Published in final edited form as:

Stat Med. 2017 January 30; 36(2): 215–224. doi:10.1002/sim.6910.

Implementation of adaptive methods in early-phase clinical trials

Gina R. Petroni¹, Nolan A. Wages¹, Gautier Paux², and Frédéric Dubois²

¹Division of Translational Research and Applied Statistics, Department of Public Health Sciences, The University of Virginia, Charlottesville, VA 22908

²Oncology Clinical Biostatistics, Institut de Recherches Internationales Servier (I.R.I.S.), 92284 Suresnes Cedex (France)

Abstract

There has been constant development of novel statistical methods in the design of early-phase clinical trials since the introduction of model-based designs, yet the traditional or modified 3+3 algorithmic design remains the most widely used approach in dose-finding studies. Research has shown the limitations of this traditional design compared to more innovative approaches yet the use of these model-based designs remains infrequent. This can be attributed to several causes including a poor understanding from clinicians and reviewers into how the designs work and how best to evaluate the appropriateness of a proposed design. These barriers are likely to be enhanced in the coming years as the recent paradigm of drug development involves a shift to more complex dose-finding problems. This article reviews relevant information that should be included in clinical trial protocols to aid in the acceptance and approval of novel methods. We provide practical guidance for implementing these efficient designs with the aim of augmenting a broader transition from algorithmic to adaptive model-guided designs. In addition we highlight issues to consider in the actual implementation of a trial once approval is obtained.

Keywords

dose-finding; adaptive methods; protocol development

1. Motivation

Statisticians and researchers working on designs for early-phase clinical trials have advocated the use of more innovative approaches to efficiently and accurately address the objectives of finding appropriate doses or dose combinations to merit further research. The statistical and medical literature is replete with reviews, justification, and recommendations on the use of more novel designs [1–12]. These designs are often referred to as ‘adaptive designs’ and are based upon the premise that, as safety and other information are acquired, the information should be used to guide dose recommendations for future study participants. In this article we do not advocate the use of any particular body of adaptive designs instead it is assumed that an adaptive design is planned. The basic framework on how these dose

recommendations are determined is understood generally, however, many barriers, some artificial some genuine, have been encountered that have hindered the use of these adaptive designs. Frequently, the technical statistical methodology and protocol specific operating characteristics that govern these adaptive designs present a challenge for clinicians to accept, reviewers to evaluate and study teams to implement. In this article we provide recommendations on the minimal information that should be included in a protocol to aid in the understanding and approval of the protocol by various review committees (scientific/protocol review committees, regulatory agencies, pharmaceutical/industry, NIH, etc.), and importantly should be considered in order to conduct the protocol specific proposed design.

2. Essential protocol information

In general the study specific protocol is a document that provides the background and rationale; the study objectives and endpoints; the target population and treatment schedule; the design criteria and methodology; the statistical considerations; and the regulatory reporting schedule. This document is the main source used to support approval and application of the clinical trial therefore it is essential that all information that is needed to understand, evaluate and implement the proposed design be included. The protocol document serves as the source to guide decisions as participants are accrued to the study and should contain a cohesive description of the decision framework and if/when deviations/exceptions are allowed.

2.1. Study objectives

The type of study design employed is directed by the objectives that the research is intended to accomplish. These objectives need to be clearly delineated at the start of the design process and consequently specified in the protocol.

2.1.1. Endpoints

2.1.1.1. Safety: In all early-phase clinical trials, whether or not the treatment is truly first-in-human or is intended for a new disease indication and regardless of whether or not the treatment dose or regimen is expected to be safe, adverse events (AE) must be captured. Attribution of AEs can be posited but if they are severe the impact on decision rules for dose escalation/de-escalation/continuation is usually the same since safety is a mandatory endpoint in early-phase research. The frequency of dose limiting toxicities/AEs (DLT) is a measure often used to assess safety and as such must be defined clearly in the protocol. This includes specification of the system used for the coding and grading of AEs as well as the time frame for assessing DLTs. For example, the Common Terminology Criteria for Adverse Events (CTCAE) is the system most commonly employed in oncology trials with the first cycle or two of treatment defining the time frame. If AE exceptions will be allowed based upon classification, grade or attribution then it must be part of the definition of a DLT, see supplemental materials Table S1 for an example.

2.1.1.2. Beyond safety: Historically in oncology dose finding studies the main objective was to identify the maximum tolerated dose among a range of predefined dose levels. Each increasing dose level was associated with an assumed increasing probability of toxicity/

adverse events, thus, decisions about which dose to recommend for further research was based upon safety outcomes. Today the probabilities may or may not be assumed to be increasing and decisions about which dose to recommend may depend upon other outcome measures. If doses/agents under investigation are assumed safe overall the study specific objective may not even refer to safety however we must still monitor for the unexpected. For example the main objective could be to identify the optimal biologic dose among a range of predefined dose levels which is defined as the dose at the beginning of the plateau of the dose efficacy curve [13]. Thus, conditional on safety, other endpoints may be the main summary endpoint that is used to determine which dose(s)/dose combination(s) to carry forward. These main measures need to be pre-defined in the protocol. Examples include an early measure of efficacy (clinical response); pharmacokinetic/pharmacodynamics (PK/PD); targets (immune response [14], biologic response that satisfies a threshold); and more.

2.1.2. Optimal or ideal dose/dose combination (optimum dose)—In dose finding studies, in order to find an optimum dose one must define the conditions that characterize the optimum dose, including the associated target probability that the study hopes to ascertain. The target probability is dependent upon the specific optimum dose criteria and is a fundamental parameter in the study design. A common example comes from research into cytotoxic agents where the goal is to find the highest effective dose that can be tolerated. Thus the optimum dose may be the maximum tolerated dose (MTD) defined as the highest dose that yields serious but reversible dose limiting toxicity (DLT) in no more than 33% of participants. Often omitted from a protocol is direct specification of the associated target probability which in traditional cytotoxic studies has varied from 16.7 – 33.3%. This is most common in the application of the frequently referred to algorithm-based ‘3+3 design’ where the safety target is the MTD which is defined by the end decision rule and varies depending upon the choice of cohort size and the DLT limit per cohort [15, 16]. Other examples of proper specification include: the optimal biologic dose (OBD) which in targeted therapy could be, the lowest dose where a target-mediated biologic pathway is altered by a prespecified amount in at least 50% of participants; or the optimal dose combination (ODC) which is the dose combination that produces DLTs in fewer than 25% of participants and attains the highest clinical response rate. Many new agents are not expected to produce serious AEs thus it is essential that the study safety target and associated probability be specified in the protocol. In model based designs sample size and development of the decision process is directly related to the target and threshold.

2.1.3. Assessment period—Along with characterization of the optimum dose, the minimum required window of observation for assessment conditions should be identified in the protocol. For example, observation of DLTs through the first 2 cycles or 6 weeks from start of treatment could be used to define the minimum timeframe for evaluation. Specification of which study participants this minimum observation period is mandatory needs to be stipulated in the protocol. The required timeline could be applied to all participants in a cohort or between the 1st and 2nd participants accrued to a dose level. How this constraint is used to guide decision rules is another component that should be described in the protocol. Is the information observed in the minimum timeline considered fixed or can it be updated to determine participant allocation? How and if conditions that occur outside of

the initial observation window, such as DLTs during cycle 4 of treatment, are used to direct participant allocation must be noted.

If participant allocation is based upon other summary measures then the assessment period associated with the other measures as well as how the information is incorporated into the decision process needs to be stated in the protocol. The timeline for assessment could be “response at 2 months”; PK/PD data in 1st cycle; or target met after three treatments. In determining the timeline for incorporation of these endpoints into the decision process one should consider and specify whether it is safety first then expansion after a minimal number of participants have been accrued; inclusion at the start of the decision process; or some other permutation.

2.1.4. Analysis populations—Inclusion and exclusion criteria are used to define the study population. If the study design uses information from study participants differently then it must be stated in the protocol. Do all participants contribute equally to the decision process or is their information used differently to guide the decision process? Are there other requirements that must be satisfied in order for a participant’s information to be used for some or all aspects of the trial? For example, differences by groups (good or poor prognosis [17], adult or pediatric, previously treated or untreated) may be weighted differently or be considered for only a subset of dose levels, or modified for “expansion”. A requirement could be to limit escalation decisions to the set of participants who a) received at least some predetermined amount of the planned doses during the assessment period (e.g., 70%, 85%, 95%, etc.) or b) experienced a DLT.

2.2. Decision process

2.2.1. Overall sample size—In early-phase research total sample size is an outcome measure and cannot be specified in advance. However, it is possible to decide upon a study maximum accrual or to estimate possible ranges. Estimates can be based upon operating characteristic results determined by simulation. See Cheung [18] for recommendations for Bayesian continual reassessment method designs.

2.2.2. Dose or dose combination levels—The range and starting dose or dose combinations considered for study are most often predefined and therefore should be listed in the protocol. For single agent escalation studies, $dose\ i, i = 1, \dots, k$ could be used to define the starting dose, $dose\ 1$, with k possible levels to consider, or $i = -2, -1, 1, \dots, k$ could be used to distinguish a range of lower doses from the starting dose that may be assessed if DLTs are observed early on in the decision process (Table 1). If multiple agents define a level then it should be stated whether one agent (or known regimen) is fixed and the other escalated; whether all are escalated or whether there is interest in studying a specified combined ordering. When multiple agents are being evaluated it is important to state where escalations can occur after the initial dose combination is accrued to and assessed. This may be predefined or could involve randomization (e.g., within a set of possible paths). Expansion from the single agent goal may require a modified naming convention for understanding allocation paths (e.g., instead of ‘ $dose\ i$ ’ it could be ‘ $arm\ i$ within a $zone\ j$ ’) (Table 2).

2.2.3. Cohort sample size—If the decision process is based upon data from cohorts then the target cohort size needs to be defined and importantly it must be stated whether there is a cohort minimum (1, 2, 3 etc.) and maximum size required for the decision process. The size requirements may differ if deviations are allowed during the dose finding process. For instance, a minimum of 1 participant per cohort must be assessed until the first DLT is observed and then at least 3 participants per cohort are required. Allowances must be spelled out in the protocol and include stipulations on when and how often a deviation is allowed. For instance, a participant is in the pipeline but the minimum window of observation has not yet been satisfied so will this participant be allowed into the study, and if so, at what dose? If allowed, is there a limit to the number of times it is permitted, such as, allowing only 1 deviation per dose?

2.2.4. Escalation/de-escalation method—All studies should clearly state what method of escalation/de-escalation will be employed in the study. For adaptive/model-based designs more description is needed so that everyone, be it reviewers or users, can understand the study approach. The study description should include references to the method employed; address basic model assumptions; the assumed prior distribution on model parameters; when the modeling process begins; whether the process allows the possibility to skip doses, and whether or not a safety endpoint (e.g., DLT) must be observed for the modeling component. See supplemental materials Table S2 for an example.

2.2.4.1. Design assumptions and modifications from published methods: Design operating characteristics are unique to the design being implemented, and dependent upon design goals and assumptions. The construction of study specific design operating characteristics depends upon the protocol assumptions and whether modifications from published methods were needed to address the trial goals. Modifications to any method should be delineated and may necessitate a more detailed and wider range of scenarios to include in the reporting of operating characteristics. The program code used to generate the operating characteristics should be made available either upon request or by placement in an area that is accessible via the internet.

2.2.5. Starting/stopping the trial—Trial participants may be allocated by different rules at the beginning of a trial compared to those that govern allocation decisions later in the trial. The reasons vary but include adjustment for rapid allocation through early dose levels or the requirement that a minimum number of accruals or events (DLTs) must be observed in order to estimate model parameters. If there is an initial allocation stage then it should be specified in the protocol. The allocation plan should include details on whether a minimum number of accruals or observed DLTs are required before the modeling process begins, or whether other conditions must be satisfied in order to determine participant allocation. It is equally important to describe the conditions that should be satisfied to halt accrual. This could be based upon reaching a maximum number of accruals to a dose level or a probability assessment that the dose recommendation will not change if a few more participants are accrued [19]. Once accrual is halted, what analysis method will be employed to determine which dose(s)/dose combination(s) warrant further study (i.e., what defines the MTD or OBD or ODC).

2.3. Design operating characteristics

2.3.1. Simulation studies—The operating characteristics of a proposed design provide the justification for the choice of design and provide reviewers with information to determine if the study design is sufficient to achieve the objectives of the clinical trial. They are also used in the design phase to evaluate the appropriateness of design constraints such as stopping rules or allowance of skipped doses. Adaptive designs are evaluated through simulations (recommend 1000 trials) and are based upon choosing a maximum sample size (e.g., $n=45$) or by assessing a variety of sizes (e.g., $n=15, 30, 45, 60$). The range of sample sizes to include will depend upon many factors including the number of dose/dose combinations and the design method. To start the process, we recommend a sample size of 5 times the number of dose combinations under investigation, up to a maximum amount that could be attained within the constraints of the institution/budget/expected time line). It is crucial that the simulations cover a wide range of scenarios including anticipated or ideal scenarios; worst case (a lot of toxicity) scenario; and best case (none to little toxicity) scenario. Reported simulation results should contain the average number or percent of participants treated at each dose or dose combination in the range (including low doses, high doses, at or near the ‘target’ dose (MTD, OBD, ODC, etc.); the average (percentiles) trial size; the percent of trials that stop at the 1st dose level, the percent of participants with a DLT; and the percent of participants with other protocol defined endpoints [20]. Simulation results from a study (NCT02419560) designed to find the optimal dose combination (ODC), defined by an acceptable toxicity profile (25% DLT rate) and the highest clinical response rate, are shown in Table 3.

2.3.2. Dose escalation decision paths—The adaptive design should recommend reasonable decisions during the conduct of the clinical trial based on the observed endpoints. To demonstrate that the model will not lead to decisions that are unethical or in disagreement with clinical practice, a decision path(s) based upon the trial design could be included in the protocol. This can be useful to investigators and reviewers to see how the model will behave at the beginning of the study. In this situation the protocol should present an illustration of on-study dose allocation based on several scenarios for the first cohorts. An illustration of dose escalation decision paths for a single agent fixed cohort design where no DLTs have been observed in the prior cohorts is given in (Table 4).

2.3.3. Software/programs—Correct design and implementation of any adaptive design is dependent upon the statistical programs used to generate the decision rules. Availability of the programs for both the design stage (to estimate sample size and generate operating characteristics) and for direct protocol implementation (to determine the decision rule) is required. Often executable programs (e.g., R-package executable [21–23]) are made available through the internet or upon request. It is important to note that whatever code is used that, if requested, you need to make the code, not just the executable, available to reviewing entities (regulatory agency, pharmaceutical) that may need to verify the implementation of the method or run additional scenarios.

3. Clinical trial implementation

It takes a lot of time and effort to design a clinical trial to satisfy requirements from the principal investigator (PI), co-investigators, statisticians, reviewing bodies and sponsors. During the design process many changes occur and the final design may be very different from the one proposed during the initial design stage. There are many reasons for this. The PI goals may be very different from the sponsor goals, such that the best design or hypothesized outcomes may not be of interest to the sponsor. Review and assessment of design operating characteristics may reveal that the study cannot be accomplished with the limited participant pool or within a specified time frame, both of which could be related to limited funding. Modifications to design assumptions may be required by reviewers or based upon feasibility assessment. With that noted, equal attention must be applied to the actual implementation of the final approved protocol. If not, many roadblocks may still be encountered during the start-up and accrual portion of the clinical trial.

3.1. Data capture and data flow

3.1.1. Clinical research management system—Identification of the management system that will be used to capture participant accrual and outcome information, whether it is internal to the sponsoring institution or external, is required. Once identified many areas related to application should be considered. If access to the local system is for local use only (i.e., a single site trial) then system permission and training can be handled at the local site. If the study involves many sites with each site needing access then there may be many system permissions/hurdles that need to be addressed to allow access to the system. These can be time consuming and require more time to develop standard operating procedures and train users on how to use the system. If a multi-site trial is planned it is recommended that the trial start at the PI institution first.

With adaptive designs it is important that the system can be set-up to capture current accruals and their endpoint data (e.g., DLTs) in a timely manner. This could be accomplished by requiring information from a case report form that **MUST** be entered into the system at the end of the day of a participant visit that provides the timeframe of the assessment and whether or not a DLT was reported. Not only is information needed about current accruals it is important to capture information on potential accruals that are in the pipeline. It is recommended that potential participant accruals be entered into the system at the time of consent so that the number of potential candidates coming into the trial is always known. Once eligibility is confirmed and the participant is eligible for the study, information about the estimated start of treatment date should be captured. Another option that is useful for multisite trials is to allow potential slots to be tagged with a time limit set to consent and verify patient eligibility. When slots are full, stand-by queues can be used to claim the next open spot based upon time in the queue.

Adaptive designs require participant specific tracking to guide decisions about the recommended allocation for the next participant. Thus, one must have the ability to export data into a summary report to verify all required data have been captured for the participants already put on study, as well as the capability to assign the next recommended dose/dose combination to the future participant.

3.1.2. Responsibility of personnel—There are many system and data capture requirements that must be adhered to in a timely manner so that the adaptive design can be applied properly and as designed. Who has main responsibility for each component should be reviewed at study start-up. Too often when things go wrong the reason given is that it was ‘assumed’ that someone else was supposed to do that duty. Clearly, the PI has ultimate responsibility for the correct conduct of the study however it is good team work that makes a clinical trial run well. In general, it should be discussed who is responsible for participant specific DLT and “response” assessment (e.g., PI and research nurses, local PIs), and who is responsible for DLT and “response” data capture and submission into the system (e.g., clinical research coordinator, data manager). How each responsibility is handled at other participating sites must be addressed and if possible consideration should be given to inclusion of a protocol manager to help track participant eligibility and data capture over all possible participating sites.

Study statisticians are essential for the development of adaptive-based designs, and they should be equally involved in the management and implementation of the design. They are a part of the study team. As such it is important to determine if the study statistician(s) or designee is responsible for the generation of summary reports and determination of next dose/dose combination to be accrued to. Most often, study statisticians have read only access to the data in the system. However, for adaptive/model-based designs it may be necessary for them to have access to the system to input the ‘next’ allocation. If allocation determination is directly programmed then the clinical research management system must have the capability to perform the function.

Who has responsibility for developing and maintaining study specific programs both at the design stage and the implementation stage has to be addressed early on, especially at institutions just beginning to use these type of designs. Often the study statistician(s) alone or in conjunction with a statistical programmer develop and verify the code. Once the study is started it should be decided, how and where the code will be maintained, and who has access/permission to run the code to make the next dose/dose combination recommendation.

3.2. Safety assessment reviews and the decision process

It is the responsibility of the PI and study team to have a process in place to review escalation decisions. This can be accomplished in many ways and can include email, conference calls or team meetings, either routine or as needed. An example of a pre-planned assessment could occur once a cohort of participants has been completed and the minimum requirements have been met (e.g. the minimum required number of evaluable participant has reached the end of the assessment period). A meeting between the investigators and the sponsor could take place to jointly decide the next dose level to allocate to the next cohort. During this meeting, data from the currently evaluated cohort of participants should be carefully reviewed, along with new data from previously treated participants who are still ongoing in the study. The statistician in charge of the study should provide a statistical summary report and inform participants to the next dose level proposed by the model. As the meeting takes place a few weeks after the end of the assessment period of the last participant included in the cohort, the statistician could use the observed data of all evaluable

participants to update the model. However, to avoid potential disagreement(s) on endpoints assessment during the meeting (e.g. disagreement on a DLT), the statistician should be prepared to provide the next dose allocation to be accrued to for any possible scenarios of outcomes for the participants included in the cohort evaluated during the meeting. For example, in a clinical trial with a single endpoint (DLT) and 2 participants accrued at each cohort, the statistician should provide all possible dose allocation scenarios according to the number of DLTs observed in the currently evaluated cohort, namely 0, 1, or 2 (see Figure 1).

Protocol specific deviations may result in a modification to the recommended path. A discussion between all meeting members should take place and the final decision should be taken after having reviewed all relevant data. An agreement on the next recommended dose is required between all members to proceed to the accrual of the new cohort of study participants.

3.3. Audit trail

Regardless of which approach is used to determine patient allocation a written audit trail should be kept. Program outputs, flow charts, outcomes of the decision process and meeting minutes if applicable, should be documented and stored in a file shared location for access during the accrual period and locked at study closure.

4. Summary

For several years now, investigators have asserted that more innovative approaches are needed to address early-phase research questions. In response, statisticians have continued to develop more flexible, efficient and relevant study designs, however, the application of these designs has been limited. There have been several recent published trials that have successfully implemented innovative designs into complex dose-finding studies. The approach of Braun et al. [24] was used to design a dose and schedule finding study of de Lima et al. [25]. The method of Conaway, Dunbar and Peddada [26] was implemented in a Phase I trial investigating induction therapy with VELCADE and Vorinostat in patients with surgically resectable non-small cell lung cancer (NSCLC) [27]. A recent editorial by Mandrekar [28] described the use of the method of Ivanova and Wang [29] in a Phase I study of neratinib in combination with temsirolimus in patients with human epidermal growth factor receptor 2-dependent and other solid tumors [30]. Our goal is to see wider acceptance and application of these designs [31]. We and others [32] have provided researchers with guidelines to include in protocols to improve understanding, acceptance and approval of studies using these novel designs, as well as guidelines to efficiently execute the design. Suggestions are summarized in a checklist in Table 5. We have emphasized issues related to model-based designs, however, the recommendations should be applied to algorithm-based designs as well, especially when deviations from the algorithm-based are allowed in the protocol.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

Research reported in this publication was supported by the National Cancer Institute of the National Institutes of Health under award numbers R01CA142859 (GRP), P30 CA044579 (GRP) & K25CA181638 (NAW).

References

1. Paoletti X, O'Quigley J, Maccario J. Design efficiency in dose-finding studies. *Computational Statistics and Data Analysis*. 2004; 45:197–214.
2. Rogatko A, Schoeneck D, Jonas W, Tighiouart M, Khuri FR, Porter A. Translation of innovative designs into phase I trials. *Journal of Clinical Oncology*. 2007; 25:4982–86. [PubMed: 17971597]
3. Iasonos A, Wilton A, Riedel E, Seshan V, Spriggs D. A comprehensive comparison of the continual reassessment method to the standard 3+3 dose escalation scheme in phase 1 dose-finding studies. *Clinical Trials*. 2008; 5:465–77. [PubMed: 18827039]
4. Le Tourneau C, Lee JJ, Siu LL. Dose escalation methods in phase I cancer clinical trials. *Journal of the National Cancer Institute*. 2009; 101:708–20. [PubMed: 19436029]
5. O'Quigley J, Conaway M. Continual reassessment and related dose-finding designs. *Statistical Science*. 2010; 25(2):202–216. [PubMed: 21731191]
6. O'Quigley J, Conaway M. Extended model based designs for more complex dose-finding studies. *Statistics in Medicine*. 2011; 30:2062–2069. [PubMed: 21351287]
7. Jaki T, Clive S, Weir CJ. Principles of dose finding studies in cancer: a comparison of trial designs. *Cancer Chemotherapy and Pharmacology*. 2013; 71:1107–14. [PubMed: 23299793]
8. Wages NA, Conaway MR, O'Quigley J. Performance of two-stage continual reassessment method relative to an optimal benchmark. *Clinical Trials*. 2013; 10:862–875. [PubMed: 24085776]
9. Ji Y, Wang SJ. Modified toxicity probability interval design: a safer and more reliable method than the 3+3 design for practical phase I trials. *Journal of Clinical Oncology*. 2013; 31:1785–91. [PubMed: 23569307]
10. Iasonos A, O'Quigley J. Adaptive dose-finding studies: a review of model-guided phase I clinical trials. *Journal of Clinical Oncology*. 2014; 32:2505–11. [PubMed: 24982451]
11. Liu S, Yuan Y. Bayesian optimal interval designs for phase I clinical trials. *Journal of the Royal Statistical Society – Series C (Applied Statistics)*. 2015; 64:507–23.
12. Paoletti, X.; Ezzalfani, M.; Le Tourneau, C. Statistical controversies in clinical research: requiem for the 3 + 3 design for phase I trials. *Annals of Oncology*. 2015. [epub ahead of print] <http://dx.doi.org/10.1093/annonc/mdv266>
13. Zang Y, Lee JJ, Yuan Y. Adaptive designs for identifying optimal biological dose for molecularly targeted agents. *Clinical Trials*. 2014; 11:319–327. [PubMed: 24844841]
14. Wages NA, Slingluff CL Jr, Petroni GR. A phase I/II adaptive design to determine the optimal treatment regimen from a set of combination immunotherapies in high-risk melanoma. *Contemporary Clinical Trials*. 2015; 41:172–79. [PubMed: 25638752]
15. Reiner E, Paoletti X, O'Quigley J. Operating characteristics of the standard phase I clinical trial design. *Computational Statistics and Data Analysis*. 1999; 30:303–315.
16. Kelly WK, Halabi S. *Oncology Clinical Trials*. DemoMedical. 2010; Chapter 8
17. Wages NA, Read PW, Petroni GR. A Phase I/II adaptive design for heterogeneous groups. *Pharmaceutical Statistics*. 2015; 14:302–10. [PubMed: 25962576]
18. Cheung YK. Sample size formulae for the Bayesian continual reassessment method. *Clinical Trials*. 2013; 10:852–61. [PubMed: 23965547]
19. O'Quigley J, Reiner E. A stopping rule for the continual reassessment method. *Biometrika*. 1998; 85:741–748.
20. Iasonos A, Gönen M, Bosl GJ. Scientific Review of Phase I Protocols With Novel Dose-Escalation Designs: How Much Information Is Needed? *Journal of Clinical Oncology*. 2015; 33:2221–5. [PubMed: 25940721]
21. Cheung, K. *dfcrm*: dose-finding by the continual reassessment method. R package version 0.2-2. 2013. Available at: <http://CRAN.R-project.org/package=dfcrm>

22. Sweeting M, Mander A, Sabin T. bcrm: Bayesian continual reassessment method designs for phase I dose-finding trials. *Journal of Statistical Software*. 2013; 54:1–26.
23. Wages NA, Varhegyi N. pocrm: an R package for phase I trials of combinations of agents. *Computer Methods and Programs in Biomedicine*. 2013; 112:211–18. [PubMed: 23871691]
24. Braun TM, Thall PF, Nguyen H, de Lima M. Simultaneously optimizing dose and schedule of a new cytotoxic agent. *Clinical Trials*. 2007; 4:113–24. [PubMed: 17456511]
25. de Lima M, Giralt S, Thall PF, et al. Maintenance therapy with low dose azacitidine after allogeneic hematopoietic stem cell transplantation for recurrent acute myelogenous leukemia or myelodysplastic syndrome: a dose and schedule finding study. *Cancer*. 2010; 116:5420–31. [PubMed: 20672358]
26. Conaway MR, Dunbar S, Peddada SD. Designs for single- or multiple-agent phase I trials. *Biometrics*. 2004; 60:661–69. [PubMed: 15339288]
27. Jones DR, Moskaluk CA, Gillenwater HH, Petroni GR, Burks SG, Philips J, Rehm PK, Olazagasti J, Kozower BD, Bao Y. Phase I Trial of Induction Histone Deacetylase and Proteasome Inhibition Followed by Surgery in Non-Small-Cell Lung Cancer. *Journal of Thoracic Oncology*. 2012; 7:1683–90. [PubMed: 23059775]
28. Mandrekar SJ. Dose-finding trial designs for combination therapies in oncology. *Journal of Clinical Oncology*. 2014; 32:65–67. [PubMed: 24323038]
29. Ivanova A, Wang K. A non-parametric approach to the design and analysis of two-dimensional dose-finding trials. *Statistics in Medicine*. 2004; 23:1861–70. [PubMed: 15195320]
30. Gandhi L, Bahleda R, Tolaney SM, Kwak EL, Cleary JM, Pandya SS, Hollebecque A, Abbas R, Ananthakrishnan R, Berkenblit A, Krygowski M, Liang Y, Turnbull KW, Shapiro GI, Soria JC. Phase I study of neratinib in combination with temsirolimus in patients with human epidermal growth factor receptor 2-dependent and other solid tumors. *Journal of Clinical Oncology*. 2014; 32:68–75. [PubMed: 24323026]
31. Wages NA, Conaway MR, Slingsluff CL Jr, Williams ME, Portell CA, Hwu P, Petroni GR. Recent developments in the implementation of novel designs for early-phase combination studies. *Annals of Oncology*. 2015; 26:1036–37. [PubMed: 25697216]
32. Lorch U, O’Kane M, Taubel J. Three steps to writing adaptive study protocols in the early phase clinical development of new medicines. *BMC Medical Research Methodology*. 2014; 14:84.doi: 10.1186/1471-2288-14-84 [PubMed: 24980283]
33. O’Quigley J, Shen LZ. Continual reassessment method: a likelihood approach. *Biometrics*. 1996; 52:673–684. [PubMed: 8672707]

FIGURE 1.

Dose allocation for cohort 3

One agent, Single endpoint (DLT)

Next recommended dose is the dose having a mean posterior probability of DLT closest to 30%

2 participants per cohort until a DLT is observed, then 3 participants per cohort dose skipping not allowed

TABLE 1

Dose escalation schedule

Dose level	Agent (mg twice daily)
level -2	50
level -1	75
level 1 (starting dose)	100
level 2	200
level 3	300
level 4	400
level 5	500
level 6	600

TABLE 2

Zone and Arm Designation by Combination

Agent A (mg per day)	400	Zone 2/Arm C	Zone 3/Arm E	Zone 4/Arm F
	200	Zone 1/Arm A	Zone 2/Arm B	Zone 3/Arm D
		280	420	560
		Agent B (mg per day)		

TABLE 3

Two agents, Multiple endpoints (DLT, Response)
 Defined by an acceptable toxicity profile (25% DLT rate) and the highest clinical response rate
 Stop when an 11th participant recommended to a dose combination that already has 10 participants
 Maximum total accrual set at 28, 1000 simulations

Scenario		(true %DLT, true %Response) % combination recommended Avg # pts treated on combination	Avg Size, %-tiles	% stop	% DLT	% Rsp
	Agent A (mg per day)	Agent B (mg per day)				
		280	420	560		
1 "Ideal"	400	(0.07, 0.55)	(0.15, 0.70)	(0.25, 0.95)	19.1, 25 th = 17	0.2 16.4 74.2
		0.12	0.12	0.67	50 th = 18	
	200	1.72	2.30	7.46	75 th = 20	
		(0.05, 0.40)	(0.07, 0.55)	(0.15, 0.80)	90 th = 24 95 th = 27	
2 "Worst"	400	(0.25, 0.25)	(0.30, 0.40)	(0.40, 0.50)	18.1, 25 th = 16	7.0 25.6 38.3
		0.06	0.08	0.06	50 th = 18	
	200	2.54	2.18	1.63	75 th = 21	
		(0.20, 0.25)	(0.25, 0.40)	(0.30, 0.50)	90 th = 25 95 th = 27	
3 Best Guess	400	(0.12, 0.55)	(0.18, 0.80)	(0.25, 0.95)	19.8, 25 th = 17	0.6 18.4 75.0
		0.02	0.10	0.58	50 th = 18	
	200	1.97	2.77	6.72	75 th = 22	
		(0.10, 0.40)	(0.15, 0.55)	(0.20, 0.80)	90 th = 26 95 th = 28	

Scenario		(true %DLT, true %Response) Avg # pts treated on combination			Avg Size, %-tiles	% stop	% DLT	% Rsp
	Agent A (mg per day)	Agent B (mg per day)						
		280	420	560				
4 "Plateau"	400	(0.25, 0.35)	(0.30, 0.50)	(0.40, 0.50)	18.2, 25 th = 17 50 th = 18	9.0	25.7	44.8
		0.06	0.08	0.03				
	200	2.55	2.19	1.46	75 th = 21 90 th = 25 95 th = 28			
		(0.20, 0.35)	(0.25, 0.50)	(0.30, 0.50)				
5 "Toxic"	400	0.20	0.38	0.15				
		3.83	5.47	2.92				
	200	(0.30, 0.40)	(0.40, 0.50)	(0.50, 0.75)	18.4, 25 th = 16 50 th = 19	11.8	32.9	52.9
		0.09	0.09	0.03				
6 "Very Toxic"	400	2.94	2.20	1.10	75 th = 23 90 th = 26 95 th = 28			
		(0.25, 0.40)	(0.30, 0.50)	(0.40, 0.75)				
	200	0.26	0.25	0.17				
		4.41	4.59	3.12				
	400	(0.75, 0.40)	(0.85, 0.50)	(0.95, 0.75)	4.3, 25 th = 2 50 th = 2	100.00	74.0	42.5
		0.00	0.00	0.00				
	200	0.95	0.09	0.00	75 th = 6 90 th = 9 95 th = 11			
		(0.70, 0.40)	(0.75, 0.50)	(0.85, 0.75)				
		0.00	0.00	0.00				
		1.98	1.07	0.17				

Notation:

Average (avg); percentiles (%-tiles); Response (Rsp); Average trial sample size (avg size); Percent of trials stopped early for safety (% stop); Percent of DLTs (% DLT); Percent of responses (% Rsp)

TABLE 4

One agent, Single endpoint (DLT)

Next recommended dose is the dose having a mean posterior probability of DLT closest to 30%

3 participants by cohort, dose skipping allowed (limited to 200mg)

Cohort	Dose (mg)	Number of observed DLT	Next dose level proposed (mg)
1	100	0	300
		1	200
		2	50
		3	50
2 (0 DLT in cohort 1)	300	0	500
		1	500
		2	300
		3	100
3 (0 DLT in cohort 1 and 2)	500	0	700
		1	600
		2	500
		3	400
4 (0 DLT in cohort 1, 2 and 3)	700	0	900
		1	800
		2	700
		3	600

TABLE 5

Adaptive study design methods definition/identification checklist

Protocol		
<i>Study objective</i>	Endpoints	
	Optimum dose definition	
	Assessment period	
	Analysis population	
<i>Decision/allocation process</i>	Dose levels	
	Cohort size(s)	
	Escalation/de-escalation method	
	Starting/stopping rules	
<i>Operating characteristics</i>	Parameter(s) definitions	
	Evaluation metric(s)	
	Simulation scenarios	
	Initial decision paths	
Simulation outcomes checklist		
<i>Overall</i>	'Target'	
	Threshold	
	Maximum sample size	
	Number of simulations	
	Stopping rule	
<i>Scenarios</i>	Include a wide range and specify assumed true probabilities: Anticipated or ideal	
	Worst	
	Best	
	Others	
<i>Dose level specific results (within each scenario)</i>	Average number or percent of participants treated at each dose	
	Percent dose recommended as the 'target' dose	
<i>Scenario results</i>	Average (percentiles) of overall trial size	
	Percent of times trial stopped at 1 st dose level	
	Percent of participants with a DLT	
	Percent of participants with other stopping/decision criteria	