A Case Study for Radiation Therapy Dose Finding Utilizing Bayesian Sequential Trial Design

Author's Details: [(1)]Fuyu Song and [(2)(3)]Shein-Chung Chow

[1]Peking University Clinical Research Institute, Peking University Health Science Center, Beijing, China

[2]Duke University School of Medicine, Durham, North Carolina, USA

[3]Friendship Hospital, Beijing Capital Medical University, Beijing, China

Corresponding Author: Fuyu Song, M.S. Peking University Clinical Research Institute, Peking University Health Science Center, No.38 Xueyuan Road, Haidian District, Beijing,100191, China.

Abstract

Dose escalation trials for identifying the maximum tolerable dose (MTD) is commonly considered in phase 1 cancer clinical research. For this purpose, an algorithm-based design such as a standard escalation design with traditional escalation rule (TER) and a model-based design such as the method of continued reassessment method (CRM) under a well-established dose toxicity model are commonly employed. In practice, relative merits and limitations of these two different types of designs are not fully understood. Besides, most dose escalation studies do not provide scientific justification for sample size and design selection. In this article, the validity and efficiency of these two different types of study designs are evaluated based on the criteria of the number of subjects expected, the number of DLT expected, the probability of correctly achieving the MTD, and the probability of overdosing. A case study regarding a radiation therapy for treatment of certain solid tumors is discussed to illustrate the criteria for design selection..

Key Words: Algorithm-based design; Model-based design; "3+3" TER design; CRM design; Bayesian sequential design.

1. INTRODUCTION

In clinical research and development, dose response studies play an important role for identifying the minimum effective dose (MED) and maximum tolerable dose (MTD) in early phase of clinical development. The information regarding MED and MTD are useful for later phase clinical development.

For dose response trials, ICH E4 (1994) recommends the use of several study designs. These study designs include randomized parallel dose-response designs, crossover dose-response design, forced titration design (dose escalation design), optimal titration design, and placebo-controlled titration to endpoint. Detailed discussion of these study designs can be found in Ting (xxx). For analysis of dose response trials, analysis of variance (ANOVA) or covariance (ANCOVA) with appropriate contrast (Spriet and Dupin-Spriet 1996) is commonly employed. Cheng et al. (2006) proposed the use of slope approach based on the evaluation of slope between adjacent doses or with respect to the origin (zero dose). Cheng et al. (2006) claimed that slope approach is able to characterize different kind of dose response curve (both linear or non-linear) within the dose range under study.

In cancer research, the primary objectives of dose response trials are to determine (i) is there any evidence of the drug effect? (2) what is the nature of the dose-response? and (3) what is the optimal dose? Following the principles that (1) there are less patients to be exposed to the toxicity and (2) there are more patients to be treated at potential efficacious dose levels. For this purpose, algorithm-based and/or model-based dose escalation trial with limited number of subjects are often considered for identify MTD which is often considered as optimal dose for later phase of clinical development.

In this article, our emphasis will be placed on the discussion and comparison of algorithm-based design such as a traditional "3+3" dose escalation design and model-based dose escalation design such as a design utilizing continued re-assessment method (CRM) in conjunction with a Bayesian approach. For illustration

purpose, a case study regarding a radiation therapy dose finding trial is presented.

2. DOSE ESCALATION TRIAL DESIGN

2.1 Algorithm-Based Trial Design

The most commonly employed study design in cancer dose escalation trials is probably the design utilizing traditional escalation rule (TER). Among these TER designs, the "3+3" TER design is the most popular trial design for identifying the MTD based on limiting dose toxicity (DLT) of the test treatment under investigation. DLT is referred to as unacceptable or unmanageable safety profile which is pre-defined by some criteria such as Grade 3 or greater hematological toxicity according to the US National Cancer Institute's (NIH) Common Toxicity Criteria (CTC). As a result, MTD is the highest possible but still tolerable dose with respect to some pre-specified DLT. The "3+3" TER trial design is to enter three patients at a new dose level and then enter another three patients when a DLT is observed. The assessment of the six patients is then performed to determine whether the trial should be stopped at the level or to escalate to the next dose level.

TER design is considered a standard algorithm-based dose escalation design. TER design is simple and easy to implement. However, there are some drawbacks. These drawbacks include, but are not limited to, that (1) there are no room for dose de-escalation, (2) there is no sample size justification, (3) it does not require further analysis of data, (4) there is no objective estimation of MTD with statistical model, and (5) there is no sampling error and consequently no confidence interval. It should be noted that, in addition to the standard "3+3" TER which does not allow for dose-de-escalation, there is another type of "3+3" TER design that allows dose de-escalation if two of three patients have DLT. This trial design is referred to as strict traditional escalation rule (STER) trial design. The typical "3+3" TER design and can be generalized to the "m+n" TER design with and without dose de-escalation, where $m + n = 6$.

Following similar idea of dose escalation and de-escalation, Storer (1989, 1993, 2001) proposed single-stage up-and-down phase I designs (namely design A, design B, and design D as described in his papers), two-stage up-and-down phase I designs (i.e., design BD), and accelerated titration designs. As an example, for design A, we start with a group of three patients, who are treated at the lowest dose level. At the second step, if no pre-specified DLT is observed in all three patients, then the dose for the next group of three patients is escalated to the next higher dose level. Otherwise, the next group of three patients is treated at the same dose level. In step 3, the dose of the next group of three patients is escalated to the next higher dose level if the pre-specified DLT is observed at most in one patient of the six patients, otherwise, the trial stops. For step 4, we repeat steps 2 and 3 with two consecutive groups of three patients until the trial stops.

2.2 Model-Based Study Design

The model-based study design is referred to a study design utilizing the continued reassessment method (CRM) under the assumption that there is well established dose response relationship between dose and response (mainly toxicity), i.e., $\text{Tox} = \text{y} = f(\text{dose})$. The dose-response relationship is continually reassessed based on accumulative data collected from the previous subjects. The next patient who enters the trial is then assigned to the dose level with potential MTD.. In other words, based on $\text{Tox} = f(\text{dose})$, we can solve dose for the corresponding toxicity, i.e., $\text{dose} = g(\text{Tox})$. Now if Tox = DLT, then the corresponding dose is MTD.

Basically, the model-based CRM involves (1) dose toxicity modeling, (2) dose level selection, (3) reassessment of model parameters (or estimation of MTD), and (4) assignment of next patient

For does toxicity modeling, it is often assumed that (1) there is monotonic relationship between dose and toxicity and (2) the biologically inactive dose is lower than the active dose, which is in turn lower than the toxic dose. The commonly considered dose toxicity model in cancer research is either a linear (for continuous outcomes) or logistic model (for binary responses). As an example, we may consider a logistic model as follows

$$p(x) = [1 + b\exp(-ax)]^{-1},$$

where $p(x)$ is the probability of toxicity associated with dose x, and a and b are positive parameters to be determined. Let $\theta = p(x)$. Now solve x.for θ. Then, the $x = \text{MTD}$ can be expressed as

$$MTD = \frac{1}{a}\ln\left(\frac{b\theta}{1-\theta}\right).$$

where θ is the probability of observing DLT at MTD. As indicated by Crowley (2001), for an aggressive tumor and a transient and non-life-threatening DLT, θ could be as high as 0.5, while for persistent DLT and less aggressive tumors, θ could be as low as 0.1 to 0.25. A commonly used value for θ is somewhere between 0 and 1/3=0.33

For dose level selection, the following general principles are helpful. First, the selected dose should be low enough to avoid severe toxicity. At the same time, it should be high enough for observing some activity or potential efficacy in humans. In practice, the commonly used starting dose is the dose at which 10% mortality (i.e., LD_{10}) occurs in mice. The subsequent dose levels are usually selected based on the following multiplicative set

$$x_i = f_{i-1}x_{i-1},\ i = 1, 2, \ldots, K,$$

where f_i is called dose escalation factor. In practice, Fibonacci sequence (1, 1, 2, 3, 5, 8, 13 …) is often considered.

For reassessment of model parameters, the key is to estimate the parameter a in the response mode. An initial assumption or prior about the parameter is necessary in order to assign patients to the dose level based on the toxicity relationship. The estimate of a is continually updated based on cumulative data observed from the trial. The estimation method could be a frequentist (e.g., maximum likelihood estimate or least square estimate) or Bayesian approach, which requires a prior distribution about the parameter. The Bayesian approach provides posterior distribution and predictive probabilities of MTD.

The updated dose-toxicity model is usually used to choose the dose level for the next patient. In other words, the next patient enrolled in the trial is assigned to the current estimated MTD based on dose-response model. Assignment of patient to the most updated MTD leads to majority of the patients assigned to the dose levels *near* MTD, which allows a more precise estimate of MTD with a minimum number of patients. In practice, this assignment is subject to safety constraints such as limited *dose jump* and delayed response.

Alternatively, Chang and Chow (2005) proposed a hybrid Bayesian adaptive design by developing a utility function that incorporate anything that would affect the outcomes or decision making such as a treatment, a withdrawal of a treatment arm, a protocol amendment, stopping the trial, etc.

3. CRITERIA FOR DESIGN SELECTION

In clinical research, criteria for selecting an appropriate design among available designs are based on either a power approach or a precision analysis. For the power approach, for a fixed sample size, the design with highest power is considered the most appropriate design. On the other hand, for a fixed power, the study design have best precision is considered the most appropriate design. These criteria for design selection, however, may not be appropriate for design selection

between an algorithm-based design and a model-based design.

For most dose escalation studies, not only no sample size justification is provided in the study protocol, but also *no* justification regarding the selection of study design between an algorithm-based TER design or a model-based CRM design is provided. In addition, n*o* information regarding the possible safety issues such as how many subjects are expected to experience DLT and the possibility of overdose are provided. As a result, it is not clear what the most appropriate criteria for design selection between an algorithm-based design and a method-based design is. Based on informal communication with FDA statistical/medical reviewers, it is suggested that one of a few of the following criteria be considered for selecting the most appropriate study design for dose escalation trials (Chow, 2009)..

Number of patients expected – In cancer trials, the number of subjects available is usually very limited. For a given study design, the number of subjects expected for reaching the MTD is always a concern. If a smaller number of subjects under a given design can lead to the MTD, the principal investigator would select the design over the other design which requires more subjects for achieving MTD. Thus, the number of subjects expected for achieving the MTD has become one of criteria for design selection.

Number of DLT expected – In the interest of not to have too many subjects exposed to the toxic treatment, the number of DLT expected has become an important indication for design selection. With a similar number of subjects expected, the PI would select a study design with a less number of DLT expected.

Probability of correctly achieving the MTD – Despite of the consideration of the number of subjects expected and the number of DLT expected, the probability of correctly achieving the MTD is probably the most concern of the principal investigator.

Probability of overdosing – Another consideration for design selection is on the safety of possible overdosing. The possibility of overdosing could occur in both the algorithm-based design and the model-based design. Thus, the probability of overdosing provides useful information for design selection, especially for those relatively toxic treatments.

Remarks –In practice, there exists no closed forms evaluation of the number of subjects expected, the number of DLT expected, and the probability of correctly achieving the MTD under either an algorithm-based design or a model-design. Thus, it is suggested that a clinical trial simulation be conducted under the algorithm-based design and the model-based design of interest.

4. CASE STUDY

4.1 Radiation Therapy

For illustration purpose, consider the example concerning radiation therapy for treatment of certain solid tumor. A pharmaceutical company is interested in conducting a dose escalation study for identifying the MTD of a test treatment under investigation. The identified MTD will be considered as the optimal dose for subsequent clinical trials conducted for later phase clinical development. Since the test treatment is very toxic, the principal investigator wishes to have a study design with small size cohort for lower dose levels. Ideally, it is to minimize the number of patients at lower dose groups and have majority patients near the MTD (ideally, the last two dose cohorts under study). The principal investigator also wishes to have flexibility for dose de-escalation. If a model-based CRM design is used, there should be a limited dose jump. In addition, the principal investigator would like to have higher probability of reaching the MTD and at the same time have lower probability of overdosing.

To account for the principal investigator's wishing list, the "3+3" TER design (does not allow dose de-escalation), the "3+3" STER (allows dose de-escalation), and the CRM in conjunction with a Bayesian approach are considered. Based on a pilot study, the DLT rate at MTD is assumed to be 1/3=0.33. The initial dose is selected for 0.5 mCi/kg with six dose levels in a dose range from 0.5 mCi/kg to 4.5 mCi/kg. A modified Fibonacci sequence dose escalation factor is considered. That is, dose levels are 0.5, 1, 1.6, 2.5, 3.5, and 4.7. Note that for CRM, the dose level close to the estimated MTD (based on the updated toxicity model) will be used for assignment of the next subject. The three candidate designs seem reasonable to the principal investigator and yet there are no quantitative criteria for design selection.

4.2 Clinical Trial Simulation

As indicated earlier, there exist no closed forms for evaluation of the number of subjects expected, the number of DLT expected, and the probability of correctly achieving the MTD. Thus, a clinical trial simulation with 5,000 runs was conducted under the following assumptions

(1) Initial dose for the test treatment is assumed to be 0.5 mCi/kg;
(2) Dose range is from 0.5 mCi/kg to 4.5 mCi/kg;
(3) The number of dose levels is 6.
(4) Maximum dose de-escalation allowed is 1 (for STER);
(5) DLT rate at MTD is assumed to be 1/3=33%;
(6) Logistic toxicity model given in Section 2.2 is assumed for the CRM;
(7) For the CRM in conjunction with a Bayesian approach, a simple uniform prior is considered.
(8) No dose jump is allowed.

Under the above assumptions, the clinical trial simulation results are summarized in Table 1. As it can be seen from Table 1 that CRM in conjunction with a Bayesian approach gave the smallest number of subjects expected (N=13.82), while the "3+3" TER design has the smallest number of DLT expected (2.8). All of the study designs seem to underestimate the MTD. The CRM in conjunction with a Bayesian approach has the highest probability of correctly achieving the MTD (69.6%). As a result, the CRM in conjunction with a Bayesian approach is recommended for the radiation therapy dose escalation study.

5. CONCLUDING REMARKS

In cancer phase 1 dose escalation studies, the "3+3" TER design and CRM design (with or without Bayesian approach) are commonly employed. In the past, no justifications for sample size and/or design selection are provided for validity and efficiency of the selected study design. As illustrated in the case study of radiation therapy, it is suggested that sample size and/or design selection should be justified/evaluated based on the criteria of (1) number of subjects expected, (2) number of DLT expected, (3) probability of correctly achieving the MTD, and (4) probability of overdosing through a clinical trial simulation.

As indicated in Section 4.2, CRM has an acceptable probability of correctly reaching the MTD. TER is always under estimate the MTD. STER which allows dose de-escalation does not improve the probability of correctly reaching the MTD. CRM generally performs better than that of TER design.

Depending upon the study objectives and the wishing list of the principal investigator, similar studies designs such as "$m+n$" TER design (with $m+n=6$) and CRM (k), where k is the number of subjects in each dose cohort. In the case study, we consider $k=1$. When $k \geq 2$, the accuracy and precision of the updated (estimated) MTD will improve. The comparison of these similar study designs requires further research.

REFERENCES

ICH E4 (1994) Guideline on Dose-response Information to Support Drug Registration

Ting, N.T. (2006). *Dose Finding in Drug Development*. Springer, New York.

Chang, M. and Chow, S.C. (2005). Journal of Biopharmaceutical Statistics, 15, 667-691

Cheng B, Chow SC, and Su WL. (2006). On the assessment of dose proportionality: a comparison of two slope approaches. *Journal of Biopharmaceutical Statistics*, 16, 385-392.

Chow, S.C. (2009). One-day short course on Adaptive Design Methods in Clinical Research. Food and Drug Administration, Rockville, Maryland, September 11, 2009.

Crowley, J. (2001). *Handbook of Statistics in Clinical Oncology*, Marcel Dekker, Inc., New York, New York.

ExpDesign (2002). Clinical Trial Design with ExpDesign Studio, CTriSoft Intl. Boston, MA. www.ctrisoft.net.

Spriet, A. and Dupin-Spriet, T. (1996). *Good Practice of Clinical Drug Trials.* Karger, S. Karger AG, Medical & Scientific Publication, Basel.

Storer, B. E. (1989). Design and analysis of phase I trials. *Biometrics*, 45, 925-937.

Storer, B. E. (1993). Small-sample confidence sets for the MTD in a phase I clinical trial. *Biometrics*, 49, 1117-1125.

Storer, B. E. (2001). An evaluation of phase I clinical trial designs in the continuous dose-response setting. *Statistics in Medicine*, 20, 2399-2408.

Table 1. Summary of a clinical trial simulation (based on 5,000 simulation runs)

Design	# patients expected (N)	# of DLT expected	Mean MTD (SD)	Prob. of selecting correct MTD
"3+3" TER	15.23	2.8	1.94 (0.507)	0.392
"3+3" STER*	17.59	3.2	1.70 (0.499)	0.208
CRM**	13.82	3.2	2.33 (0.451)	0.696

* Allows dose de-escalation

** Uniform prior was used

Simulation was performed using ExpDesign (2002).