Published in final edited form as:

Contemp Clin Trials. 2016 March ; 47: 217–227. doi:10.1016/j.cct.2016.01.009.

Innovations for phase I dose-finding designs in pediatric oncology clinical trials

Adelaide Doussau¹, Birgit Geoerger^{2,3}, Irene Jiménez⁴, and Xavier Paoletti^{5,6}

¹National Institutes of Health, Clinical Center, Department of Bioethics, Bethesda, Maryland, USA

²Gustave Roussy, Pediatric and Adolescent Oncology, Villejuif, France

³CNRS UMR8203, Univ. Paris-Sud, Univ. Paris-Saclay, Villejuif, France

⁴Institut Curie, Pediatric, Adolescent and Young Adults Department, Paris, France

⁵Gustave Roussy, Biostatistics and Epidemiology unit, Villejuif, France

⁶INSERM U1018, CESP, Univ. Paris-Sud, Univ. Paris-Saclay, Villejuif, France

Abstract

Phase I oncology clinical trials are designed to identify the optimal dose that will be recommended for phase II trials. In pediatric oncology, the conduct of those trials raise specific challenges, as the disease is rare with limited therapeutic options. In addition, the tolerance profile is known from adult trials. This paper provides a review of the major recent developments in the design of these trials, inspired by the need to cope with the specific challenges of dose finding in cancer pediatric oncology. We reviewed simulation studies comparing designs dedicated to address these challenges. We also reviewed the design used in published dose-finding trials in pediatric oncology over the period 2009–2014. Three main fields of innovation were identified. First, designs that were developed in order to relax the rules for more flexible inclusions. Second, methods to incorporate data emerging from adult studies. Third, designs accounting for toxicity evaluation at repeated cycles in pediatric oncology. In addition to this overview, we propose some further directions for designing pediatric dose-finding trials.

Keywords

Dose-finding clinical trials; Phase I; Adaptive designs; Oncology; Pediatrics; Review

Corresponding author: Adelaide Doussau, Department of Bioethics, Clinical Center, National Institutes of Health, 10 Center Drive, Building 10/1C118, Bethesda, MD 20892 – 1156 USA, Tel. (+1) 301 594 5523, adelaide.doussaudebazignan@nih.gov, Permanent adelaide.doussau@gmail.com.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

Disclaimer: The views expressed are those of the authors and do not necessarily reflect those of the Clinical Center, the National Institutes of Health, the Public Health Service or the Department of Health and Human Services.

1. Introduction

Phase I trials aim at defining the maximum tolerated dose (MTD) in order to determine the dose recommended for phase II studies (RP2D). Phase I dose-finding trials are performed separately in adults and in children. A major difference between adult and pediatric oncology dose-finding trials is that pediatric trials usually start after completion of adult trials, therefore providing knowledge about the RP2D in adults, the toxicity profile, the mechanisms of action (1–3).

Those trials present specific challenges. Pediatric cancer is rare with limited therapeutic options. Only three targeted treatments have been approved. Children are heterogeneous in size and age, amplifying potential variability of results. The ethical aspects of these trials are different as well. Parents have to consent on behalf of their child for enrollment in trials where the primary objective is the assessment of toxicity (4–7).

The main design used in pediatric dose finding trials has been the 3+3 design (1, 2). In this algorithm-based design, the RP2D is usually the highest dose at which no more than one in three to six patients experience dose limiting toxicities (DLT). This design was constructed under the assumptions that the risk of toxicity increase with the dose of the drug, and that the RP2D would be the dose with less than 33% DLT occurring during the first cycle of treatment. However, in the last decade, alternative designs have been proposed in the setting of pediatric trials. New approaches have been encouraged by the scientific community and patient advocacy groups in order to increase access to trials and to improve the efficiency of the development pipeline of pediatric oncology drugs (8). The purpose of this paper is to review the challenges of pediatric dose-finding trials, and discuss the recent design innovations that can improve scientific efficiency or support pragmatic implementation.

We first reviewed the statistical design used in phase I cancer pediatric trials published between 2009 and 2014, i.e. after the publication of the Rolling 6, one of the first design developed specifically in the field of pediatric oncology (9). The search strategy on the Pubmed database contained the words ‘phase I trials’ AND ‘children OR pediatrics’. The design was extracted from the method section or deduced from the results whenever possible. Among the 88 trials retrieved in Pubmed (Figure 1), the 3+3 design was the most frequently used for pediatric phase I trials (72%) but this proportion strongly decreased in 2013 and 2014.

In the traditional 3+3 design, accrual is suspended between cohorts of three patients, allowing for assessment of DLT before enrolling the next dose level cohort (10). Escalation to the next dose level occurs when no DLT is observed in three patients. If one DLT occurs, then three additional patients are assigned to the same dose level. If DLT occurs in one or more of these extra three patients, this dose is considered unacceptably toxic and is de-escalated. An example is represented through diagrams on Figure 2.A. In the 3+3 design, enrollment is suspended between each cohort, which raises several difficulties:

- Clinical: children in early phase trials have rapidly progressing cancer which argues for limited waiting time before starting treatment; few alternative treatments are available for this population; these treatments are often the last therapeutic option

before palliative care; an eligible child and his parents willing to participate may not understand why the investigator has to postpone treatment for design reasons (1).

- Efficiency: as cancer is a rare disease in children, suspensions of accrual prolong an already long trial.
- Logistics: pediatric trials, even at the early phase, are mostly multicenter trials in which suspension of enrollment creates organizational complexities.

Therefore the community expressed a need for designs that can accelerate the trial and limit trial enrollment suspensions while assuring safety of the included patients. In part 2, we present designs that were developed in order to allow for more flexible inclusions. Two other domains were identified as recent developments in the field of pediatric oncology: In part 3 and 4 we respectively present methods to incorporate data emerging from adult studies and designs accounting for toxicity evaluation at repeated cycles. Finally, in part 5 we discuss our results and propose some further directions for designing pediatric dose-finding trials.

2. Relaxing the rules for more flexible inclusions

Several algorithm-based and model-based designs have been proposed to address this issue (9, 11–15).

1) Rolling 6

The Rolling 6 design is an algorithm-based extension of the 3+3 design proposed in 2008. This design has been reported in the literature for testing anticancer agents in children since 2011; it is increasingly used and was reported in 15 (17%) of the published pediatric clinical trials since 2009 (Figure 1).

In the Rolling 6 design, a total of six children can be enrolled concomitantly at the same dose level (9). The dose level allocated to a new patient is based on the number of patients currently enrolled and evaluated, the number of patients experiencing DLT, and the number of patients whose evaluation is pending at the time of new patient entry. De-escalation occurs when two or more DLT occur at a dose level, whereas escalation can be performed when 3/3, 4/4, 5/5, 5/6 or 6/6 patients are evaluated without DLT. In all other situations, patients can be included at the same dose level, up to a total of six (Figure 2.B). Once 6 patients have been included at the current dose level, inclusions are suspended until at least 5 of those 6 children have completed the first cycle without DLT.

2) Continual Reassessment Method

The use of the Continual Reassessment Method (CRM) has been reported in the pediatric setting since 2003 (19). Four CRM trials were published over the period 2009–2014 (Figure 1).

The CRM is a model-based method. Let x_i be the dose d_k allocated to patient i , among a set of K increasing dose to be studied. The basic idea is to sequentially update the estimates of the probabilities of toxicity of the increasing doses ($d_1, d_2, \dots, d_K$). This is performed after

each inclusion using a working model, $\phi(d_k, a)$, relating the dose d_k to the probability of DLT by a parameter a that depends on the model used. As an example, a one-parameter logistic model with a fixed slope b_c can be used (16):

$$\varphi(d_k, a) = \frac{e^{(b_c \times d_k + a)}}{1 + e^{(b_c \times d_k + a)}}$$

O'Quigley et al. proposed a CRM using Bayesian inference (17) and further extended the method with likelihood approach (18). The maximum likelihood estimator can be used to obtain $\hat{a}$ estimate (18). Given $\hat{a}$, probabilities of DLT can be estimated by $\phi(k, \hat{a})$ for all k . Let τ be the target probability of toxicity of the dose that would be considered appropriate for phase II trials. τ is usually set to 20% in model-based design, which usually corresponds to identifying the dose below the traditional 33% threshold of toxicity used in the 3+3 design; however, in a model-based approach such as the CRM, τ could be set to any clinically relevant threshold. The dose level recommended for the next patient is the level which minimizes the distance between the estimated probability of toxicity at this level and the target τ . Each patient is then included at the best current estimate of the MTD. Inference is based on all available data from all patients previously evaluated (Figure 2.C).

A modification of the CRM has been proposed to minimize accrual suspensions in pediatric brainstem gliomas (12). Onar-Thomas (19) proposed another modification of the CRM in which patient specific body surface area (BSA) adjusted doses were used in the calculation instead of assigned doses (CRM-BSA). This design was developed in the setting of oral agents where it may not be possible to administer the exact theoretical BSA-adjusted dose, given the constraints of limited tablet strength available and the variability of BSA of children of various ages.

3) Time-to-event Continual Reassessment Method

The CRM has been extended to handle toxicity that would occur after a long period without suspending accrual. The time-to-event CRM (TiTE-CRM) provides estimates of the cumulative probability of toxicity over several cycles. It was first proposed to address the issue of late toxicities in radiotherapy dose-finding trials (20). New patients can be enrolled in a trial even though evaluation of some previous patients is still pending. This design was also proposed as an option to relax the rules of inclusion and allow for continuous inclusions in pediatric trials (14).

In the TiTE-CRM, observations of patients who have no DLT at the time of the analysis but have not completed the DLT assessment period are down-weighted in the likelihood. Several weighting schemes have been proposed. The most common consists in weighting the observation proportionally to the length of follow-up; for instance, if the follow-up window is six cycles, a patient who is evaluated for three cycles at the interim analysis and who has not experienced severe toxicity is attributed a weight of $3/6=0.5$.

With this method, the data of all children previously included could be used, even if some follow-up for DLT are still incomplete. In theory, this method would therefore be

appropriate irrespective of the DLT observation window (first cycle, two cycles, or longer), leading to a more flexible access to the trial. Importantly if the TiTE-CRM is used with a DLT assessment period of one cycle, the DLT should be continuously reported during the cycle, which is not common practice. Furthermore, this method relies on the assumption that all patients without DLT are followed up for the whole DLT assessment period. If the DLT assessment is extended over several cycles, this assumption might be violated as a large proportion of patients get off-study due to disease progression after two cycles (21).

4) Escalation with overdose control and Time-to-event Escalation with overdose control

The escalation with overdose control design (EWOC) is an adaptation of the Bayesian CRM, and was proposed to limit the number of patients above the MTD (Figure 2.D). In EWOC designs, the expected proportion of patients treated at doses above the MTD is equal to a specified value α , and the dose chosen for a new patient is the dose for which the probability of DLT is inferior to this predefined α threshold (22).

The EWOC design was further combined with the TiTE approach, resulting in the TiTE-EWOC (13). The observations of patients who have not completed the pre-defined follow-up are down-weighted in the likelihood. It considerably shortens the trial duration as compared to the EWOC design, with no significant increased risk of toxicity. This approach was also proposed as an efficient way to avoid waiting lists in pediatric cancer chemoradiation trials, and to prevent the risk of treating children at doses above the MTD (13).

The EWOC design has been used in one pediatric trial of escalation of irinotecan and gefitinib combination (23). In the trial, the target probability of toxicity defining the RP2D was set as $\tau=25\%$, and the risk of overdose was controlled to be less than $\alpha=30\%$. To our knowledge, the extension for time to event (TiTE-EWOC design) has not been implemented yet in pediatric dose-finding.

5) Comparative analysis

a. Methods—Several studies have compared most of the designs described above through simulations. Due to the sequential and adaptive nature of dose finding trials on small sample sizes, the best way to evaluate the efficiency and safety of designs is to compare them through simulations, such as the “discrete event simulation” method (24, 25). These simulations reproduce thousands of trials under a given realistic dose-toxicity relationship. For each scenario the probability of toxicity at each dose is specified and constitutes the dose-toxicity curve. Simulated patients are sequentially included in the trial at dose levels specified by the allocation rule of the design, according to the toxicities experienced by the patients previously included. At the end of each trial, the RP2D is recorded. The key outcome is the capacity of the design to identify the true target dose.

We reviewed simulations studies comparing designs that specifically addressed dose-finding studies in pediatric oncology. Ten simulation studies were identified (Appendix Table A.1), and evaluated approximately 50 different dose-toxicity curve scenarios (median, 4 scenarios per simulation study, range 1 to 12). We extracted results from nine simulation studies (9, 11, 13, 14, 19, 26–29). In the tenth study, results of simulations were expressed as

distribution of the difference of outcomes between two designs (30). The outcomes of each design could therefore not be extracted.

In each simulation study, we selected a scenario where the true MTD was the third dose and extracted the outcomes of each design. This scenario was selected because dose-finding clinical trials in pediatric oncology frequently select the third dose as RP2D (1, 2). The operating characteristics of a given design evaluated in simulations are directly impacted by the scenario (dose-toxicity curve, starting dose, number of doses to explore, stopping rules...). For instance, the steepness of the dose-toxicity curve, i.e. the true probability of toxicity of the doses adjacent to the true MTD will impact the proportion of trials that correctly identify the MTD or the dose above or below. Therefore, even if we selected scenarios where the true MTD is the third dose in order to identify patterns in the performances of those designs across the statistical literature, the performances of a given design can only be compared within a given scenario, where the other sources of variability are controlled.

For each study we extracted the target probability of toxicity of the dose that would be considered appropriate for phase II trials (τ), the stopping rules, the duration of a cycle of treatment (DLT window), the mean time between patients presenting for enrollment, and the probability of toxicity at each dose for the selected scenario. Most of these features are listed in the appendix (Table A.2, Figure A.1). For each design evaluated in the simulation study, we retrieved the distribution of the dose recommended at the end of the simulations, the toxicity experienced during the trial, and the length of the studies. We presented results per simulation study for each design. For the distribution of the RP2D chosen at the end of a trial and the number of patients allocated per dose level, we represented graphically the results, according to the true probability of dose-limiting toxicity of each dose.

b. Results—Seven studies compared the 3+3, the Rolling 6, the CRM, the TiTE-CRM, or the TiTE-EWOC. In Table 1, we report the outcomes of simulations when the true MTD is the third dose. We summarized the main advantages and limits of the various designs in Table 2.

Algorithm-based designs: Skolnik (9) and Doussau (11) showed that the Rolling 6 indeed achieves the aim of facilitating access to the trial by avoiding delay of enrollment for trial design constraints (Table 2). As a result, the total number of patients included in Rolling 6 trials is slightly higher than in 3+3 trials. The risk of toxicity is similar to the traditional 3+3 design (9, 11). Two simulation studies showed a shorter overall duration as compared to the 3+3 (9, 19). However, with the Rolling 6, it is likely that the number of patients treated in the lowest doses increases as compared to the 3+3 (Appendix Figure A.2), and the proportion of trials correctly identifying the target dose is not improved (Figure 3) (11, 30). As an example, Doussau et al. (11) found that the Rolling 6 led to the inclusion of a median number of 6 children at the levels 1 and 2, while the 3+3 only led to the inclusion of a median of 3 children per level; meanwhile, the proportion of correct recommendations of the true MTD (dose with 20% toxicity) did not vary (38% of simulations with 3+3 and 39% with the Rolling 6) (11). Sposto (30) even suggested that the Rolling 6 design might be more

conservative than the 3+3 (i.e., leading to the recommendation of a dose below the target more often than the 3+3).

Model-based designs: Three simulation studies further challenged the fact that even if the Rolling 6 might be shorter than the 3+3 design, it is not shorter than the CRM (11, 19) or TiTE-CRM (14). Model-based designs tend to identify the target dose more frequently than algorithm-based designs in two simulation studies (from 39 to 45% and 32 to 51% of correct recommendation of the MTD, respectively (11, 14) and no difference in the third simulation (19)) (Figure 3). Of note, the gain in the number of simulations correctly identifying the target dose with model-based designs as compared to algorithm-based designs could be accompanied with a slight increase in the likelihood of recommending the next higher dose above the true MTD. As an example, Onar-Thomas (19) reported that a dose with a risk of toxicity of 49% could be recommended in 6% of the simulations with CRM, as compared to 4% or less with 3+3 or Rolling 6 in the scenario that we have studied (Figure 3). As well, Doussau (11) reported that the CRM could recommend the next higher dose above the true MTD (a dose with 35% risk of DLT) in 21% of simulations with CRM, as compared to less than 15% with 3+3 or Rolling 6. Conversely, Onar and colleagues (28) did not find that the CRM leads to recommendation of doses higher than the true MTD. This risk is therefore limited. In addition the number of patients treated at the lower dose levels can be minimized with the CRM as compared to the Rolling 6 (11, 30) (Appendix Figure A.2), as well as the total number of patients included (Table 2).

Last, Onar-Thomas (19) compared the 3+3, Rolling 6 and CRM, and CRM-BSA. The CRMBSA showed very similar results as compared to traditional CRM (Table 1, Figure 3).

3. Incorporating data from adult trials

Data from adult trials can be incorporated in various ways. First, the safety profile observed in adult data serves to refine the DLT definition or modify the protocol to handle or prevent certain side effects such as nausea. The main difference between DLT definitions in pediatric versus adult studies was a higher tolerance in hematological toxicity. While this was relevant in the area of cytotoxic agents it is less applied now when exploring molecularly targeted agents that often have non-hematologic dose limiting toxicities.

Second, adult data can help to define the starting dose and the range of doses to explore. The most traditional way to build on the knowledge acquired with adults is to start with 80% of the recommended dose in adults, adjusted for the weight or body surface area (BSA), as defined by an international consensus in 1996 (31). In children the dose increments are usually additive (+25% or 30% of the BSA-adjusted recommended dose in adults at each dose increment); this differs from adult trials where the dose increments are usually multiplicative. A review of the literature mostly including chemotherapeutic agents, reassessed that pediatric trials should explore no more than four levels of dose (0.7, 1.0, 1.3 and 1.6 times the adult RP2D) as the RP2Ds in children and in adults were strongly correlated (1). A comprehensive review of pediatric dose-finding trials of molecularly targeted agents also showed that in 13 of 19 trials (69%), the pediatric RP2D was very close to the adult RP2D, ranging between 90% and 130% of the adult RP2D (2) and main

differences were reported in trials of targeted agents where DLTs were not observed and the MTD could not be determined.

Third, the data can be directly incorporated in the dose-escalation scheme in bridging studies (26, 29). The general idea of bridging studies is to share data from several subgroups to improve the accuracy of the MTD estimate. Two types of bridging studies have been proposed: parallel bridging study where a trial can be conducted simultaneously in two subgroups (26, 29); sequential bridging study, where the second group begins after the first group has completed inclusions and uses data from the first group (29). Bridging designs were developed in the setting of acute leukemia where it had been previously observed that children better tolerated aggressive doses adjusted for their weight compared to adults (26, 29). The method has been applied in trials carried out both in adults and children or in heavily pretreated and less heavily pretreated children (32, 33), or in children receiving a new treatment with and without adjunction of radiotherapy (12). Bridging designs could be used for the evaluation of the MTD in teenagers, in trials where adult and teenagers could be included in specific strata.

Statistical techniques include the use of some common data structure through hierarchical models for parallel bridging studies (34) and the elicitation of prior distribution for sequential trials. Hierarchical model is a powerful tool to share information in a unique trial with multiple strata; a model of the dose-toxicity relationship is parameterized by a stratum-specific parameter, whose variance is common to all strata. A smaller variance is assumed when the various strata are similar. This variance parameter is updated throughout the trial. An important limit of the exchangeable formulation is the difficulty to specify orders in the various strata; for instance, one cannot specify that less heavily pre-treated patients have lower risk of DLT as compared to heavily pretreated patients.

Several alternative methods have been proposed to share data structure for bridging studies, although not specifically in the field of pediatric population. An isotonic approach (35) based on the up and down method has been developed to relax assumptions related to parametric models and to avoid the elicitation of prior distribution that is challenging in early phase trials. More recently, Wages and colleagues (36) have developed an approach where several models connecting the different strata at each dose level are hypothesized prior to the trial. The probability of each model is updated based on collected data.

In sequential bridging studies, prior elicitation is crucial as the knowledge gained in the first group is incorporated in the estimation of the dose-toxicity risk of the following group. Two fundamental parameters are specific to these studies. The first is the bridging parameter itself that connects the two groups. The second parameter is a design parameter called the diminishing parameter and allows for down weighting the data used in one group as compared to the other. Prior distribution offers a natural tool to specify a known ordering in the overall risk across strata. This is also applicable in the setting of parallel bridging studies when the expected difference between the strata can be elicited in an informative prior.

Two simulation studies were performed in the pediatric setting (26, 29). O'Quigley and Iasonos (29) evaluated parallel and sequential bridging designs. Bridging designs were

compared to the conduct of a single study including all patients regardless of the subpopulation information, or the conduct of two independent CRM in each subpopulation. Both type of bridging designs increased the probability that the correct dose is identified as the RP2D in the pediatric strata (Figure 3), as compared to a trial conducted independently in a pediatric sample. Nevertheless the gains were modest except when some prior knowledge was elicited on the expected difference between the two strata via some mildly informative prior distribution in a Bayesian framework.

Broglio compared a parallel bridging design in adult and pediatric populations to the conduct of an independent CRM in children (26). The simulation study from Broglio also showed a better ability to select the true MTD with the bridging approach as compared to an independent CRM conducted in the pediatric population (Figure 3).

Bridging designs can potentially lead to the inclusion of a smaller number of children. However, those simulation studies showed that designs are sensitive to the presence of heterogeneity between the subpopulations. This implies that the model should be flexible enough to avoid the risk of bias in the estimation of the RP2D in children, as the risk of toxicity might be different between adults and children (29).

4. Accounting for toxicity evaluation at repeated cycles in pediatric oncology

Several publications have highlighted the need to account for the side effects beyond cycle 1 (20, 21, 27, 37–43). The TiTE-CRM is a possible approach, as presented previously. A key feature is that the risk of DLT is evaluated globally, i.e. the repeated severe toxicities are counted once and the risk of cumulative toxicity requires alternative approaches (43).

A second approach recently proposed is to extend the CRM through utilization of data from the grade of toxicity of all the cycles of treatment (27). In this design, the endpoint is the highest grade toxicity experienced for each patient, repeated over each cycle of treatment. It differs from the traditional DLT endpoint in two ways: 1) it is repeated over time, 2) it is recorded as an ordinal variable (no / moderate / severe toxicity), thus using the information provided by moderate grade 2 toxicities. This approach provides investigators with a probability of “severe” and “moderate or severe” toxicity at each cycle and each dose. Due to repeated data on the same patients, probabilities of toxicity are estimated with a random intercept model. A proportional odds model mixed model (POMM) is used to account for the ordinal nature of the endpoint.

For patient i at cycle j , the outcome Y_{ij} can take three modalities from 1 (mild toxicity) to 3 (severe toxicity). The dose allocated to patient i , x_i , was considered constant over time, i.e. dose modification or temporary dose interruptions were not modelled. Let's denote $m=1, 2$, or 3 the outcome; the proportional mixed effect model including dose, time (cycle) and a random intercept (27) is then:

$$\log\left(\frac{P(Y_{ij} \leq m | x_i, t_{ij})}{1 - P(Y_{ij} \leq m | x_i, t_{ij})}\right) = \alpha_k - \beta_1(x_i) - \beta_2(t_{ij}) - u_i$$

with $u_i \sim N(0, \sigma_0^2)$

β_1 and β_2 are common to all cumulative modalities of toxicity m , as this model relies on the proportional odds assumption, i.e. assumption that the effects of the dose as well as of the time are the same for severe toxicities and for moderate or severe toxicities. Under the assumption that (i) the risk of going off-study for a patient is mainly due to previous toxicities that are observed or progressive disease, (ii) the risk of progression at a dose is independent on the risk of toxicity, maximum likelihood estimates of mixed models provides unbiased estimates (44).

In our simulations (27), the POMM-CRM design led to +15% increase in the number of simulations correctly identifying the true target dose as compared to the CRM (Figure 3). The POMM model can also be used to analyse the final data of any trial in order to have a more precise estimate of the RP2D, or to study the probability of toxicity over time. As an example, the reanalysis of a pediatric trial of Erlotinib+radiotherapy on 20 children showed that the precision of the estimation of the risk of severe toxicity per cycle at the RP2D was strongly increased, from 16% (95% CI 4% to 45%) based on the DLT window of the first two cycles, to 8.6% per cycle (95% CI 4% to 19%) using data over a total of 96 cycles (38). In addition, this reanalysis showed that the risk of toxicity did not increase with time.

5. Discussion

Our study reviews the designs proposed and used in recent pediatric dose-finding oncology trials, and shows an increasing use of new designs such as the Rolling 6 or the CRM. It also shows that various innovative designs have been proposed to tackle the specific challenges of dose-finding trials in pediatric oncology. Three main axes of development have been to limit accrual suspension before escalating to the next higher dose level, while ensuring children's safety, to extrapolate data from adults' trials in bridging studies and to integrate richer endpoints such as ordinal variables and repeated measurements.

To avoid trial suspension, the Rolling 6, or a modification of the CRM seems to appropriately decrease the number of patients who cannot enter the trial. Due to their slightly better ability to identify the target dose, while not increasing the total number of subjects, we recommend that model-based design should be preferred over algorithm-based designs. Furthermore, because the Rolling 6 might lead to a high number of patients at the lower levels to explore, the choice of the starting dose is of paramount importance in order to minimize the number of children treated at doses lower than the optimal dose. Of note, rapid recruitment may also lead to study suspension in the Rolling 6 design, as shown in a simulation study (11); conversely, the rolling 6 is equivalent to the 3+3 when recruitment is very slow.

Our review however underlines the fact that most dose-finding trials perform relatively poorly with traditional sample sizes, as the simulations rarely exceed 60% of recommendation of the correct target dose, even when some very refined designs are implemented. Despite the amount of methodological studies supporting the advantages of model-based designs, their implementation is relatively infrequent (45), but they are gaining popularity, and are now recognized by regulatory agencies as designs to consider for dose-finding trials.

Importantly, our analysis is based on a review of simulation studies comparing various designs. The performance of any design is influenced by the dose-toxicity curve, and various aspects of each design such as stopping rules. In addition, there are many choices that can be made to tailor a CRM to a specific clinical trial setting. Therefore quantitative comparisons between various articles should be considered carefully even when a similar scenario is used across studies as in our analysis.

In the era of molecularly targeted agents, the hypothesis of “the more, the better” that is underlying all algorithmic or model-based approaches based on toxicity endpoints is challenged, as the therapeutic index might be wider, and the optimal biologic dose might not be the MTD. It is therefore important to use activity or efficacy endpoints as early as possible in the development of new drugs. Two approaches might be promising: first, develop dose-finding designs for co-primary endpoints using both toxicity and efficacy; second, promote seamless phase I/II dose-finding trials (46). These approaches should be encouraged. Several statistical publications provide innovative designs considering both safety and activity criteria (36, 47–52). We did not include them in this review as those aspects were not developed specifically in pediatric oncology dose-finding trials. As well, we did not detail methods for dual agents combination (48, 53–55) for the same reason, even if associations of two investigational agents are increasingly considered in the pediatric population.

Importantly, the literature comparing the RP2D in children to adults could justify challenging the traditional rule of starting the trials at 80% of the BSA-adjusted adult dose, especially in the era of molecular targeted agents (1, 2). When the PK/PD profile of the drug is not expected to be that different from adults, and when no cumulative toxicity is expected, it would seem reasonable to start upfront giving 100% of the BSA-adjusted adult dose to children, with rules to decrease the dose to 80% if necessary. However, such an approach might not apply to neonates or infants for whom the metabolism might differ a lot. An alternative approach could be to administer one cycle of the treatment at 80% of the adult dose and do intrapatient dose-escalation to 100% after the second cycle. The RP2D in children rarely exceeds 130% of the adult approved dose; therefore, it could even be proposed not to investigate doses higher than the adult approved dose if biodisponibility as measured with PK is similar and no other reasons such as lower sensitivity of the molecular target in the pediatric disease are present. This would dramatically simplify the design, as no or very limited dose escalation would be necessary, and the trial could be a phase I/II trial, with an early stopping rule for safety upon completion of the phase I part of the study. All approaches that minimize the number of cycles administered below the adult approved dose while not compromising safety should be encouraged. In pediatric oncology phase I trials

where parents have to consent on the behalf of their child, there is an incentive to maximize as much as possible the potential clinical benefit for the children. A framework of options to consider in the choice of the design is summarized as a flow-diagram (Figure 4).

There is a growing amount of literature showing that pediatric oncologists are willing to reform the development path of drugs (8). As an example, hematologists argued that phase I trials should not be necessary for children with leukemia if the drug was already tested in phase I trials for children with solid tumors (56). The development of molecularly targeted agents is transforming this rare disease in even more scarce subgroups of patients sharing a biologic target. Therefore, the need for innovations on how to conduct the research as efficiently as possible in the era of personalized medicine is exemplified.

Acknowledgments

We would like to thank Bradley Otterson for English editing and Christine Grady for comments on earlier drafts of this paper.

Appendix

Characteristics of the simulation studies

Table A. 1

Reference	Designs	τ (%)	Pediatric MTD dose level	DLT window (days)	MIP (days)	Ratio MIP/DLT window	N simulations
Skolnik, 2008 (9)	3+3, Rolling 6	NA [*]	3	21	10	2.1	1000
Onar, 2009 (28)	3+3, CRM	20	3	NA	NA	NA	10000
Onar-Thomas, 2010 (19)	3+3, Rolling 6, CRM, CRM-BSA	20	3	28	20	1.4	1000
Sposto, 2011 ^{**} (30)	3+3, Rolling 6	NA [*]	**	28	56	0.5	1000
Doussau, 2012 (11)	3+3, Rolling 6, CRM	20	3	21	10	2.1	1000
Zhao, 2011 (14)	Rolling 6, TITE-CRM	25	3	42	36.5	1.2	2000
Mauguen, 2011 (13)	EWOC, TITE-EWOC	33	2 ^{***}	21	14	1.5	1000
Doussau, 2013 (27)	CRM, POMM-CRM	20	4 ^{***}	NA	NA	NA	1000
O'Quigley 2014 (29)	CRM, Bridging design	20	3 [§]	NA	NA	NA	1000
Broglio, 2015 (26)	CRM, Bridging design	30	3 [§]	NA	NA	NA	-

^{*} No model-based design.

^{**} The results of this study where expressed as paired difference between designs and do not provide individual results for each design.

^{***} We selected an alternative scenario when the simulation study was not evaluating a scenario where the true MTD was the third dose.

[§] In bridging studies, the true MTD in the adult cohort was the second dose.

MIP: mean interpatient arrival time.

MTD: maximum tolerated dose. CRM Continual Reassessment Method. BSA: body surface area. POMM: proportional odds mixed effect model. TITE: Time to event. EWOC: escalation with overdose control.

τ : target probability of toxicity of the dose that would be considered appropriate for phase II trials.

Table A. 2
Stopping rules for model-based designs implemented in simulation studies

Reference	Designs	Cohort size / escalation rules	Dose-skipping in escalation	Stopping-rule of the trial
Onar, 2009 (28)	CRM	3 slots; 2 patients needed to be evaluated before next escalation.	Not allowed	Dynamic: “ <i>once six or more patients are treated at a dose, we do not stop the trial unless treating two more hypothetical patients at that dose level would not lead to escalation</i> ”
Onar-Thomas, 2010 (19)	CRM, CRM-BSA	3 slots; 2 patients needed to be evaluated before next escalation.	Not allowed	Dynamic: when at least 6 patients have been treated at the MTD and inclusion of 2 additional patients would not lead to escalation.
Doussau, 2012 (11)	CRM	Two patients at the first dose level. When the evaluation of three patients was pending at the current dose, accrual for new patients was temporarily suspended until the first DLT. No restriction to accrual after the first DLT (model-based allocation).	Not allowed	Dynamic: stop once 10 patients have been included at a given dose (nonconsecutive patients).
Zhao, 2011 (14)	TITE-CRM	Escalation from the current dose not allowed until at least one patient at the current dose has completed follow-up.	Not allowed	Fixed: maximum of 24 patients
Mauguen, 2011* (13)	EWOC, TITE-EWOC	EWOC: new patient cannot be enrolled before the previous patient enrolled has finished follow-up. TITE-EWOC: no restriction regarding pending evaluations (“ <i>a weight allows to take into account all available information, including incomplete information, in the estimation of the dose-toxicity relationship</i> ”).	Not allowed	Fixed: 25 patients
Doussau, 2013 (27)	CRM, POM M-CRM	New patient cannot be enrolled before the previous patient	Not allowed	Fixed: 30 patients

Reference	Designs	Cohort size / escalation rules	Dose-skipping in escalation	Stopping-rule of the trial
		enrolled has finished follow-up.		
O'Quigley 2014 (29)	CRM, Bridging design	Unspecified	Unspecified	Fixed: 16 in Group 1 and 16 in Group 2. Total = 32.
Broglio, 2015 (26)	CRM, Bridging design	3 adult patients enrolled per month vs. 1 child/month. Dose-toxicity curves are updated once each subject has completed evaluation. At the 3 lower dose levels, information about 2 patients is required prior to escalation. At dose levels higher than the 3 rd level, information is required from at least 3 patients.	Not allowed for adults. Allowed for children provided that there is complete information for 2 adults at this level and the dose level above, and both levels are considered safe.	Dynamic; <i>"First, a minimum number of patients must be treated at or around the MTD. In the adult population, at least 8 patients must be treated at the MTD and at least 20 patients treated within one dose level (lower or higher) of the MTD. In the pediatric population, at least 5 patients must be treated at the MTD and at least 10 patients treated within one dose level of the MTD. The second condition is to ensure that the selection of the MTD is robust and that a higher dose level should not still be considered. For this, we assume three additional hypothetical patients treated at the current MTD with no DLT. If, with this additional information, an update of the CRM would not recommend escalation, the MTD is considered well characterized."</i> Maximum sample size is 35 adults and 25 children.

MTD: maximum tolerated dose. CRM Continual Reassessment Method. BSA: body surface area. POMM: proportional odds mixed effect model. TITE: Time to event. EWOC: escalation with overdose control.

Figure A.1. Dose-toxicity curves of the reported scenarios. The 3rd dose level is the true MTD, unless specified. * In Doussau 2013, no scenario was reported with the third dose as true MTD, we report a scenario where the true MTD was the fourth dose. ** In Mauguen 2011, no scenario was reported with the third dose as true MTD, we report a scenario where the true MTD was the second dose. #In Onar-Thomas 2010, level 0 indicates a dose level that was below the starting dose.

Probability of toxicity (%)

Figure A.2. Patient allocation according to the toxicity of the dose, when the 3rd dose level is the true MTD. Short dashed lines correspond to model-based designs. The horizontal axis represents true probability of dose-limiting toxicity of the dose level. The upper part represents the proportion of patients allocated to a dose, plotted against the probability of toxicity of this dose. The lower part represents the mean or median number of patients allocated to the dose, plotted against the probability of toxicity of this dose.

Reference: Onar 2009 (28), Onar-Thomas 2010 (19), O`Quigley 2014 (29), Doussau 2012 (11), Doussau 2013 (27), Broglio 2015 (26).

Reference

1. Lee DP, Skolnik JM, Adamson PC. Pediatric phase I trials in oncology: an analysis of study conduct efficiency. *J Clin Oncol.* 2005; 23(33):8431–8441. [PubMed: 16293874]
2. Paoletti X, Geoerger B, Doz F, Baruchel A, Lokiec F, Le Tourneau C. A comparative analysis of paediatric dose-finding trials of molecularly targeted agent with adults' trials. *Eur J Cancer.* 2013; 49(10):2392–2402. [PubMed: 23540589]
3. Smith M, Ho PT. Pediatric drug development: a perspective from the Cancer Therapy Evaluation Program (CTEP) of the National Cancer Institute (NCI). *Invest New Drugs.* 1996; 14(1):11–22. [PubMed: 8880389]
4. Cousino MK, Zyzanski SJ, Yamokoski AD, Hazen RA, Baker JN, Noll RB, et al. Communicating and understanding the purpose of pediatric phase I cancer trials. *J Clin Oncol.* 2012; 30(35):4367–4372. [PubMed: 23071225]
5. Estlin EJ, Cotterill S, Pratt CB, Pearson AD, Bernstein M. Phase I trials in pediatric oncology: perceptions of pediatricians from the United Kingdom Children's Cancer Study Group and the Pediatric Oncology Group. *J Clin Oncol.* 2000; 18(9):1900–1905. [PubMed: 10784630]
6. Hazen RA, Zyzanski S, Baker J, Drotar D, Kodish E. Communication about the Risks and Benefits of Phase I Pediatric Oncology Trials. *Contemporary clinical trials.* 2015
7. Yap TY, Yamokoski AD, Hizlan S, Zyzanski SJ, Angiolillo AL, Rheingold SR, et al. Informed consent for pediatric phase 1 cancer trials: physicians' perspectives. *Cancer.* 2010; 116(13):3244–3250. [PubMed: 20564626]
8. Vassal G, Zwaan CM, Ashley D, Le Deley MC, Hargrave D, Blanc P, et al. New drugs for children and adolescents with cancer: the need for novel development pathways. *Lancet Oncol.* 2013; 14(3):e117–e124. [PubMed: 23434337]
9. Skolnik JM, Barrett JS, Jayaraman B, Patel D, Adamson PC. Shortening the timeline of pediatric phase I trials: the rolling six design. *J Clin Oncol.* 2008; 26(2):190–195. [PubMed: 18182661]
10. Storer BE. Design and analysis of phase I clinical trials. *Biometrics.* 1989; 45(3):925–937. [PubMed: 2790129]
11. Doussau A, Asselain B, Le Deley MC, Geoerger B, Doz F, Vassal G, et al. Dose-finding designs in pediatric phase I clinical trials: comparison by simulations in a realistic timeline framework. *Contemporary clinical trials.* 2012; 33(4):657–665. [PubMed: 22521954]
12. Geoerger B, Hargrave D, Thomas F, Ndiaye A, Frappaz D, Andreiuolo F, et al. Innovative Therapies for Children with Cancer pediatric phase I study of erlotinib in brainstem glioma and relapsing/refractory brain tumors. *Neuro-oncology.* 2011; 13(1):109–118. [PubMed: 20974795]
13. Mauguen A, Le Deley MC, Zohar S. Dose-finding approach for dose escalation with overdose control considering incomplete observations. *Stat Med.* 2011; 30(13):1584–1594. [PubMed: 21351289]
14. Zhao L, Lee J, Mody R, Braun TM. The superiority of the time-to-event continual reassessment method to the rolling six design in pediatric oncology Phase I trials. *Clin Trials.* 2011; 8(4):361–369. [PubMed: 21610004]
15. Neuenschwander B, Branson M, Gsponer T. Critical aspects of the Bayesian approach to phase I cancer trials. *Stat Med.* 2008; 27(13):2420–2439. [PubMed: 18344187]
16. Paoletti X, Kramar A. A comparison of model choices for the Continual Reassessment Method in phase I cancer trials. *Statistics in medicine.* 2009; 28(24):3012–3028. [PubMed: 19672839]
17. O'Quigley J, Pepe M, Fisher L. Continual reassessment method: a practical design for phase 1 clinical trials in cancer. *Biometrics.* 1990; 46(1):33–48. [PubMed: 2350571]
18. O'Quigley J, Shen LZ. Continual reassessment method: a likelihood approach. *Biometrics.* 1996; 52(2):673–684. [PubMed: 8672707]
19. Onar-Thomas A, Xiong Z. A simulation-based comparison of the traditional method, Rolling-6 design and a frequentist version of the continual reassessment method with special attention to trial

- duration in pediatric Phase I oncology trials. *Contemporary clinical trials*. 2010; 31(3):259–270. [PubMed: 20298812]
20. Cheung YK, Chappell R. Sequential designs for phase I clinical trials with late-onset toxicities. *Biometrics*. 2000; 56(4):1177–1182. [PubMed: 11129476]
21. Postel-Vinay S, Collette L, Paoletti X, Rizzo E, Massard C, Olmos D, et al. Towards new methods for the determination of dose limiting toxicities and the assessment of the recommended dose for further studies of molecularly targeted agents - Dose-Limiting Toxicity and Toxicity Assessment Recommendation Group for Early Trials of Targeted therapies, an European Organisation for Research and Treatment of Cancer-led study. *Eur J Cancer*. 2014; 50(12):2040–2049. [PubMed: 24880774]
22. Babb J, Rogatko A, Zacks S. Cancer phase I clinical trials: efficient dose escalation with overdose control. *Statistics in medicine*. 1998; 17(10):1103–1120. [PubMed: 9618772]
23. Brennan RC, Furman W, Mao S, Wu J, Turner DC, Stewart CF, et al. Phase I dose escalation and pharmacokinetic study of oral gefitinib and irinotecan in children with refractory solid tumors. *Cancer Chemother Pharmacol*. 2014; 74(6):1191–1198. [PubMed: 25257509]
24. Barrett J, Skolnik J, Jayaraman B, Patel D, Adamson P. Discrete event simulation applied to pediatric phase I oncology designs. *Clin Pharmacol Ther*. 2008; 84(6):729–733. [PubMed: 18923388]
25. Barrett JS, Jayaraman B, Patel D, Skolnik JM. A SAS-based solution to evaluate study design efficiency of phase I pediatric oncology trials via discrete event simulation. *Comput Methods Programs Biomed*. 2008; 90(3):240–250. [PubMed: 18276034]
26. Broglio KR, Sandalic L, Albertson T, Berry SM. Bayesian dose escalation in oncology with sharing of information between patient populations. *Contemporary clinical trials*. 2015; 44:56–63. [PubMed: 26173185]
27. Doussau A, Thiebaut R, Paoletti X. Dose-finding design using mixed-effect proportional odds model for longitudinal graded toxicity data in phase I oncology clinical trials. *Stat Med*. 2013; 32(30):5430–5447. [PubMed: 24018535]
28. Onar A, Kocak M, Boyett JM. Continual reassessment method vs. traditional empirically based design: modifications motivated by Phase I trials in pediatric oncology by the Pediatric Brain Tumor Consortium. *J Biopharm Stat*. 2009; 19(3):437–455. [PubMed: 19384687]
29. O'Quigley J, Iasonos A. Bridging Solutions in Dose Finding Problems. *Stat Biopharm Res*. 2014; 6(2):185–197. [PubMed: 25071878]
30. Sposto R, Groshen S. A wide-spectrum paired comparison of the properties of the Rolling 6 and 3+3 Phase I study designs. *Contemp Clin Trials*. 2011; 32(5):694–703. [PubMed: 21554993]
31. Smith M, Bernstein M, Bleyer WA, Borsi JD, Ho P, Lewis IJ, et al. Conduct of phase I trials in children with cancer. *J Clin Oncol*. 1998; 16(3):966–978. [PubMed: 9508179]
32. Geoerger B, Vassal G, Doz F, O'Quigley J, Wartelle M, Watson AJ, et al. Dose finding and O6-alkylguanine-DNA alkyltransferase study of cisplatin combined with temozolomide in paediatric solid malignancies. *Br J Cancer*. 2005; 93(5):529–537. [PubMed: 16136028]
33. Vassal G, Doz F, Frappaz D, Imadlou K, Sicard E, Santos A, et al. A phase I study of irinotecan as a 3-week schedule in children with refractory or recurrent solid tumors. *J Clin Oncol*. 2003; 21(20):3844–3852. [PubMed: 14551303]
34. Neuenschwander B, Wandel S, Roychoudhury S, Bailey S. Robust exchangeability designs for early phase clinical trials with multiple strata. *Pharmaceutical statistics*. 2015
35. Yuan Z, Chappell R. Isotonic designs for phase I cancer clinical trials with multiple risk groups. *Clin Trials*. 2004; 1(6):499–508. [PubMed: 16279290]
36. Wages NA, Read PW, Petroni GR. A Phase I/II adaptive design for heterogeneous groups with application to a stereotactic body radiation therapy trial. *Pharmaceutical statistics*. 2015; 14(4):302–310. [PubMed: 25962576]
37. Benda N. Model-based approaches for time-dependent dose finding with repeated binary data. *Stat Med*. 2010; 29(10):1096–1106. [PubMed: 20082362]
38. Doussau A, Thiebaut R, Geoerger B, Schoffski P, Floquet A, Le Deley MC, et al. A new approach to integrate toxicity grade and repeated treatment cycles in the analysis and reporting of phase I dose-finding trials. *Ann Oncol*. 2015; 26(2):422–428. [PubMed: 25403589]

39. Fernandes LL, Taylor JM, Murray S. Adaptive Phase I Clinical Trial Design Using Markov Models for Conditional Probability of Toxicity. *J Biopharm Stat.* 2015
40. Paoletti X, Doussau A, Ezzalfani M, Rizzo E, Thiebaut R. Dose finding with longitudinal data: simpler models, richer outcomes. *Stat Med.* 2015
41. Postel-Vinay S, Gomez-Roca C, Molife LR, Anghan B, Levy A, Judson I, et al. Phase I trials of molecularly targeted agents: should we pay more attention to late toxicities? *J Clin Oncol.* 2011; 29(13):1728–1735. [PubMed: 21444876]
42. Simon R, Freidlin B, Rubinstein L, Arbus SG, Collins J, Christian MC. Accelerated titration designs for phase I clinical trials in oncology. *J Natl Cancer Inst.* 1997; 89(15):1138–1147. [PubMed: 9262252]
43. Sinclair K, Whitehead A. A Bayesian approach to dose-finding studies for cancer therapies: incorporating later cycles of therapy. *Stat Med.* 2014; 33(15):2665–2680. [PubMed: 24590816]
44. Carriere I, Bouyer J. Choosing marginal or random-effects models for longitudinal binary responses: application to self-reported disability among older persons. *BMC Med Res Methodol.* 2002; 2:15. [PubMed: 12466027]
45. Mussai FJ, Yap C, Mitchell C, Kearns P. Challenges of clinical trial design for targeted agents against pediatric leukemias. *Frontiers in oncology.* 2014; 4:374. [PubMed: 25610810]
46. Chow SC, Chang M. Adaptive design methods in clinical trials - a review. *Orphanet J Rare Dis.* 2008; 3:11. [PubMed: 18454853]
47. Braun TM. The bivariate continual reassessment method. extending the CRM to phase I trials of two competing outcomes. *Control Clin Trials.* 2002; 23(3):240–256. [PubMed: 12057877]
48. Mandrekar SJ, Cui Y, Sargent DJ. An adaptive phase I design for identifying a biologically optimal dose for dual agent drug combinations. *Stat Med.* 2007; 26(11):2317–2330. [PubMed: 17016867]
49. Mandrekar SJ, Qin R, Sargent DJ. Model-based phase I designs incorporating toxicity and efficacy for single and dual agent drug combinations: methods and challenges. *Stat Med.* 2010; 29(10):1077–1083. [PubMed: 20419760]
50. Thall, PF.; Cook, JD. Using both efficacy and toxicity for dose-finding. In: Chevret, S., editor. *Statistical methods for dose-finding experiments.* Chichester, West Sussex (England): Hoboken (NJ): Wiley Press; 2006. p. 275-285.
51. Thall PF, Russell KE. A strategy for dose-finding and safety monitoring based on efficacy and adverse outcomes in phase I/II clinical trials. *Biometrics.* 1998; 54(1):251–264. [PubMed: 9544520]
52. Zhang W, Sargent DJ, Mandrekar S. An adaptive dose-finding design incorporating both toxicity and efficacy. *Stat Med.* 2006; 25(14):2365–2383. [PubMed: 16220478]
53. Wages NA, Conaway MR, O'Quigley J. Continual reassessment method for partial ordering. *Biometrics.* 2011; 67(4):1555–1563. [PubMed: 21361888]
54. Wang K, Ivanova A. Two-dimensional dose finding in discrete dose space. *Biometrics.* 2005; 61(1):217–222. [PubMed: 15737096]
55. Yin G, Yuan Y. Bayesian dose finding in oncology for drug combinations by copula regression. *Journal of the Royal Statistical Society.* 2009; 58:211–224.
56. Devidas M, Anderson JR. Considerations in the design of clinical trials for pediatric acute lymphoblastic leukemia. *Clin Investig (Lond).* 2013; 3(9)
57. Le Tourneau C, Lee JJ, Siu LL. Dose escalation methods in phase I cancer clinical trials. *J Natl Cancer Inst.* 2009; 101(10):708–720. [PubMed: 19436029]

Figure 1. Designs used in pediatric phase I cancer clinical trials

Comprehensive review of published phase I trials conducted in children between 2009 and 2014: 88 publications, 90 trials, as 2 trials included 2 independent substudies. Search through Pubmed performed in December 2014.

Figure 3. Distribution of the dose recommended in 8 simulations studies, when the 3rd dose level is the true MTD

Horizontal axis = True probability of dose-limiting toxicity of the dose level. Vertical axis = Proportion of simulated trials recommending this dose as recommended phase II dose (RP2D). Short dashed lines correspond to model-based designs. BSA: body surface area. CRM: Continual reassessment method. POMM: proportional odds mixed model. Tite: Time to event. The scenarios are described in appendix. Reference: Skolnik 2008 (9), Onar 2009 (28), Onar-Thomas 2010 (19), Doussau 2012 (11), Zhao 2011 (14), Doussau 2013 (27), Broglio 2015 (26), O'Quigley 2014 (29).

Figure 4. Flow diagram: Designing a phase I dose-finding trial in pediatric oncology

Table 1

Comparative analysis of designs based on simulation studies

Results correspond to a scenario where the true MTD is the third level, unless specified. The scenarios are described in appendix.

Study	Design	DLT rate per trial (mean or median, %)	Trial duration (mean or median, days)	N patients enrolled (mean or median)	N patients skipped (mean or median)	N DLT (mean or median)
Skolnik 2008 (9)	3+3	18,8 -	350 (SD, 84)	17 (SD, 4)	-	3,2 (SD, 1,1)
	Rolling 6	16,5 -	294 (SD, 75)	20 (SD, 5)	-	3,3 (SD, 1,1)
Onar 2009 (28)	3+3	19,1 (range, 5–67)	-	18 (range, 3–36)	-	-
	CRM	18,2 (range, 0–100)	-	15 (range, 2–37)	-	-
Onar-Thomas 2010 (19)	3+3	22 (IC95%, 13–67)	461 (IC95%, 85–847)	14 (IC95%, 3–22)	-	-
	CRM	22 (IC95%, 11–67)	401 (IC95%, 79–768)	15 (IC95%, 3–25)	-	-
	CRM-BSA	22,2 (IC95%, 13–67)	401 (IC95%, 57–695)	12 (IC95%, 2–20)	-	-
	Rolling 6	22,2 (IC95%, 13–67)	406 (IC95%, 57–706)	12 (IC95%, 2–20)	-	-
Doussau 2012 (11) *	3+3	19,2 -	271 -	17 -	8 -	3 -
	Rolling 6	16,7 -	238 -	20 -	3 -	3 -
	CRM	18,2 -	265 -	21 -	3 -	4 -
Zhao 2011 (14)	Rolling 6	22 -	246 (Q1–Q3, 180–300)	15 -	5 (Q1–Q3, 2–8)	5 -
	TITE-CRM	21 -	279 (Q1–Q3, 250–310)	24 -	0 (Q1–Q3, 0–0)	3,3 -
Mauguen 2011 (MTD=2) ** (13)	EWOC	- -	909 (Q1–Q3, 814–982)	25 #	-	10 (Q1–Q3, 8–11)
	TITE-EWOC	- -	357 (Q1–Q3, 342–369)	25 #	-	10 (Q1–Q3, 9–12)
Doussau 2012 (MTD=4) * (11)	3+3	17 (Q1–Q3, 13–19)	323 (Q1–Q3, 268–287)	20 (Q1–Q3, 17–23)	11 (Q1–Q3, 8–14)	3 (Q1–Q3, 2–4)
	Rolling 6	13 (Q1–Q3, 11–17)	305 (Q1–Q3, 247–360)	25 (Q1–Q3, 21–29)	4 (Q1–Q3, 3–7)	3 (Q1–Q3, 2–4)
	CRM	15 (Q1–Q3, 13–19)	299 (Q1–Q3, 259–342)	20 (Q1–Q3, 20–26)	4 (Q1–Q3, 2–7)	4 (Q1–Q3, 3–4)

* In Doussau 2012 study, the precision of all parameters was given for the scenario with the fourth dose as true MTD. We therefore report that scenario in addition to the scenario with the third dose as true MTD.

** In Mauguen 2011, no scenario was reported with the third dose as true MTD, we report a scenario where the true MTD was the second dose.

In Mauguen 2011, the sample size was fixed to n=25 patients.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

BSA: Body-surface area. CRM: Continual reassessment method. EWOC: escalation with overdose control. MTD: maximum tolerated dose. POMM: proportional odds mixed model. TTE: Time to event. SD: standard deviation. Q1–Q3: interquartile range. IC: Confidence interval.

Table 2
Comparative analysis of four competing designs

Adapted and extended from (57).

	Algorithm-based designs		Model-based designs	
	3+3	Rolling 6	CRM	TITE-CRM
Dose-escalation	Algorithm based, using the data of patients previously evaluated on current dose-level.		Model-based, using the data of all patients previously evaluated (all doses).	Model-based, using the data of all patients previously included (all doses). Incomplete follow-up are down-weighted in the model, as compared to complete evaluations.
Suspension of inclusion	Every 3 patients	If 6 patients are included and data on evaluable patients does not allow dose escalation	Flexible, can be designed or relaxed on demand. Suspension might be required for the first stage of frequentist CRM.	Flexible, can be designed or relaxed on demand.
RP2D	Dose at which less than 2 DLT in 6 patients occurred		Dose which probability of DLT is closest to the target toxicity rate (e.g. 20%)	
Main advantages	Easy to implement and safe.		Target toxicity level is defined explicitly. More rapid dose-escalation. Patients inclusions are more concentrated around true MTD. RP2D can be estimated with its confidence interval (better precision).	
		Limited suspension of inclusion. Moderately shorter trial (as compared to 3+3)	Moderately shorter trial (as compared to 3+3)	Allows toxicity information to contribute to dose escalation before all patients finished follow-up. Take into account late onset toxicities
Main limits	Appropriate to identify a dose level inferior to 33% toxicity. Target toxicity level is not defined explicitly. High proportion of patients treated at low, potentially ineffective doses. Low efficiency to identify the true MTD. Long trial if several doses have to be explored before the RP2D.		Computation after each patient or cohort of patients. Need biostatistical support for dose escalation decisions.	
	Suspension of inclusions.	Increased number of patients included, therefore increased financial costs as compared to 3+3 (30) Extra patients included at lower doses as compared to CRM.	Increased number of patients included, therefore increased financial costs as compared to 3+3	

CRM: Continual reassessment method. DLT: dose limiting toxicity. MTD: maximum tolerated dose. RP2D: Recommended phase II dose. TITE: Time to event.