available at www.sciencedirect.com

ScienceDirect

journal homepage: www.ejconline.com

Operating characteristics of two independent sample design in phase I trials in paediatric oncology ☆

Mathilde Raphaël [a], *Marie-Cécile le Deley* [a,c], *Gilles Vassal* [a,c], *Xavier Paoletti* [b,*]

[a] *Institut Gustave Roussy, 39 rue Camille Desmoulins, 94805 Villejuif, France*
[b] *Institut Curie, INSERM U900, 26 rue d'Ulm, 75005 Paris, France*
[c] *University Paris-Sud, Le Kremlin-Bicêtre, France*

ARTICLE INFO

Article history:
Received 10 April 2009
Received in revised form 25 September 2009
Accepted 15 January 2010
Available online 20 February 2010

Keywords:
Dose finding
Heterogeneity
Continual reassessment method
Evaluation

ABSTRACT

Purpose: The European medicines agency (EMEA) has stated that the degree of pre-treatment could modify the patient's tolerance to new treatments in paediatric oncology. It is current practice to divide a phase I trial into two groups to identify the maximum tolerated dose (MTD) in each group separately. The aim of this study was to investigate the relevance of this approach.

Methods: We reanalysed a large phase I trial of Irinotecan that included 80 children (32 heavily pretreated patients and 48 less heavily pretreated). An extended simulation study was performed to investigate the robustness of the conclusions in the context of small sample sizes. Dose recommendations were studied according to scenarios with group differences, as measured by odds ratio (OR), ranging from 1 (no difference) to 10 (large difference) and sample sizes increasing from 20×2 to 60×2 patients.

Results: This study shows a high risk of misidentification of the MTD in each of the two groups, regardless of the group difference. With a group difference corresponding to OR = 8 and balanced sample sizes (20×2 patients), the same MTD was identified in 11% of the simulations. Even with larger sample sizes (40×2 patients), this figure reached 24% for OR = 3. There is also a very high risk of identifying two different MTD (52% for 40×2 patients) although the risk is similar in both groups.

Conclusions: Two independent sample designs in paediatric phase I trials should be avoided or reserved to limited situations when there is a strong rationale possibly based on adult data.

1. Introduction

The purpose of a phase I clinical trial is to identify the optimal dose to be recommended for further studies on a limited sample of patients. For most compounds, this is generally taken as the maximum tolerated dose (MTD), usually defined as a percentile of some unknown increasing dose–toxicity curve. This dose will be used for further investigations.[1] The European medicines agency (EMEA)[2] has stated that the degree of pre-treatment could modify the patient's tolerance to new treatments in paediatric oncology, and then suggested that exploration of this potential difference should be considered. Assuming that heavily pretreated patients have a higher risk of severe toxicity than less heavily pretreated patients, it is often expected that the MTD should be lower for heavily pretreated patients than for less heavily pretreated patients.[3]

☆ This work was partly presented during the 2007 ASCO meeting.

* *Corresponding author*: Tel.: +33 1 56 24 56 47; fax: +33 1 53 10 40 20.
E-mail addresses: raphael.mathilde@gmail.com (M. Raphaël), marie-cecile.ledeley@igr.fr (M.-C. le Deley), gilles.vassal@igr.fr (G. Vassal), xavier.paoletti@curie.net (X. Paoletti).

0959-8049/$ - see front matter © 2010 Elsevier Ltd. All rights reserved.
doi:10.1016/j.ejca.2010.01.024

Phase I trials are therefore commonly designed to investigate several groups of patients with different histories of previous chemotherapies to determine different MTDs, requiring larger sample sizes for an unknown benefit.[4–6] This two-sample design is not systematically reported in publications either because of lack of accrual in one group or because the results are published separately.

Is it relevant to consider two separate groups based on the degree of pre-treatment in phase I clinical trials with inevitably limited sample sizes? Whilst being standard practice, this approach has never been evaluated.

To address this issue, the results of a large phase I paediatrics clinical trial of IRINOTECAN were reviewed.[4] Relationships between degree of pre-treatment, dose and patient tolerance were described on the basis of these data. An extended simulation study was then performed to investigate the impact of a two-sample design on correct identification of the MTD under various scenarios. If the degree of pre-treatment is truly a prognostic factor of the risk of severe toxicity, how likely is it to be detected? Conversely, if the two groups are homogeneous, what is the risk of identifying two distinct MTD?

2. Materials and methods

2.1. Materials: the Irinotecan trial

The Irinotecan trial was a phase I paediatric trial to determine the maximum tolerated dose (MTD) of IRINOTECAN administered as a single intravenous infusion.[4] Eighty patients with malignant solid tumours were enrolled in this study. Two groups were considered separately including 'less heavily pretreated' ($n = 48$, group L) and 'heavily pretreated' ($n = 32$, group H) patients, as defined by prior craniospinal irradiation and/or high-dose chemotherapy with stem cell rescue. Patients were required to have adequate blood cell counts (i.e. neutrophils > 1500 mm^{-3} [group L], or >1000 mm^{-3} [group 2]; and platelets > 100,000 mm^{-3} [group L], or >75,000 mm^{-3} [group H]) except in the case of bone marrow involvement, satisfactory liver function, normal renal function and no radiotherapy or chemotherapy within the last 4 weeks before study entry (or 6 weeks if NU).

Toxic side-effects were assessed after each cycle of treatment but only dose-limiting toxicities (DLTs) after the first cycle were used to identify the MTD. Dose finding was performed in both groups. In group L, the standard 3 + 3 method was used and in group H, the continual reassessment method (CRM)[7] was used. As shown in Table 1, children received IRINOTECAN at doses ranging from 200 to 720 mg/m^2 and the main DLTs experienced were delayed diarrhoea in less heavily pretreated patients, and neutropenia in heavily pretreated patients. The maximum tolerated dose (MTD) was 600 mg/m^2 in both groups. This level was expanded to further characterise the toxicity profile and was recommended for further studies.

2.2. Methods

To investigate the effect of the degree of pre-treatment on the dose–toxicity relation of Irinotecan, a logistic model was fitted to the Irinotecan data with a covariate g for the group that takes value $g = 0$ for less heavily pretreated patients (group L) and $g = 1$ for heavily pretreated children (group H). This provides the adjusted odds ratio estimate for the group effect.

To investigate the prognosis of baseline characteristics on the risk of DLT, we carried out only bivariate analyses adjusted on the dose due to the limited number of events. We estimated the odds ratio of the degree of pre-treatment as well as of the different baseline variables after adjustment on the dose level. The 95% confidence interval (95% CI) is provided together with each estimate. Significant baseline characteristics were further adjusted on the degree of pre-treatment.

To investigate the relevance of considering two parallel groups of patients and the reproducibility of our results in a more general framework, a simulation study was conducted. The risk of erroneous final recommendations was assessed with the following parameters:

(i) The risk of not recommending the correct dose level in at least one group (risk of wrong identification).
(ii) The risk of identifying different MTDs in the two groups when there is, in fact, no different risk of toxicity.

Table 1 – Dose levels and DLT outcomes in the less heavily pretreated and heavily pretreated patients of the CPT11 trial.

Dose		Group L, less heavily pretreated ($n = 48$)		Group H, heavily pretreated ($n = 32$)	
Level	mg/m^2	Nb of pts	DLTs	Nb of pts	DLTs
Level 1	200	3	–	3	–
Level 2	240	3	–	4	–
Level 3	300	3	–	3	–
Level 4	350	6	–	7	–
Level 5	420	5	–	3	–
Level 6	500	6	1 heart failure Gr-3	3	1 Gr-3 nausea/vomiting
Level 7	600	16	1 neutropenia/Gr-3 infection 1 Gr-3 nausea/vomiting	10	2 Gr-4 persistent neutropenia
Level 8	720	6	2 delayed diarrohea (Gr-3 and Gr-4) 1 Gr-3 cholinergic syndrome	–	

Gr-3: grade 3 and Gr-4: grade 4.

(iii) The risk of identifying the same MTD in the absence of group difference, or even of identifying a higher MTD in group H than in group L.

Simulations consist in several steps: first assume a 'true' dose–toxicity relation (a scenario), i.e. the probability of toxicity at the different dose levels, next simulate the process of a trial, that is (i) the patient allocation to a dose, (ii) the simulation of the outcome (DLT or not) based on the 'true' probability of toxicity at the allocated dose level and (iii) the recommendation of the next dose level to be visited, last repeat the trial simulation a large number of times to obtain average operating characteristics expressed as the distribution of the final recommendations, or the dose allocation, the number of DLTs, etc.

Several scenarios were explored. The true group-specific dose–toxicity relation for IRINOTECAN in children was assumed to follow the multivariate logistic model described above with the dose taken as continuous and the parameters estimated from Irinotecan data. The MTD was the dose level for which the probability of toxicity was closest to 20%. In the group L, the MTD was then dose level 7 and the 'true' dose–toxicity relation was the same in all scenarios The true dose–toxicity model in the other group ranged from OR = 1 ($\beta = 0$) corresponding to no group effect (same risk of severe toxicity in the two groups), to OR = 10 ($\beta = 2.3$) corresponding to a major excess risk in group H compared to group L. Scenarios are illustrated in Fig. 1. In other words, in group H, the MTD was also dose level 7 for the scenario OR = 1 (same risk of severe toxicity in both groups). It was dose level 6 for a moderate excess risk compared to group L (OR = 3 or 4) and dose level 5 for a major excess risk. The impact of the sample size on the final recommendations was also explored. Trials enrolling between 20×2 and 60×2 patients were simulated for an OR of either 1 or 8. Only balanced sample sizes were studied. CRM was used to conduct dose escalation independently in groups L and H. Doses were escalated after each simulated patient tolerated treatment until the first DLT was observed. A logistic model with fixed slope was then fitted after each new simulated patient. The dose allocated to the following patient had a probability of toxicity closest to 20%. No dose skipping was allowed. The simulated trials came to a halt after the fixed number of patients. For each scenario, 1000 simulated trials were repeated, resulting in a distribution of final recommendations, i.e. the frequency of simulations that recommended each dose level as the MTD. The precision of the simulations, that is of the estimate of the probability of identifying a dose level closest to 20% of DLT, was then between 1% and 1.5% depending on how frequently the dose level was recommended.

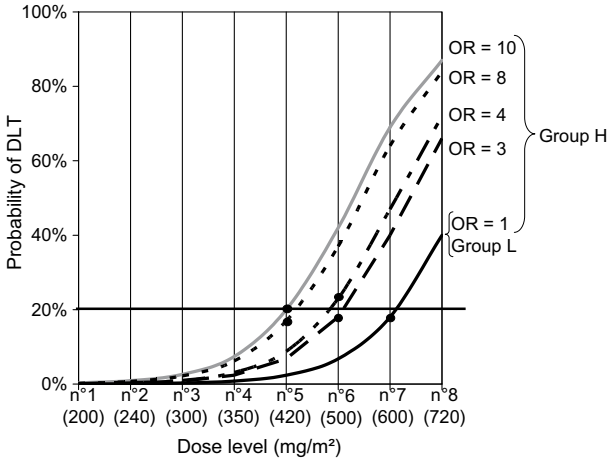

Fig. 1 – Dose–toxicity curve of the reference group L (less heavily pretreated patients) drawn from the Irinotecan data and dose–toxicity curves in group H (heavily pretreated) obtained under the assumption of increasing group difference (OR = 1–10).

3. Results

3.1. *Irinotecan trial*

Patient characteristics on inclusion in the Irinotecan trial according to the degree of pre-treatment are described in Table S1 of the Supplementary material. No statistically significant difference was observed between groups L and H in terms of clinical data at entry, while platelet count only appeared to be better in the less heavily pretreated group. No statistically significant association was found between the degree of pre-treatment and the overall risk of toxicity after one course (OR = 2.2, 95% confidence interval 0.4–12.6). This was in agreement with the conclusions of the trial that had recommended the same MTD in both groups. No other covariates (clinical, haematological or biochemical), apart from age (and therefore BMI) were statistically associated with the risk of DLT after adjustment on the dose. Nevertheless a lack of power cannot be ruled out.

3.2. *Simulation study*

Table 2 describes the operating characteristics in the case of no group difference (OR = 1) in terms of risk of severe toxicity. Strikingly, a very high frequency of simulations that did not recommend the correct doses was observed. For example, for a sample size of 30 patients per group ($n = 60$), the probability of recommending the right MTD (dose level 7) was 57% in each group and then the probability of correct identification of both MTD in the two independent groups was 32% (68% led to a wrong identification in at least one of the two groups). The risk of identification of different MTDs, when the groups did not differ, was greater than 50% when the trial recruited 40×2 patients or less. When increasing the sample size, this risk decreased up to 37% for 60×2 patients; this risk remained quite high for an untypical sample in phase I.

As illustrated in Fig. 2, the risk of incorrect identification in at least one group decreased with the sample size but was independent of the OR for a given sample size. In the presence of a strong group effect (OR = 8), this risk was still as high as 65% after 30×2 patients.

As described in Section 2.2, an odds ratio equal to 8 resulted in a two-dose level difference in terms of MTDs (dose level 7 in group L and dose level 5 in group H). With such a large difference between groups, the probability of identifying

Table 2 – Final recommendations for different sample sizes in the scenario of an odds ratio of 1.

	Scenarios									
	$N = 20 \times 2$		$N = 30 \times 2$		$N = 40 \times 2$		$N = 50 \times 2$		$N = 60 \times 2$	
Dose level	L	H	L	H	L	H	L	H	L	H
Frequency of recommendation in each group										
Level 1	–	–	–	–	–	–	–	–	–	–
Level 2	–	–	–	–	–	–	–	–	–	–
Level 3	–	–	–	–	–	–	–	–	–	–
Level 4	–	1%	–	–	–	–	–	–	–	–
Level 5	2%	2%	1%	1%	–	–	–	–	–	–
Level 6	24%	24%	21%	21%	18%	18%	15%	14%	10%	10%
Level 7	46%	45%	57%	57%	65%	65%	72%	73%	79%	79%
Level 8	28%	28%	21%	21%	17%	17%	13%	14%	11%	12%
Group ordering	*Distribution of joint recommendation*									
$MTD_{Li} > MTD_{Hi}$	34%		29%		26%		21%		19%	
$MTD_{Li} = MTD_{Hi}$	32%		41%		48%		56%		63%	
$MTD_{Li} < MTD_{Hi}$	34%		30%		26%		23%		18%	
Risk of identifying different MTDs in the two groups when there is no difference in the risk of toxicity										
	68%		59%		52%		44%		37%	

For each scenario and each group, highlighted percentages represent the probability of correctly identifying the MTD.

the same MTD was 9% after two samples of 20 patients; a sample size of at least 40×2 was required to maintain this risk below 3%. There was even a 2% risk of identifying a higher MTD in group L than in group H (cf. Table 3).

Table 4 highlights the impact of group differences, considering a sample size of 40×2 patients. A 24% probability of wrongly identifying the same MTD or even a higher MTD in group L than in group H was observed for an odds ratio equal to 3. This risk was less than 5% only in the situation of a marked excess risk (OR $\geqslant$ 6, results for OR = 6 not shown).

4. Discussion

There has been a growing interest in the design of paediatric phase I clinical trials either to improve the accuracy of MTD identification, or to speed up the trial with the rolling six design,[8] or to more effectively take into account the heterogeneity of the population. Regarding this last concern, our analysis of the Irinotecan data failed to clearly associate clinical or laboratory characteristics with the degree of pre-treatment. This result can be partly explained by the very strict inclusion criteria in the Irinotecan study that tended to reduce possible differences between the two groups. Moreover, no correlation was observed between toxicity and degree of pre-treatment. However, the observed toxicity profile was mainly digestive and the impact of pre-treatment, if any, is probably more marked on haematologic toxicity than on digestive toxicity. The effect of pre-treatment on bone marrow toxicity could have been masked by the development of other toxicities and the association between pre-treatment and risk of DLT may have been diluted. This negative result may also be due to the random variation that has a strong impact on inference due to the limited sample sizes of phase I clinical trials.

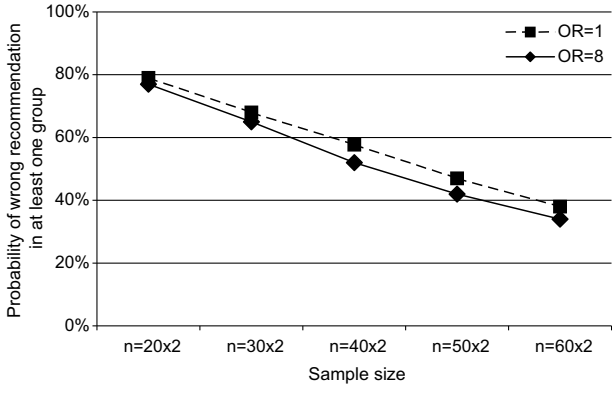

Fig. 2 – Risk of incorrect identification of the MTD in at least one group for an odds ratio of 1 and 8 in terms of sample sizes.

Results of the simulations show that the conclusions of the Irinotecan trial should be interpreted with caution even though the sample size was exceptionally large. Considering the large risk of error after 80 patients, we cannot exclude that heavily pretreated children were indeed at greater risk of toxicity than less heavily pretreated children. Small samples are a major limitation to conclusions of phase I clinical trials. Splitting the population into two samples multiplies the risk of incorrect identification of MTD in at least one of the two groups, even for large sample sizes. The risk of error is a general but neglected feature of phase I clinical trials. Single arm trials have limited precision: over a large range of situations, it has been shown that about 50% of the simulations identify a dose that does not correspond to the MTD.[9,10] This success rate can get even worse for flat dose–toxicity relations,[11] which are not uncommon with targeted agents. For two-arm trials, in addition to incorrect MTD identification, there is a significant risk of concluding on either a group dif-

Table 3 – Final recommendations for different sample sizes in the scenario of an odds ratio of 8.

	Scenarios									
	$N = 20 \times 2$		$N = 30 \times 2$		$N = 40 \times 2$		$N = 50 \times 2$		$N = 60 \times 2$	
Dose level	L	H	L	H	L	H	L	H	L	H
Frequency of recommendation for each group separately										
Level 1	–	–	–	–	–	–	–	–	–	–
Level 2	–	–	–	–	–	–	–	–	–	–
Level 3	–	2%	–	–	–	–	–	–	–	
Level 4	–	19%	–	15%	–	11%	–	7%	–	5%
Level 5	3%	48%	1%	64%	–	74%	–	80%	–	83%
Level 6	24%	28%	27%	20%	18%	15%	15%	13%	12%	12%
Level 7	47%	3%	54%	1%	65%	–	73%	–	80%	–
Level 8	26%	–	19%	–	17%	–	12%	–	9%	–
Group ordering	*Distribution of joint recommendation*									
$MTD_{Li} > MTD_{Hi}$	89%		93%		98%		98%		99%	
$MTD_{Li} = MTD_{Hi}$	9%		6%		2%		2%		1%	
$MTD_{Li} < MTD_{Hi}$	2%		1%		–		–		–	

For each scenario and each group, highlighted percentages represent the probability of correctly identifying the maximum tolerated dose.

ference when there is no such difference or on the absence of difference when a difference exists. Doubling the sample size with two groups appears to be unproductive and an escalation ignoring the implied biological constraints may lead to logically inconsistent dosing recommendations.

To limit these large errors of joint recommendations, several authors[12–14] have recently proposed experimental designs that incorporate prior assumption regarding the expected effect of the prognostic factor. The extreme case is to consider that a group cannot be inferior to the other by construction. Nevertheless, simulations of the operating characteristics of such methods have shown that only very little is gained unless strong and unverifiable prior information on the group difference is elicited. Such prior assumptions increase the risk of biased final recommendations in the case of misspecification.

The rationale for running multi-arm dose-finding studies designed to investigate population heterogeneity can also be questioned. Lee and colleagues[15] underlined that in the past 15 years, dose-intensive therapies have been routinely administered in high-risk tumours and patients included in early clinical trials were generally heavily pretreated. To our knowledge, no studies on large populations, such as those eligible for phase I trials, have shown that the degree of pretreatment is associated with the risk of DLT. Likewise, we are not aware of paediatric treatments that are registered with different doses according to degree of previous treatments. Conversely, the main restrictions to dosages are

Table 4 – Final recommendations for different odds ratio and 40 patients per group

	Scenarios: OR				
	Reference group	OR = 3	OR = 4	OR = 8	OR = 10
Dose level	L	H	H	H	H
Frequency of recommendation for each group separately					
Level 1	–	–	–	–	–
Level 2	–	–	–	–	–
Level 3	–	–	–	–	0.1%
Level 4	–	–	–	11%	18%
Level 5	–	14%	30%	74%	75%
Level 6	18%	76%	68%	15%	7%
Level 7	65%	10%	2%	–	–
Level 8	17%	–	–	–	–
Group ordering		Distribution of joint recommendation			
$MTD_{Li} > MTD_{Hi}$	–	76%	86%	98%	99%
$MTD_{Li} = MTD_{Hi}$	–	23%	14%	2%	1%
$MTD_{Li} < MTD_{Hi}$	–	1%	–	–	–

For each scenario and each group, highlighted percentages represent the probability of correctly identifying the maximum tolerated dose.

related to organ dysfunctions, especially for some toxicity profile such as cardiac toxic side events with anthracyclines or renal toxicity with platinum-based regimens. Very few factors have been related to the risk of developing toxic side events. On a pooled analysis of 154 patients included in phase I, Bachelot and colleagues[16] only found the age as prognostic factor after adjustment on the dose level. Organ function disorders have much stronger influence on the tolerability of the treatment than other known variables and could be good candidates supposing that there is a need to define the MTD in this specific subpopulation. The influence of previous lines of treatments is probably strongly reduced when controlling for good organ functions. This raises the delicate issue of inclusion criteria as to whether strict selection should or should not be performed. Phase I trials usually enrol a highly selected population and this likely has a major impact on dose intensity achieved and treatment activity observed.[17] Usually, phase I patients are not representative of the whole phase I population since they have good performance status, good medullar, renal or hepatic functions. This limits the risk of severe and irreversible adverse events and enables to better characterise the toxicity profile of the treatment under study. This population is also felt to be more representative of the general less advanced population. However, these drastic inclusion criteria homogenise the population making it questionable to target different MTDs. The two alternatives should be either to investigate the MTD variation in a largely defined population and then to use a 2-sample design or to restrict the inclusion criteria and to use a 1-sample design. This latter solution appears much more practical and efficient.

Although there is physiologic and biologic evidence for this association, the number of previous lines of chemotherapy may not be the most appropriate factor. This somewhat surrogate marker of patient susceptibility to support chemotherapy should be reconsidered on a case-by-case basis. It is likely that a unique prognostic factor adapted to all diseases and all therapeutic classes cannot be identified. Information given by previous trials performed in adults provides a unique opportunity to identify strong prognostic factors of toxicity that could be taken into account in paediatric trials. Better use of the knowledge accumulated in adults is a crucial challenge to improve the efficiency of phase I trials in children as shown by Lee and colleagues[15] who studied the correlation between adult and paediatric studies and the way adult data are used. It could provide a strong rationale for investigating subgroups of patients or, preferably, speed up the trial by restricting efforts to a limited number of levels. Ultimately, phase I trials in children could possibly be based on extrapolation from adult data and could be designed to verify this extrapolation.

Is it therefore relevant to identify MTD in two independent groups? One should bear in mind that such a design leads to double the number of children enrolled in toxicity trials. At this stage of the clinical development, few proofs of the activity of the treatment under study are available and it may be more relevant to concentrate efforts on faster and more informative trials. Two independent sample designs in phase I trials should be avoided or reserved to limited situations when there is a strong rationale allowing for a hypothesis on group ordering. Parallel dose search may only be fruitful when the parallel groups are known to have a clear different toxicity profile. We believe that the EMEA recommendations should be carefully reviewed.

Conflict of interest statement

None declared.

Acknowledgements

We are indebted to all children and parents who accepted to participate in this study; all investigators who participated in the trial; Anthony Saul for editing; Catherine Hill for her helpful comments and suggestions; the Fondation pour la Recherche Médicale the Institut National du Cancer and the ligue contre le cancer for their financial support.

Appendix A. Supplementary material

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.ejca.2010.01.024.

REFERENCES

1. Roberts Jr TG, Goulart BH, Squitieri L, et al. Trends in the risks and benefits to patients with cancer participating in phase 1 clinical trials. *JAMA* 2004;**292**:2130–40.
2. Committee for proprietary medical products: note for guidance on evaluation of anticancer medical products in man, *addendum* on paediatric oncology. EMEA/CPMP/EWP/569/02.
3. Adamson PC, Blaney SM. New approaches to drug development in pediatric oncology. *Cancer* 2005;**J11**:324–30.
4. Vassal G, Doz F, Frappaz D, et al. A phase I study of irinotecan as a 3-week schedule in children with refractory or recurrent solid tumors. *J Clin Oncol* 2004;**21**:3844–52.
5. Schwartz GH, Patnaik A, Hammond LA, et al. A phase I study of bizelesin, a highly potent and selective DNA-interactive agent, in patients with advanced solid malignancies. *Ann Oncol* 2003;**14**:775–82.
6. Broniscer A, Gururangan S, MacDonald TJ, et al. Phase I trial of single-dose temozolomide and continuous administration of o6-benzylguanine in children with brain tumors: a pediatric brain tumor consortium report. *Clin Cancer Res* 2007;**13**(22 Pt 1):6712–8.
7. O'Quigley J, Shen LZ. Continual reassessment method: a likelihood approach. *Biometrics* 1996;**52**:673–84.
8. Skolnik JM, Barrett JS, Jayaraman B, et al. Shortening the timeline of pediatric phase I trials: the rolling six design. *J Clin Oncol* 2008;**26**:190–5.
9. Christian MC, Korn EL. The limited precision of phase I trials. *J Natl Cancer Inst* 1994;**86**:1662–3.
10. Paoletti X, O'Quigley J, Maccario J. Design efficiency in dose finding studies. *Comput Stat Data Anal* 2004;**45**:197–214.
11. Ahn C. An evaluation of phase I cancer clinical trial designs. *Stat Med* 1998;**17**:1537–49.
12. O'Quigley J, Paoletti X. Continual reassessment method for ordered groups. *Biometrics* 2003;**59**:430–40.

13. Yuan Z, Yuan R. Isotonic designs for phase I cancer clinical trials with multiple risk groups. *Clin Trial* 2004;**1**:499–508.
14. Meille C, Gentet JC, Barbolosi D, et al. New adaptive method for phase I trials in oncology. *Clin Pharmacol Ther* 2008;**83**:873–81.
15. Lee DP, Skolnik JM, Adamson PC. Pediatric phase I trials in oncology: an analysis of study conduct efficiency. *J Clin Oncol* 2005;**23**:8431–41.
16. Bachelot T, Ray-Coquard I, Catimel G, et al. Multivariable analysis of prognostic factors for toxicity and survival for patients enrolled in phase I clinical trials. *Ann Oncol* 2000;**11**:151–6.
17. Hamberg P, Verweij J. Phase I drug combination trial design: walking the tightrope. *J Clin Oncol* 2009;**27**:4441–3.