Available online at www.sciencedirect.com
ScienceDirect
Contemporary Clinical Trials 29 (2008) 32–41
Contemporary Clinical Trials
www.elsevier.com/locate/conclintrial

# Optimal and minimax three-stage designs for phase II oncology clinical trials

Kun Chen [a,*], Michael Shan [b]

[a] Global Biometric Science, Pharmaceutical Research Institute, Bristol Myers Squibb Company, 5 Research Parkway, Wallingford, CT 06492, USA
[b] Department of Global Biometry, Bayer Pharmaceutical Corporation, 400 Morgan Lane, West Have, CT 06516, USA

Received 5 February 2007; accepted 24 April 2007

**Abstract**

The common objective of oncology phase II trials is to evaluate the anti-tumor activity of a new agent and to determine whether the new drug warrants further investigation. For cancer drugs that significantly shrink tumors, response (CR and PR) rate is usually the primary endpoint in cancer phase II trials for testing $H_0$: $P \le P_0$ vs $H_1$: $P \ge P_1$, where $P_0$ and $P_1$ are response rates which does not or does warrant further investigation given the rate of false positive ($\alpha$) and false negative ($\beta$). Multiple-stage designs including two-stage and three-stage have been developed by several authors. For example, Simon's optimal two-stage design [Simon R. Optimal two-stage designs for phase II clinical trials. Control Clin Trials 1989;10:1–10], Ensign et al. optimal three-stage design with restriction at the first stage [Ensign LG, Gehan EA, Kamen DS, Thall PF. An optimal three-stage design for phase II clinical trials. Stat Med 1994;13:1727–1736], Chen's optimal three-stage design without any restriction [Chen TT. Optimal three-stage designs for phase II clinical trials. Stat Med 1997;16:2701–2711], etc. However, all the above designs only early terminate a trial due to lack of activity of the study drug. Fleming's multiple-stage design [Fleming TR. One-sample multiple testing procedure for phase II clinical trials. Biometrics 1982;38:143–151] allows early stopping for either sufficient activity or lack of activity. But his design does not attempt to optimize its efficiency.

We extend Chen's [Chen TT. Optimal three-stage designs for phase II clinical trials Stat Med 1997;16:2701–2711] design and propose an optimal and a minimax design for three-stage cancer phase II trials which allows early stopping under both hypotheses. The design is optimal in the sense that the average sample number (ASN) is minimized under $P=P_0$. The minimax design minimizes the maximal sample size ($N$) and then given this value of $N$ minimizes the average sample number under $P=P_0$.


*Keywords:* Two-stage design; Three-stage design; Optimal; Minimax; Average sample number; Maximal sample size

## 1. Introduction

In cancer drug development, phase I trials are to determine the maximal tolerated dose which usually is the recommended dose in phase II trials. Phase II trials are intended to assess the anti-tumor activity of a new drug. Single-arm trials without control are commonly used in phase II trials for cytotoxic agents. Phase I trials usually include patients with

* Corresponding author. Tel.: +1 203 677 5010; fax: +1 203 6777279.
*E-mail address:* kun.chen@bms.com (K. Chen).

1551-7144/$ - see front matter © 2007 Elsevier Inc. All rights reserved.
doi:10.1016/j.cct.2007.04.008

different tumor types. But in phase II trials, patients with one specific type of tumor who meet inclusion criteria would be enrolled into the studies.

Typically, tumor response rate is the primary endpoint for single-arm cancer phase II trials. There are two standard criteria, the standard World Health Organization (WHO) Criteria and the Response Evaluation Criteria in Solid Tumors Group (RECIST) criteria, for evaluating tumor response to treatment. There are some differences between these two criteria. In particular, the WHO criteria uses bi-dimension (tumor area) and the RECIST criteria use uni-dimension (longest diameter) to measure tumor size. For example, by WHO criteria, Complete Response (CR) means disappearance of all known tumors. Partial Response (PR) means $\geq 50\%$ decrease in the sum of all target lesions, no new lesions. Stable Disease (SD) means $<50\%$ decrease and $<25\%$ increase in the sum of all target lesions. Progressive Disease means $\geq 25\%$ increase in the sum of all target lesions or appearance of new lesions. CR and PR are considered as 'responders' and SD and PD are considered as 'non-responders'. Response rate is calculated as the number of patients with CR or PR divided by the total number of patients.

Our objective in phase II trials is to determine whether a new drug is worth further investigation, for example, in a large more costly phase III trial. Both desirable and undesirable response rates would be specified in the protocol. Given the false-positive ($\alpha$) and false-negative ($\beta$) rates, the null hypothesis could be tested: $H_0$: $P \leq P_0$ vs $H_1$: $P \geq P_1$, where $P_0$ is the true response rate which is not clinically meaningful, and $P_1$ is the true response rate which is sufficiently high to warrant further study. The probability of rejecting a promising drug should be required to be less than or equal to $\beta$ if the alternative hypothesis is true and the probability of accepting an inactive drug should be required to be less than or equal to $\alpha$ under the null hypothesis.

The three-stage trial design, which is the topic of this manuscript, is based on a one-sample binomial statistic with the probability of success as the probability of response. The multi-stage design allows early stopping of a trial due to lack of activity or promising effect. Early termination can save drug development time, which may reduce the cost and bring efficacious treatments to patients early. Multiple-stage designs, including two-stage and three-stage designs, have been developed by the following authors among others.

Simon's optimal two-stage design for phase II trials [1] is a very popular multiple-stage design for phase II oncology trials. Simon's design does not allow early acceptance of the drug. Early termination occurs only when the drug has low activity. Two-stage design is easier than three-stage design from clinical operational standpoint because patient accrual at the end of each stage may have to be suspended. Accrual suspension may cause difficulty for future patient enrollment.

However, Simon's design does not allow early termination if there is a long run of responses at the start. Ensign et al. [2] proposes a three-stage design which permits early stopping when a moderately long sequence of initial responses occurs. They put a constraint at the first stage, if all treated patients respond, then the trial stops. If there are one or more responses in stage 1, then continue to stages 2 and 3 using the same stopping rules as in Simon's design.

Chen's [3] designs extend both the optimal and minimax two-stage designs to three-stage designs without any restriction at the first stage. His extension reduces the average sample number when the treatment is ineffective by an average of 10% from those of two-stage designs.

Our optimal and minimax three-stage designs extend Chen's design to allow early stopping due to effective or ineffective drug. They give more opportunity to terminate the trial earlier based on the number of responses at the early stages. This may facilitate the further development of effective drug if the sufficient number of responses occurs at an early stage and also terminate the trial earlier based on the insufficient number of responses at an early stage.

## 2. Proposed three-stage designs

### 2.1. General notation

Let $n_i$ denote the number of patients enrolled into a trial at the $i$th stage, $i=1, 2, 3$, and $N_i$ represent the cumulative sample size at the $i$th stage, $N_1=n_1$ and $N_2=n_1+n_2$, the total sample size $N_3=N=n_1+n_2+n_3$. Let $s_i$ denote the number of responses among the $n_i$ ($i=1, 2, 3$) patients, $S_g$ denote the number of cumulative responses observed at the $g$th stage ($g=1, 2, 3$), where $S_g$ is a binomial random variable ($N_i$, $P$). Let $a_i$ denote the acceptance points (of $H_0$) and $r_i$ denote the rejection points (of $H_0$), where $i=1, 2, 3$. The decision rules for stopping or continuing the trial are:

At stage $g$ ($g=1, 2$),
If $S_g=\sum_{i=1}^{g} s_i \leq a_g$, stop and reject $H_1$;