PharmaSUG China 2019 - Paper SP-053

Simon's Two Stage with Safety Assessment Added

Xiaojing Wu, dMed Biopharmaceutical Co., Ltd.

ABSTRACT

Simon's Two stage is a widely used clinical trial design to initially investigate if a new drug or treatment regimen has sufficient efficacy worthy conducting a confirmatory one in subsequent development stage. The clinical trial in such a design primarily focuses on efficacy evaluation but there is a need from industry, especially those who conduct trials not for submission purpose, wanting to embed a safety (DLT) assessment during the 1st stage of the two stage deign. The reason to do so is easily understood, which is time saving and financially efficient. However, Simon's Two stage is a clinical trial design having well-established probability structure with scientific rigor. Arbitrarily adding safety assessment and thus changing the decision-making process will violate the original probability structure which will inevitably result in unexpected Type I and Type II error rate inflation or deflation. This paper will formulate a set of multiple Null and Alternative hypotheses and provide a modified Simon's probability function (Simon, 1989) by incorporating an additional safety endpoint in the original Simon's design. By this way this paper will further illustrate by an example how an arbitrarily added safety assessment will alter the original specified Type I and Type II errors. At the end some suggestions on more generalized application of the modified probability function will be provided as well.

INTRODUCTION

The formal pathway of developing a new treatment in clinical trial is to firstly identify the Maximal Tolerable Dose (MTD) in a Phase I study dose finding study and then followed by a single arm Phase II study in which the preliminary efficacy results will be obtained and evaluated. The decision on whether to carry out a confirmatory study in subsequent stage will be made. Simon's two stage is such a design that is well-established for Phase II study with a binomial efficacy endpoint, i.e. Objective Response Rate (ORR). There are some practical needs from industry, which attempt to extend the Simon's two stage design to include a safety endpoint in addition to the efficacy one. The basic idea behind this need is firstly there are still some uncertainties remaining around the safety profile of the new treatment even after MTD has been established in the Phase I dose finding study and secondly there is always a need to make best use of the limited number of subjects enrolled. In this paper Simon's two stage design will be slightly modified to incorporate a safety assessment measured by Dose Limited Toxicity (DLT) and to be embedded in the 1st stage only. The Null hypothesis for this safety endpoint will be an 'acceptable toxicity' in terms of a DLT rate lower than a pre-defined value and the Alternative will be the complementary set of the Null. An unexpected high DLT rate observed in the 1st stage of the trial will lead to rejection of the Null for safety. The individual hypothesis of efficacy and safety will be combined to form a joint Null that is the new treatment has either 'unacceptable toxicity' or 'undesirable efficacy' and joint Alternative that is the new treatment has both 'acceptable toxicity' and 'desirable efficacy'. The probability of rejecting the joint Null or joint Alternative will be derived in this paper given the pre-specified joint decision rules and can be expressed in terms of the Type I and Type II error rates induced by original Simon's two stage design with the efficacy endpoint only. By using the rejection probability derived, this paper suggests some more generalized applications such as looking for most fitted trial design parameters under different error control criteria.

SIMON'S TWO STAGE DESIGN FOR EFFICACY ASSESSMENT

In his 1989 paper, Simon, R proposed a two-stage design for Phase II oncology study, in which the binomial efficacy endpoint, usually the ORR, will be assessed in two separate stages and the decision on rejecting or accepting hypothesis will be made at each stage according to the respective number of ORs(Objective Response) observed. The null and alternative are specified as below:

$$\mathrm{H0}: p \leq p_0$$

$$\mathrm{H1}: p > p_0$$

where p_0 is such an ORR value that has the minimal clinical relevance and/or the effect of Standard of Care (SoC). To reject the $\mathrm{H1}: p > p_0$, Simon provided the following probability function including a set of key parameters of the two-stage trial design:

$$B(r_1, p, n_1) + \sum_{x=r_1+1}^{\min(n_1, r)} b(x, p, n_1)\, B(r - x, p, n_2)$$

Where n_1 and n_2 are the number of subjects who are put on the test at the first and second stage of the trial respectively. r_1 is such a cut-off number that if no greater than r_1 ORs are observed at the end of 1st stage, $\mathrm{H0}$ cannot be rejected and the trial will be early terminated due to futility result. Similarly, the r is such a cut-off number that if no greater than $r - x$ number of ORs are observed at the end of the second stage, the $\mathrm{H0}$ will NOT be rejected as a final decision. B and b denote the cumulative and density function of binomial distribution respectively. Given the boundary value of p_0, the above probability function returns the probability of correctly accepting $\mathrm{H0}$ when efficacy is truly undesirable, which can be expressed also as $1 - \alpha$, where α denotes the Type I error rate.

Similarly, given p_1, which is the most expected efficacy value in $\mathrm{H1}$, the Simon probability function gives a probability of falsely rejecting H1 when efficacy is truly desirable, which is Type II error rate denoted by β .

To search for a solution among all those sets of n_1, n_2, r_1 and r that can satisfy pre-specified α and β, Simon proposed that the optimal solution should be the one that can minimize the expected sample size.

SAFETY ASSESSMENT

By using the subjects enrolled in the trial with Simon's two stage design, where the efficacy assessment is the primary focus, the safety endpoint can also be assessed meanwhile. Moreover, since the commonly used safety endpoint in early development is DLT, which follows a binomial distribution, simple exact test can be used to conduct hypothesis testing and make relevant decision. If the safety assessment is performed and assessed throughout the entire Simon's design with two stages, the complexity in probability structure and the challenge in trial operation will likely discourage its application. So, the safety assessment proposed in this paper will be limited into the 1st stage and will be using the n_1 subject only.

The hypothesis for the safety will be constructed in a way slightly different with the usual one, in which a range of acceptable DLT rates will be put on Null and its complementary set will be put on the Alternative as unacceptable DLT rates. The reason to do so is that before Phase II clinical trial, most new treatments have already been tested for toxicity with MTD having been determined in Phase I study. Therefore, the established safety profile of the new treatment can only be rejected when an unexpected high DLT rate is observed during a Phase II study. The following hypothesis will be set up for safety assessment that is embedded in the 1st stage of the Simon's two stage.

$$\mathrm{H0}: \pi \leq \pi_0$$

$$\mathrm{H1}: \pi > \pi_0$$

where π_0 is such a boundary value that being less or equal to it is a range of 'acceptable' DLT rates. Suppose r_d is such a cut-off that if greater than r_d number of DLTs are observed in the 1st stage of the trial, the Null will be rejected. Given the most expected value of π_1 in H1, the probability function $B(r_d, \pi_1, n_1)$ returns a probability of falsely rejecting $\mathrm{H1}$ when treatment is truly unacceptable and probability function of $1 - B(r_d, \pi_0, n_1)$ returns a probability of falsely rejecting $\mathrm{H0}$ when the treatment is truly acceptable.

MULTIPLE HYPOTEHSIS TESTING

In this section the joint hypotheses combining efficacy and safety assumptions will be formulated and the rejection probability given the testing rules and under joint hypothesis will be provided as well.

1. Joint Hypothesis Formulation

- Joint Null Hypothesis

  As introduced in previous section, the individual Null hypothesis for efficacy is no greater than an

undesirable ORR and the safety is no greater than an acceptable DLT rate. However, the joint Null should not be constructed as a simple union or interaction of the two individual Nulls, but a form that reflect the strategy on how to give green light to the treatment for further extensive development. In this paper it is supposed that a new treatment should be denied further development when it has either unacceptable safety or undesirable efficacy. Therefore, the joint Null hypothesis then can be constructed as a union of Null of efficacy and Alternative of safety.

$$\mathrm{H0}: p \le p_0 \; or \; \pi > \pi_0$$

To reject the joint Null and make the trial a success, the unacceptable safety and undesirable efficacy should be rejected simultaneously. Joint Alternative Hypothesis

- The joint Alternative is defined as the complementary set of the joint Null as below

$$\mathrm{H1}: p > p_0 \; and \; \pi \le \pi_0 \,,$$

which assumes the new treatment has both a desirable efficacy and acceptable safety.

2. Probability of Rejection of Joint Hypothesis

If the independency between the efficacy and safety endpoints, i.e. ORR vs.DLT rate, is assumed, the probability of rejecting the joint hypothesis can be derived giving pre-specified rejection rule.

- The joint $\mathrm{H1}$ can be rejected if either $\le r_1$ number of ORs or $> r_d$ number of DLTs are observed in n_1 subjects at 1st stage or $\le r - x$ ORs are observed in n_2 subjects at the 2nd stage. So, the probability of rejecting the join Alternative is given by:

$$B(r_d, \pi, n_1)\left[B(r_1, p, n_1) + \sum_{x=r_1+1}^{\min(n_1, r)} b(x, p, n_1)\, B(r - x, p, n_2)\right] + [1 - B(r_d, \pi, n_1)] \qquad 1.$$

When p takes value of p_1 which is the most expected value in $p > p_0$ and π takes value of π_0, the above probability becomes the type II error of falsely rejecting desirable ORR and acceptable DLTs, and can be re-written as:

$$B(r_d, \pi_0, n_1)[\beta_1 + \beta_2] + [1 - B(r_d, \pi_0, n_1)] \qquad 1.1$$

Where β_1 and β_2 are the type II errors for the original Simon's design at 1st and 2nd stage with only efficacy endpoint ORR respectively.

The probability of falsely terminate the trial early at the end of 1st stage will be given by:

$$B(r_d, \pi_0, n_1)\beta_1 + [1 - B(r_d, \pi_0, n_1)] \qquad 1.2$$

- The joint $\mathrm{H0}$ will be rejected if $> r_1$ number of ORs and $\le r_d$ number of DLTs are both observed among n_1 subjects at the 1st stage and $> r - x$ number of ORs are observed among the n_2 subjects at the 2nd stage. So, the probability of rejecting the joint Null is given by:

$$B(r_d, \pi, n_1)\left[I(r \le n_1)\,[1 - B(r - 1, p, n_1)] + \sum_{x=r_1+1}^{min(n_1, r)} b(x, p, n_1)\,[\,1 - B(r - x, p, n_2)]\right] \qquad 2.$$

Where the $I(r \le n_1)$ is an indicator function which equals 1 when $r \le n_1$.

The joint Null actually produces three different hypotheses when p and π take different pair of values. For the first two, where $p = p_0$ and $\pi = \pi_1$ or $p = p_0$ and $\pi = \pi_0$, the above Null rejection probability becomes the Type I error rates of falsely rejecting undesirable efficacy with either unacceptable or acceptable safety. They also can be re-written in terms of α_1 and α_2 as:

$$B(r_d, \pi_1, n_1)(\alpha_1 + \alpha_2) \text{ or} \qquad 2.1$$

$$B(r_d, \pi_0, n_1)(\alpha_1 + \alpha_2) \qquad 2.2$$

Where α_1 and α_2 are the type I errors for the original Simon's design at 1st and 2nd stage with efficacy endpoint only respectively.

For the third scenario, which is $p = p_1$ and $\pi = \pi_1$, the above Null rejection probability becomes the Type I error rate of falsely rejecting desirable efficacy with unacceptable safety. It can be re-written as:

$$B(r_d, \pi_1, n_1)(1 - \beta_1 - \beta_2) \quad 2.3$$

Note that the errors rate produced when rejecting the above joint Null is not the so-called Family Wise Error Rate (FWER) when performing multiple hypotheses tests. The former one defines error rate as the probability of rejecting all the individual hypothesis when at least one of them is true. The latter one defines the error rate as the probability of falsely rejecting at least one individual hypothesis.

3. An example

To illustrate how an embedded safety assessment at 1st stage will impact the error rates of a given Simon's two stage design, suppose we are going to test the efficacy of ORR at a single arm study. The Null and Alternative for this individual efficacy test is given by $\mathrm{H0}: p \leq 0.2$ and $\mathrm{H1}: p > 0.2$ respectively. The most expected value of p in $\mathrm{H1}$ is given by 0.4. Then the optimal Simon's two stage design with $\alpha = 0.05$ and $\beta = 0.2$ will have study parameters of $n_1 = 13$, $r_1 = 3$, $n_2 = 30$ and $r = 12$.

If a safety assessment for DLT will be carried out simultaneously during the 1st stage of the above Simon's two stage design, $n_1 = 13$ should be the only sample size to be considered. A DLT cut-off $r_d > 6$ will be used to reject $\pi_0 = 0.3$ in $\mathrm{H0}: \pi \leq 0.3$ which is an acceptable DLT hypotheses and $r_d \leq 6$ will be used to reject $\pi_1 = 0.5$ in $\mathrm{H1}: \pi > 0.3$ which is an unacceptable DLT hypotheses.

The joint Hypothesis will be given by:

$$\mathrm{H0}: p \leq 0.2 \; or \; \pi > 0.3$$

$$\mathrm{H1}: p > 0.2 \; and \; \pi \leq 0.3,$$

By using probability function of 1.1, we can quickly obtain the Type II error rate for this joint design which is given by: $B(6, 0.3, 13) * 0.2 + [1 - B(6, 0.3, 13)] = 0.2499$.

By using probability function of 2.1 and 2.2, we can quickly obtain the Type I error of falsely rejecting undesirable efficacy with either unacceptable or acceptable safety, which is given by:

$B(6, 0.5, 13) * 0.05 = 0.025$ and $B(6, 0.3, 13) * 0.05 = 0.0469$ respectively.

By using probability function of 2.3, we can quickly obtain the Type I error of mistakenly rejecting desirable efficacy with unacceptable safety, which is given by:

$B(6, 0.5, 13) * (1 - 0.2) = 0.4$

DISCUSSIONS

By using the probability function of 1 and 2, the Type I and Type II error rate of testing the joint hypothesis of both efficacy and safety endpoints can be expressed by a set of parameters used in the original Simon's two stage design with efficacy endpoint only plus a cut-off of the newly embedded safety assessment at the 1st stage, namely which are given by n_1, r_1, n_2, r and r_d.

For a given Simon's two stage design for efficacy only with given α and β, we can evaluate the impact of adding and evaluating a safety endpoint on the overall Type I and Type II error rate by using the simplified functions of 1.1, 2.1, 2.2 and 2.3.

A more generalized application of the probability function of 1 and 2 is that we can search solutions for n_1, r_1, n_2, r and r_d simultaneously, which can meet a set of pre-specified overall Type I and Type II error rates and other criteria if applicable, such as minimization of the expected sample size when the true situation is that the treatment has both undesirable efficacy and unacceptable safety. We can also search solution of n_1, r_1, n_2, r and r_d that can satisfy pre-specified error rate of an individual test for only one of the two endpoints, either safety or efficacy.

REFERENCES

Simon R. 1989. "Optimal Two-Stage Designs for Phase II Clinical Trials" Controlled Clinical Trials 10:1-10.

CONTACT INFORMATION

Your comments and questions are valued and encouraged. Contact the author at:

Xiaojing Wu
dMed Biopharmaceutical Company Limited
18811247771
Xiaojing.wu@dmedglobal.com