifteen patients each (to allow for inevaluable patients) were simulated for each of 1,000 trials. From binary distributions, each patient was randomly assigned a probability of developing a DLT or of being IE. On the basis of historical data,⁴ the risk of developing DLT at dose levels 1 to 9 was set at 5%, 10%, 30%, 50%, 75%, 90%, 95%, 99%, and 99%, respectively. A dose level of -1, which had a DLT risk of 2%, was also

included for possible dose de-escalation at the starting dose level. Each patient was also assigned an interpatient arrival time (a chronologic time for a patient to arrive in the clinic and be eligible for study), an on-study start time, a time to DLT, a time to inevaluability, and a study evaluation period time without DLT or IE. The actual event of record was determined by the shortest time to an event (DLT, IE, or PASS). Arrival times were sampled from a Poisson distribution, whereas on-study start time and the time to inevaluability were sampled from normal distributions. The time to DLT was sampled from a uniform distribution that centered at 20 days. As the evaluation period (cycle length) is a study design parameter defined per protocol, it was fixed to a constant 21-, 28-, or 35-day value for each study. Study simulations were generated by using SAS 9.1.3 (SAS Institute Inc, Cary, NC), and each study run populated a separate SAS data set. A detailed description of the study simulation and design logic code and an examination of the sensitivity of the simulation model to distributional assumptions and simulation sample size considerations will be published separately.⁹

To evaluate the effect of changes in interpatient arrival time and the evaluation period (cycle length) on design comparison, populations of simulated patients were generated for interpatient arrival time distribution means

Table 3. Performance Metrics for 3 + 3 v Rolling Six Design Under Varied Conditions

Condition	3 + 3 Design								Rolling Six Design							
	Duration (days)		Patients Per Trial		DLT Per Trial		Cohorts Per Trial		Duration (days)		Patients Per Trial		DLT Per Trial		Cohorts Per Trial	
Interpatient arrival time, days*	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
5	256	69	17	4	3.2	1.2	3.1	0.7	197	53	20	5	3.5	1.3	3.3	0.8
10	350	84	17	4	3.2	1.1	3.2	0.7	294	75	20	5	3.3	1.1	3.3	0.8
20	538	128	16	4	3.1	1.0	3.1	0.7	486	121	19	5	3.2	1.0	3.2	1.0
100	2,178	519	16	4	3.1	1.1	3.2	0.7	2,112	502	19	5	3.1	1.0	3.3	0.8
Cycle length, day†																
28	383	94	16	4	3.1	1.1	3.1	0.7	315	82	20	5	3.3	1.2	3.3	0.8
35	432	109	17	4	3.2	1.2	3.2	0.7	348	89	20	5	3.5	1.3	3.3	0.8
Increased inevaluability rate of 25%	404	107	18	4	3.1	1.1	3.1	0.7	322	81	22	5	3.3	1.1	3.2	0.7
Increased risk of DLT at starting dose‡	242	59	12	3	3.2	1.0	2.1	0.3	185	39	13	2	3.3	1.0	2.1	0.4

Abbreviations: DLT, dose-limiting toxicity; SD, standard deviation.

*Cycle length = 21 days.

†Interpatient arrival time average = 10 days.

‡Probability of DLT in first cohort increased from 5% to 30%.

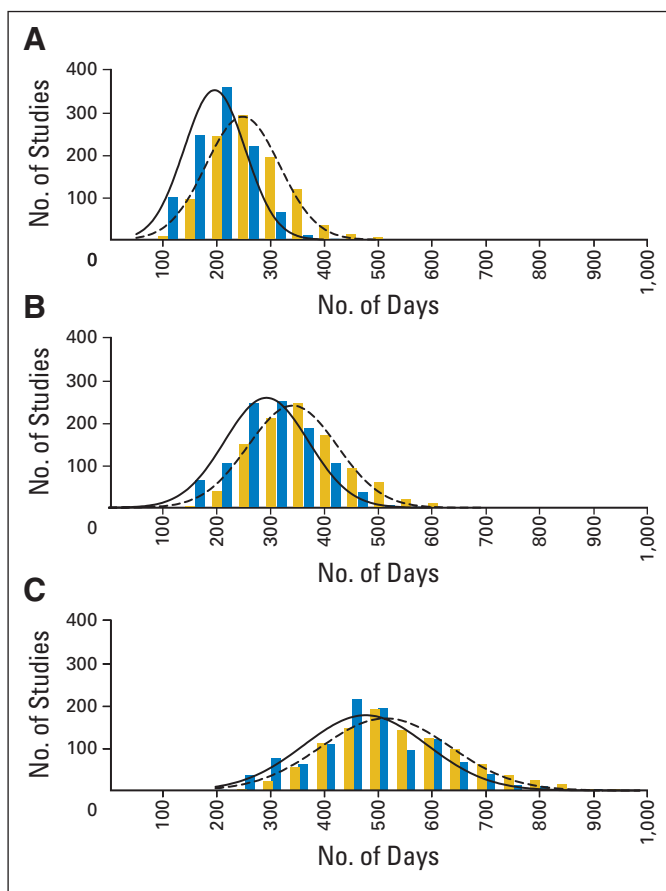


Fig 1. Distribution of elapsed time to completion for 1,000 simulated trials. Average interpatient arrival times of 5, 10, or 20 days in the A, B, and C panels, respectively. Data is plotted against a Gaussian fit. (— with blue bars) rolling six design; (--- with yellow bars) 3 + 3 design.

of 5, 10, 20, and 100 days and for evaluation periods of 21, 28, and 35 days. The effect of changing inevaluability rates on study timelines was studied by examining rates that ranged from 0% to 30%. The effect of increased risk of toxicity was evaluated by shifting the starting dose level to a risk of DLT of 30%. To verify that the simulated event probability (IE or DLT) conformed to the design logic, the number of events were summed across trials by using SAS/Proc Freq (SAS Institute Inc), and the distributions of the time to event variables (eg, time to DLT, interpatient arrival time) were verified using SAS/

Insight (SAS Institute Inc). The independence of event times and outcomes were also confirmed for each study simulation.

Comparing Study Design Logic

The performances of the 3 + 3 and rolling six designs were evaluated by simulating trials with each design that used the simulated study populations. Both designs were coded in SAS (SAS Institute Inc) with the specific dose level assignments and with event rates and times for each patient that were derived from the simulated datasets. Comparison of dose level progression properties of each design are listed in Table 1. The design logic for both approaches was based on conditional statements that check to see if a study was open to accrual. Patients were sequentially sampled from the simulated trials based on dose level assignment and were added to the evaluated trial population as follows: in the 3 + 3 design, three patients were enrolled and were fully assessed before another patient (up to three more) could enroll. To replicate real-life events, administrative time was added after the first group of three patients, during which no patients were enrolled. Dose level escalation or de-escalation proceeded through traditional up-and-down logic.

In the rolling six design, up to six patients were concurrently enrolled onto study. Accrual to the study was only suspended when awaiting data from six patients. Decisions as to whether to enroll a new participant onto the current, next highest, or next lowest dose level were made based on available data at the time of new participant enrollment. Dose level assignment was based on the number of participants currently enrolled in the cohort, the number of DLTs observed, and the number of participants at risk for developing a DLT (ie, participants enrolled but who were not yet evaluable for toxicity). For example, when three participants were enrolled onto a dose cohort, if toxicity data was available for all three when the fourth participant entered and there were no DLTs, the dose was escalated and the fourth participant was enrolled to the subsequent dose level. If data was not yet available for one or more of the first three participants, or if one DLT had been observed, the new participant was entered at the same dose level. Lastly, if two or more DLTs had been observed, the dose level was de-escalated. The process was repeated for participants five and six. In place of suspending accrual after every three participants, accrual was only suspended when a cohort of six was filled. When participants were inevaluable for toxicity, they were replaced with the next available participant if escalation or de-escalation rules had not been fulfilled at the time the next available participant enrolled onto the study. Because participants were assigned an interpatient arrival time, the selection of participants from simulated participant cohorts was not linear but rather was chronologic; if the study was not accruing new participants at the time assigned to an individual, that individual was skipped and the next available participant was assessed for arrival time, in accordance with real-life study conduct. The same patients were used for both the rolling six and the 3 + 3 designs in each pair of 1,000 simulated trials (intrasimulation), and separate groups of patients were used for each new simulation (intersimulation).

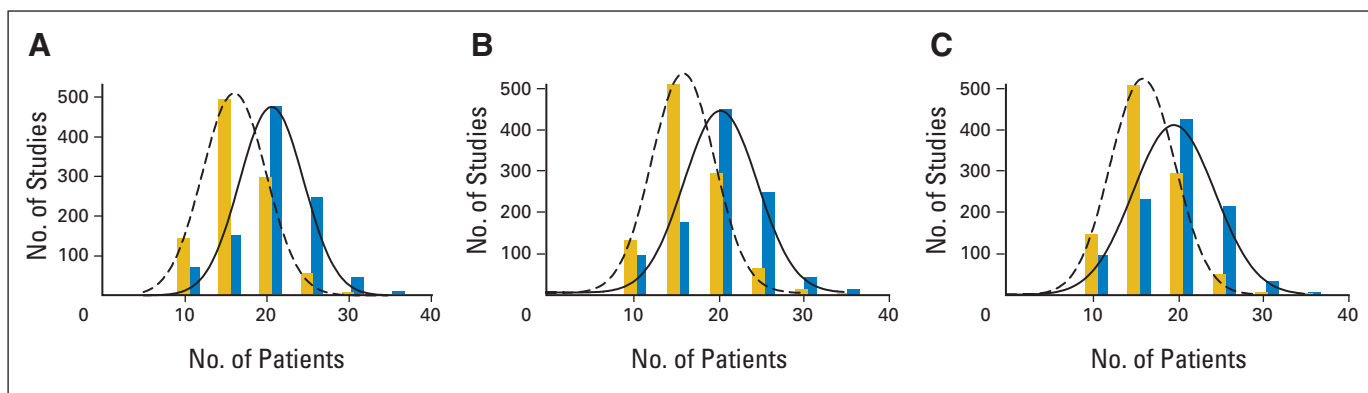


Fig 2. Distribution of number of patients for 1,000 simulated trials. Average interpatient arrival times of (A) 5, (B) 10, or (C) 20 days. Data are plotted against a Gaussian fit. (— with blue bars) rolling six design; (--- with yellow bars) 3 + 3 design.

RESULTS

Historical Phase I Database

Of the 14 COG studies completed, 12 were found to provide acceptable study characteristics and timeline data for analysis (Table 2). In two of these 12 studies, analysis was limited to cohorts that followed the traditional 3 + 3 methodology. The total number of patients in all studies was 249. The median number of patients per study was 22 (range, 11 to 33), dose levels evaluated were four (range, 2 to 8), and the dose level at which the MTD was defined was three (range, 1 to 7). Of the 12% of patients (30 of 249) who experienced a DLT, the median time to DLT was 14 days (range, 0 to 40 days). The median time to dose level and study completion was 77 days (range, 33 to 274 days) and 452 days (range, 220 to 606 days), respectively. Of note, studies were suspended to accrual at a median of 185 days (range, 86 to 430 days), which represented more than 50% of total study time.

Simulated Studies: Performance of 3 + 3 Versus Rolling Six Design

The rolling six design outperformed the 3 + 3 design for key performance metrics (Table 3). The distribution of elapsed time to complete 1,000 simulated trials, with mean interpatient arrival times of 5, 10, or 20 days, uniformly favored the rolling six design (Fig 1). For an average interpatient arrival time of 10 days—the most commonly observed interpatient arrival time for studies—the mean \pm standard deviation (SD) time to complete a pediatric phase I study was 294 \pm 75 days for the rolling six design versus 350 \pm 84 days for the 3 + 3 design. With the rolling six design, 13%, 52%, 29%, and 2.5% of studies reached the MTD at dose levels 1, 2, 3, or 4, respectively; with the 3 + 3 design, 12%, 52%, 30%, and 3% of studies reached the MTD at those dose levels, respectively. On average, the rolling six design enrolled three more patients per study than the 3 + 3 design (Fig 2) with an overall mean \pm SD patient accrual of 20 \pm 5 and 17 \pm 4 for the rolling six and 3 + 3 designs, respectively. Importantly, there was no difference in the distribution of DLTs between designs (Fig 3); a mean \pm SD of 3.3 \pm 1.1 participants developed a DLT in the rolling six design, and 3.2 \pm 1.1 participants developed a DLT in the 3 + 3 design.

Changes in Simulation Parameters

Changes in study parameters, including changes in the interpatient arrival time, the probability of a DLT, and the cycle length, did not change the performance of the rolling six design (Table 3).

Changes in the evaluation period from 21 to 28 or 35 days did not change the outcome of simulations; in all instances, the rolling six design outperformed the 3 + 3 design with respect to total study length. When interpatient arrival time was increased to a mean of 100 days, the rolling six design slightly outperformed the 3 + 3 design. The impact of increasing the rate of inevaluability was greater with the 3 + 3 design than with the rolling six design; when the rate increased to 30%, the mean \pm SD time to study completion with the 3 + 3 design was 98 \pm 85 days longer than with the rolling six design.

When the probability of a DLT at the starting dose level increased from 5% to 30%, the rolling six design was favored in terms of elapsed time to complete the trial (mean \pm SD, 185 \pm 39 v 242 \pm 59 days); there was no difference in the incidence of DLT, but, as expected, fewer patients were required per study (mean \pm SD, 13 \pm 2 v 12 \pm 3).

DISCUSSION

By using discrete event simulation, we found that the rolling six method has the potential to significantly decrease the duration of pediatric phase I studies. There was a small increase in the total number of patients enrolled when using the rolling six design ($n = 3$), but such an increase is readily acceptable because an experience of 18 to 24 pediatric patients is typical before proceeding to phase II trials. Importantly, there was no clinically significant increase in the incidence of DLTs when using the new design.

The primary reason the rolling six design shortens the overall duration of study conduct is that the number of times a study is suspended to accrual is significantly decreased compared with the 3 + 3 method. The rolling six design also decreases the likelihood that a patient who is eligible to enroll onto a pediatric phase I study is unable to do so because of study suspension to accrual.

The rolling six design performance is not sensitive to key metrics, such as increases in risk of DLTs or changes in evaluation period. At the lowest extremes of interpatient arrival times, the rolling six continued to outperform the 3 + 3 design; when interpatient arrival time was zero days, the rolling six study was on average 26 days shorter than the corresponding 3 + 3 study (data not shown). In studies with a protracted interval between patient enrollment, which occurs infrequently in pediatric phase I solid tumor studies, the rolling six performs similarly to the 3 + 3 design (Table 3).

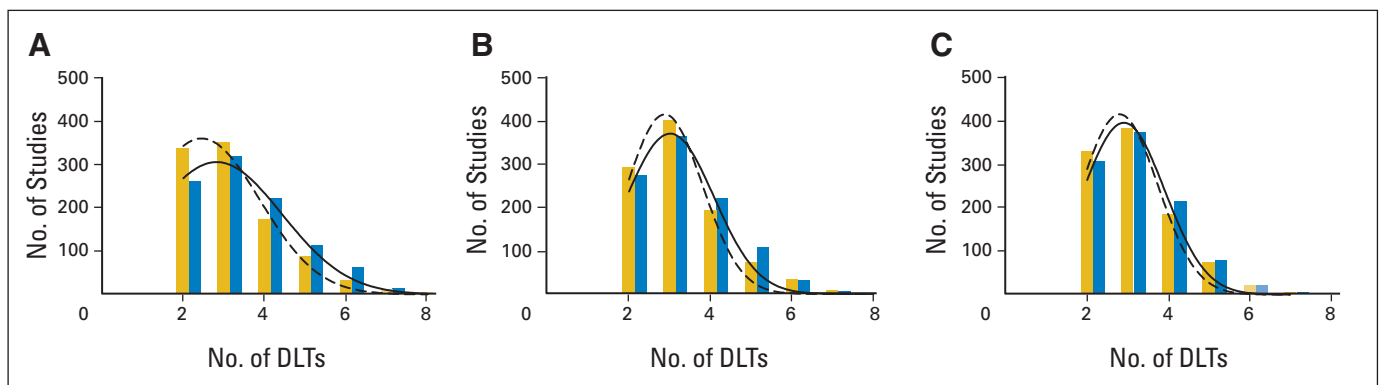


Fig 3. Distribution of number of dose-limiting toxicities (DLTs) in 1,000 simulated trials. Average interpatient arrival times of (A) 5, (B) 10, or (C) 20 days. Data are plotted against a Gaussian fit. (— with blue bars) rolling six design; (--- with yellow bars) 3 + 3 design.

In adult phase I trials, the most commonly used alternative method to the traditional 3 + 3 design is the continual reassessment method (CRM)¹⁰ or its modifications,^{11,12} which attempt to minimize the number of patients enrolled below a biologically active dose without increasing the number of DLTs. As pediatric trials are conducted only after completion of adult phase I trials, the potential dose range studied is far narrower than in adult trials,⁴ which minimizes the potential benefit of using a CRM approach to define the pediatric MTD. A fundamental difference between the rolling six and CRM methods is that the former is geared towards shortening the duration of a study and not towards refining the estimate of the MTD.

Our study utilized real-world data to generate simulated populations of patients. Unlike past studies in adults, which did not include actual study conduct data,^{13,14} we derived baseline population metrics (eg, interpatient arrival time, time to an event, risk of a DLT, risk of IE) from both a large set of published Phase I pediatric oncology studies,⁴ and a newer set of 12 recently completed, pediatric phase I studies. This qualifies our approach and allows a real-time experience with the simulations that few phase I simulations have shared in the past.

With an increasing number of new agents entering the clinical pipeline, the creation of a study design that is less sensitive to changes in metrics, such as in interpatient arrival time, probability of DLT, and cycle length, could have a significant impact on pediatric cancer drug

development. Future phase II studies are likely to center on randomized phase II selection designs¹⁵⁻¹⁷; having several new phase I agents ready for phase II evaluation in shorter time periods will facilitate the design and conduct of these studies. We plan to prospectively evaluate the rolling six design in the upcoming generation of trials conducted by the COG Phase I and Pilot Consortium.

AUTHORS' DISCLOSURES OF POTENTIAL CONFLICTS OF INTEREST

The authors indicated no potential conflicts of interest.

AUTHOR CONTRIBUTIONS

Conception and design: Jeffrey M. Skolnik, Jeffrey S. Barrett, Peter C. Adamson

Collection and assembly of data: Jeffrey M. Skolnik, Jeffrey S. Barrett, Bhuvana Jayaraman, Dimple Patel

Data analysis and interpretation: Jeffrey M. Skolnik, Jeffrey S. Barrett, Bhuvana Jayaraman, Dimple Patel, Peter C. Adamson

Manuscript writing: Jeffrey M. Skolnik, Jeffrey S. Barrett, Peter C. Adamson

Final approval of manuscript: Jeffrey M. Skolnik, Jeffrey S. Barrett, Peter C. Adamson

REFERENCES

1. Adamson PC, Weiner SL, Simone JV, et al: Making Better Drugs for Children with Cancer. Washington, DC, The National Academies Press, 2005
2. Smith M, Bernstein M, Bleyer WA, et al: Conduct of phase I trials in children with cancer. *J Clin Oncol* 16:966-978, 1998
3. Reiner E, Paoletti X, O'Quigley J: Operating characteristics of the standard phase I clinical trial design. *Comput Stat Data Anal* 30:303-315, 1999
4. Lee DP, Skolnik JM, Adamson PC: Pediatric phase I trials in oncology: An analysis of study conduct efficiency. *J Clin Oncol* 23:8431-8441, 2005
5. Ross L: Phase I research and the meaning of direct benefit. *J Pediatr* 149:S20-S24, 2006
6. Connelly LG, Bair AE: Discrete event simulation of emergency department activity: A platform for system-level operations research. *Acad Emerg Med* 11:1177-1185, 2004
7. Le Lay A, Despiegel N, Francois C, et al: Can discrete event simulation be of use in modelling major depression? *Cost Eff Resour Alloc* 4:19, 2006
8. Shechter SM, Bryce CL, Alagoz O, et al: A clinically based discrete-event simulation of end-stage liver disease and the organ allocation process. *Med Decis Making* 25:199-209, 2005
9. Barrett JS, Jayaraman B, Patel D, et al: Discrete Event Simulation (DES) as a Technique to Study Decision Rule Efficiency in Event-Driven Clinical Designs. *American College of Clinical Pharmacology* 46:1092, 2006
10. O'Quigley J, Pepe M, Fisher L: Continual reassessment method: A practical design for Phase I clinical trials in cancer. *Biometrics* 46:33-48, 1990
11. Goodman SN, Zahurak ML, Piantadosi S: Some practical improvements in the continual reassessment method for phase I studies. *Stat Med* 14:1149-1161, 1995
12. 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* 14:911-922; discussion 923, 1995
13. 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
14. Loke YC, Tan SB, Cai Y, et al: A Bayesian dose finding design for dual endpoint phase I trials. *Stat Med* 25:3-22, 2006
15. Rosner GL, Stadler W, Ratain MJ: Randomized discontinuation design: Application to cytostatic antineoplastic agents. *J Clin Oncol* 20:4478-4484, 2002
16. Morita S, Sakamoto J: Application of an adaptive design to a randomized phase II selection trial in gastric cancer: A report of the study design. *Pharm Stat* 5:109-118, 2006
17. Estey EH, Thall PF: New designs for phase 2 clinical trials. *Blood* 102:442-448, 2003

Acknowledgment

We thank Elizabeth O'Connor of the Children's Oncology Group (Arcadia, CA) for her assistance with data collection for the historical phase I data.