Published in final edited form as:

Stat Med. 2016 April 15; 35(8): 1257–1266. doi:10.1002/sim.6794.

Optimal adaptive two-stage designs for early Phase II clinical trials

Guogen Shan^{1,*}, Gregory Wilding², Alan Hutson², and Shawn Gerstenberger¹

¹Department of Environmental and Occupational Health, School of Community Health Sciences, University of Nevada Las Vegas, Las Vegas, NV 89154

²Department of Biostatistics, University at Buffalo, Buffalo, NY 14214

Abstract

Simon's optimal two-stage design has been widely used in early phase clinical trials for Oncology and AIDS studies with binary endpoints. With this approach, the second stage sample size is fixed when the trial passes the first stage with sufficient activity. Adaptive designs, such as those due to Banerjee and Tsiatis (2006) and Englert and Kieser (2013), are flexible in the sense that the second stage sample size depends on the response from the first stage, and these designs are often seen to reduce the expected sample size under the null hypothesis as compared to Simon's approach. An unappealing trait of the existing designs is that they are not associated with a second stage sample size which is a non-increasing function of the first stage response rate. In this paper an efficient intelligent process, the branch-and-bound algorithm, is used in extensively searching for the optimal adaptive design with the smallest expected sample size under the null, while the type I and II error rates are maintained and the aforementioned monotonicity characteristic is respected. The proposed optimal design is observed to have smaller expected sample sizes compared to Simon's optimal design, and the maximum total sample size of the proposed adaptive design is very close to that from Simon's method. The proposed optimal adaptive two-stage design is recommended for use in practice to improve the flexibility and efficiency of early phase therapeutic development.

Keywords

Adaptive design; Brand-and-bound algorithm; Optimal design; Simon's optimal design

1 Introduction

In Oncology studies, one important goal of early phase II clinical trials is to investigate the activity of a new treatment or a targeted therapy. In these trials, binary endpoints are often measured to calculate the response rate associated with efficacy. These studies are widely based on a one-arm design where all patients received the experimental treatment. This approach is not restricted to Oncology but has also been implemented in areas such as AIDS research [1] and gastroesophageal research [2]. Due to ethical and economic considerations,

*Corresponding author. Epidemiology and Biostatistics Program, Department of Environmental and Occupational Health, School of Community Health Sciences, University of Nevada Las Vegas, Las Vegas, NV 89154, USA. guogen.shan@unlv.edu.

a multi-stage design is often preferable given the benefits of sample size savings and patient protection in that the study stops earlier when the treatment is observed to be ineffective. The most commonly used multi-stage design is Simon's two-stage design [3] and Simon's approach has been the basis for subsequent methodological proposals [4, 5, 6].

In Simon's two-stage optimal design, the second stage sample size is fixed regardless of the number of responses observed during the first stage. This characteristic can be seen as unappealing in that treatments which have proven extremely efficacious in the first stage would not require much in terms of further evaluation as compared to treatments that were seen to be borderline in the initial cohort of patients. For this reason a number of adaptive clinical trial designs have been proposed. Lin and Shih [7] was among the first to propose adaptive two-stage design for the single arm study with binary endpoints, with type I and II error rates being respected. When the first stage response rate is seen to be larger than a pre-specified critical value, the trial proceeds to the second stage under one of two different scenarios defined by the sample space referred to as the optimistic and skeptic target response rates, each associated with a different second stage sample size. Later, Banerjee and Tsiatis [8] derived an optimal adaptive two-stage design based on a Bayesian decision theoretic construct, claiming their design to be close to optimal. Recently, Englert and Kieser [9] developed the optimal adaptive two-stage designs based on conditional error functions [10]. In their work the branch-and-bound algorithm was used to identify the global optimal design. The designs due to Banerjee and Tsiatis [8] and Englert and Kieser [9] allow the second stage sample size to change as a function of the first stage response rate, but they both suffer from the counter-intuitive feature that the second stage sample size is always not a non-increasing function of the number of first stage responses. It has been pointed out by Jin and Wei [11] that the second stage sample size from these adaptive designs could increase then immediately drop to zero as the number of response in the first stage increases.

From an intuitive and practical perspective, the second stage sample size should be a monotonic function of the number of responses in the first stage. Under this criteria, Jin and Wei [11] proposed a new adaptive one-arm two-stage design by using the conditional type I error rate and the conditional power under the restriction that the first stage sample size of this adaptive design is the same as that of Simon's optimal design. Although their designs meet overall type I error requirements, the actual power of the designs is often below the pre-specified power level even after adding the constraint on the conditional power [11, 9]. These constraints are added to reduce the computational search time for the optimal design, but the identified design may not be the global optimal design.

The remaining of this article is organized as follows. In Section 2, we introduce the branch-and-bound algorithm for searching the optimal adaptive design using conditional error functions. Detailed search methods are then described in this section. In Section 3 we compare the performance of the proposed optimal adaptive design with competitors regarding the expected sample size and the maximum sample size. We also provide detailed optimal adaptive designs for commonly used cases for practical usage. Section 4 is given to discussion regarding the proposed optimal adaptive design.

2 Optimal adaptive two stage designs for a one-arm study

In early Phase II clinical trials, one important task is to demonstrate sufficient activity of a new treatment. When the outcome is dichotomized, the commonly used hypotheses are

$$H_0: p \leq p_0, \text{ VS } H_a: p \geq p_1,$$

where p_0 is the unacceptable response rate, while p_1 is the desired response rate, $p_1 > p_0$, and we will operate under these hypotheses throughout.

Simon's optimal two-stage design is defined by four design parameters: (n_1, n, r_1, r) , where n_1 and n are the first stage and the maximum total sample size, respectively, and r_1 and r are the associated critical values to be compared with the observed number response for trial decisions. We note that the second stage sample $n_2 = n - n_1$ is fixed during trial conduct. For example, we consider the design with $(\alpha, \beta, p_0, p_1) = (0.05, 0.2, 0.2, 0.4)$. The function *ph2simon* in the package *clinfun* from the statistical software R, calculates the corresponding design parameters as $(n_1, n, r_1, r) = (13, 43, 3, 12)$. When the number of response is larger than 3 in the first stage based on 13 patients, an additional $30 = 43 - 13$ patients will then be enrolled in the study. That is, $n_2 = 30$ will be the second stage sample size regardless the number of response r_1 as long as $3 < r_1 \leq 13$. Consider two cases, one with 10 responders and the other with 4 responders, both which prompts the researcher to proceed to the second stage under this design. It would be reasonable to have smaller number of patients in the second stage in the former case than the latter case given the observed activity of the treatment. In an extreme case, all patients from the first stage respond to the treatment, which is already larger than the final critical value $r = 12$. Therefore, there is no need to enroll additional 30 patients, and the null hypothesis could be rejected after the first stage.

To make the design more flexible and efficient, we propose a new adaptive two-stage one-arm clinical trial design that allows the second stage sample size to depend on the number of response from the first stage based on the design parameters

$$(n_1, n_2(S), r(S)),$$

where $S = \{0, 1, 2, \dots, n_1\}$ is the possible number of response from the first stage based on n_1 patients, and $n_2(S)$ and $r(S)$ are the second stage sample size and the associated critical value in the second stage, both which are a function of S . It is reasonable to assume that the trial should be stopped at the first stage when no response is observed, that is, when $n_2(S) = 0$. The proposed adaptive design, similar to all other adaptive one-arm designs, allows for trial stoppage after the first stage for either futility or efficacy. For each possible observed number of response, s , the conditional probability of rejecting the null hypothesis is given as

$$CP(s|p_0) = 1 - \text{pbinom}(r(s) - s, n_2(s), p_0),$$

where $pbinom(a, b, c)$ is the cumulative probability function of a binomial distribution for observed value a with size b and probability c . Note that $CP(s) = 0$ and 1 when the study is terminated after the first stage for futility and efficacy, respectively. Then, the overall type I error rate is a weighted function of the conditional probabilities

$$TIE(p_0) = \sum_{s=0}^{n_1} CP(s|p_0) \times P(S=s|p_0),$$

where $P(S=s|p_0) = dbinom(s, n_1, p_0)$ is the null probability of s responses in the first stage given the first stage sample size n_1 , and $dbinom$ is the probability density function of a binomial distribution. Since probabilities are computed exactly from binomial distributions, one may identify all designs which meet type I error requirements: $TIE \leq \alpha$. Similarly, power of the study is defined as

$$Power = \sum_{s=0}^{n_1} CP(s|p_1) \times P(S=s|p_1).$$

Among all the designs satisfied the type I and II error rates constraints: $TIE \leq \alpha$ and $Power \geq (1 - \beta)$, the one with the smallest expected sample size under the null hypothesis is defined to be the optimal design (the same optimal criteria as in Simon's two-stage optimal design),

$$ESS(p_0) = \min \sum_{s=0}^{n_1} (n_1 + s) \times P(S=s|p_0).$$

For a given first stage sample size n_1 , the total $2n_1 + 2$ parameters $(n_2(S), r(S))$, $S = 0, 1, 2, \dots, n_1$, are needed to determine the optimal adaptive design. It is often reasonable in practice to set the maximum available sample size for the study, n_{max} . Then the maximum sample size in the second stage is $n_{2,max} = n_{max} - n_1$, and thus the range of $n_2(S)$ is from 1 to $n_{2,max}$, and the range for $r(S)$ is from 0 to $n_2(S)$. As the example given in Englert and Kieser [9] with $n_1 = 10$ and $n_{2,max} = 22$, the total number of possible designs is over 10^{19} . This number would increase exponentially as the first and second stage sample sizes increase. It is almost impossible to conduct the complete enumeration over a two-dimensional parameter space, $(n_2(S), r(S))$, even for small sample sizes. Englert and Kieser [9] proposed to search for the optimal design over the all attainable values of $CP(s|p_0)$ of type I error rate and $CP(s|p_1)$ for power. This would reduce the research from a two-dimensional space to one-dimensional, although it is still not feasible to compute the type I and II error rates over all conditional error functions.

To overcome the computational burden, Englert and Kieser [9] utilized the branch-and-bound algorithm [12] to identify the optimal design with the smallest expected sample size. However, the optimal design obtained is often counter-intuitive, in the sense that the second stage sample size $n_2(S)$ can be an increasing function of the number of responses observed in

the first stage S , or the relationship between $n_2(S)$ and S can be non-monotonic. Therefore, we add the following restriction in the optimal design search using the branch-and-bound algorithm: $n_2(S_1) \leq n_2(S_2)$ when $S_1 < S_2$.

The branch-and-bound algorithm is an intelligent algorithm for discrete and combinatorial optimization problems. The first component of the algorithm involves branching, a splitting procedure to separate the problem into several independent non-overlapping subproblems whose union is the problem. In our case, the discrete conditional error functions $CP(S)$, ($S = 0, 1, 2, \dots, n_1$), which are equivalent to $(n_2(S), r(S))$, are redefined in each branching procedure. Although it is not required in the algorithm to sort the order of all possible conditional error functions, the computational intense is reduced by sorting the conditional error functions by $n_2(S)$ in ascending order and $CP(S)$ in increasing order. Let N be the size of the union of the conditional error functions and 0 and 1, where 0 and 1 denote the study being terminated at the first stage for futility and efficacy, respectively. Suppose $B(S, W)$ is the W -th conditional error function in the union for the case of first stage response S , and $n_2(S, W)$ and $r(S, W)$ are the associated second stage sample size and critical value, $W = 1, 2, \dots, N$. It follows that $W_1 \leq W_2$ when $S_1 > S_2$ in order to meet the monotonicity feature of the proposed adaptive design. When $S = 0$, there is no response from the second stage, and we will stop the trial for futility. The branching procedure starts with $S = 1$ and $W = 1$. The value of W increases until the bounding conditions are satisfied. Then, the branching procedure goes to the next S , and the W value starts from the W value from the previous S , not from 1. This step would guarantee that the optimal design identified meets the monotonicity feature. The branching step is recursively applied with the following the bounding procedure.

The bounding procedure calculates the lower and upper bounds of the target functions in order to discard candidates which are not optimal. In our application, three target functions are considered: the minimum type I error rate, the maximum power, and the minimum expected sample size under the null hypothesis. These three functions are computed at each branching procedure. Suppose the current branching result is $B(0, 1), B(1, W_1), B(2, W_2), \dots, B(k, W_k), B(k+1, 1), \dots, B(n_1, 1)$. The maximum of power is calculated as

$$Power_{max} = \sum_{s=0}^k CP(s|p_1, n_2(s, w_i), r(s, w_i)) \times P(s=s|p_1) + \sum_{i=k+1}^{n_1} P(S=s|p_1);$$

and the minimum expected sample size under the null is

$$EN_{min}(p_0) = \sum_{s=0}^k (n_1 + n_2(s, w_i)) \times P(S=s|p_0) + \sum_{i=k+1}^{n_1} n_1 P(S=s|p_0).$$

It is always the case that the conditional type I error rate $CP(S|p_0)$ is a nondecreasing function of S . In other words, the more responses observed in the first stage, the higher the

conditional type I error will be [9]. This restriction is used in the optimal design search. To satisfy this restriction, the minimum type I error rate is given as

$$TIE_{min}(p_0) = \sum_{s=0}^k CP(s|p_0, n_2(s, w_i), r(s, w_i)) \times P(S=s|p_0).$$

The three target functions are used in the bounding procedure to discard subproblems that do not lead to the optimal design. The subproblems do not need to be computed when the minimum type I error rate is above α , or the maximum power is less than $1 - \beta$, or minimum expected sample size is larger than EN_{min} , which is often chosen to be the value that is slightly larger than the expected sample size from Simon's optimal two-stage design. Discarding the unpromising subproblems in each branching procedure is the key to the branch-and-bound algorithm in finding the optimal design efficiently since the complete enumeration of all possible subproblems is impossible.

In the optimal adaptive design by Englert and Kieser [9], it is guaranteed the expected sample size is less than or equal to that of Simon's two-stage optimal design, and the value from Simon's design can be used as EN_{min} for their design search. It easy to show that Simon's optimal design belongs to the adaptive design when the trial is only stopped for futility at the first stage, and the second stage sample size $n(S)$ is the same for $S > r_1$. For this reason, the expected sample size of the proposed optimal adaptive design should be smaller than or equal to that from Simon's optimal design.

3 Results

We compare the proposed optimal adaptive design to the one due to the Englert and Kieser [9] (referred to as the EK design), and Simon's optimal two-stage design. The EK adaptive optimal design was compared with the adaptive design due to Banerjee and Tsiatis (referred to as the BT design), and the EK design generally requires a smaller average sample size as compared the BT design, although the improvement was shown to be small to moderate. It should be noted that both the BT adaptive design and the adaptive EK design do not respect the monotonicity feature. The design due to Jin and Wei [11] has the monotonicity trait, but it does not meet the desired type II error. It is reasonable to only compare designs which meet the type I and II error rates requirement and therefore the design due to Jin and Wei is not included in the comparison.

From a logistical point of view having an upper bound on the sample size is important from both a budgeting and regulatory point of view. Funding proposals generally require a bound on the sample size. Institutional Review Boards (IRB's) approve studies based on a fixed maximum sample size. Deviations from these numbers require protocol amendments and can delay the completion of a trial. The utilization of this trial design will be an acceptable alternative for IRB's as compared to the Simon two-stage design. It is reasonable to set the upper bound of the total sample size in the design due to limited resources or time restrictions. As recommended by Banerjee and Tsiatis [8] and other researchers, the maximum sample size is set as 110% of that from Simon's design, and we apply this

restriction in our design in order to make a fair comparison with competitors. Englert and Kieser obtained the optimal design with the first stage sample size in the range of the first stage sample size of the design due to Banerjee and Tsiatis, plus and minus 4. The first stage sample size in Simon's design is very close to that from Banerjee and Tsiatis' design, and Simon's optimal design is computationally easy and available in major statistical software packages. For this reason, we define our optimal design to have a first stage sample size in the range of the Simon's first stage sample size ± 4 .

We present the three designs (Simon's optimal design, the EK design, and the proposed optimal adaptive design) for the parameters $(\alpha, \beta, p_0, p_1) = (0.05, 0.2, 0.4, 0.6)$ in Figure 1. This figure shows the second stage sample size $n_2(S)$ as a function of the number of responses from the first stage, S . All three designs require 16 patients in the first stage, and all of them will fail to reject the null hypothesis when $S = 7$. When $S = 8$, additional 30 patients are needed in Simon's design. The second stage sample size in Simon's design is a constant represented in the figure by the at line when $S = 8$. The EK optimal adaptive design and the newly proposed optimal adaptive design allow the second stage sample to change with the observed responses in the first stage. For the EK design, $n_2(S)$ is not a monotonic function of S when $S = 8$, which illustrates the counter-intuitive feature of the EK design. For the newly proposed adaptive optimal design, $n(S)$ is a non-increasing function of S when $S = 8$. It can be seen from the figure that the proposed adaptive design generally needs a smaller sample sizes as compared to Simon's design, with an average sample size 17.88 for the proposed optimal adaptive design and 22.88 for Simon's optimal design. The maximum total sample size for both designs is the same, but the proposed adaptive design needs a smaller total sample sizes in majority of S values.

After the comparison between optimal designs using the specific example, we compare our design with the EK design and Simon's optimal design, in terms of the expected sample size under the null hypothesis. These designs are compared at $\alpha = 0.05$, at 80% and 90% power for various values of null response rate p_0 . Table 1 shows the expected sample size under the null for the three design, and the expected sample size under the alternative for Simon's design and the proposed design, under each configuration for $p_1 - p_0 = 0.2$. The proposed optimal adaptive design adds the intuitive feature from the EK design, thus, the expected sample size under the null of our design should be larger than or equal to that from the EK design. The difference of the expected sample size under the null between these two adaptive designs is from 0 to 1.16. As compared to Simon's design, our design is associated with a smaller expected sample size under the null in all cases. For each obtained design based on the criteria with the smallest expected sample size under the null, we also computes the expected sample size under the alternative, $ESS(p_1)$, between Simon's design and the proposed design. The sample size saving under the alternative is considerable, which could be more than 25% for the proposed design as compared to Simon's design. Similar results are observed in the cases with $p_1 - p_0 = 0.15$ in Table 2. Although the EK design can have a smaller expected sample size under the null than our design, the difference is often negligible, and the EK design does not satisfy the monotonicity feature which could lead the limited application of the EK design.

The proposed adaptive design and Simon's optimal design are presented in Table 3 for each configuration of p_0 and power when $p_1 = p_0 + 0.2$. The nominal type I error rate α is set to be as 0.05. We use one example from this table to show how to use this table for the one-arm two-stage adaptive clinical trial design, for the case with $p_0 = 0.5$ and 90% power in Table 3. For Simon's optimal design is used, $n_1 = 24$ patients are enrolled in the first stage. When the response from the first stage is larger than $r_1 = 13$, the trial goes to the second stage with additional $n_2 = 61 - 24 = 37$ patients tested. The final decision will be made after comparing the overall observed number of response to the critical value $r = 36$. For our proposed optimal design, the first stage sample size is $n_1 = 21$. The trial is rejected for futility if the number of first stage response $S \leq 11$, or for efficacy if $S \geq 17$. The second stage sample size depends on the first stage outcome S . For example, $n_2(S)$ is 38 when $S = 14$, which makes the total sample size $n(S) = 59$ in that case.

The maximum total sample size between our proposed design and Simon's optimal design is also compared. Out of total 16 cases in Table 3, our design has a smaller, equal, and larger maximum total sample size than Simon's design in 6, 9, and 1 cases, respectively. It should be noted that maximum of $n(S)$ is the same for S where $S > r_1$ in Simon's design, while $n(S)$ is non-increasing in the proposed optimal adaptive design. In the 1 case where Simon's design has smaller maximum total sample size, this occurs only in one S value and only one more subject for the proposed design. We also present the detailed design parameters for $p_1 - p_0 = 0.15$ in the Supporting Information.

As suggested from one of the reviewers, it would be interesting to compare the performance of the proposed design and Simon's design while the actual underlying response rate of p_1 is larger than the pre-specified value. For the proposed design and Simon's design when $(\alpha, \beta, p_0, p_1) = (0.05, 0.2, 0.2, 0.4)$ in Table, their power can be drawn as a function of p , $Power(p)$, where p is above p_1 . Figure 2 presents the power of the two methods. As expected, power for each method should be larger than the nominal level, $1 - \beta = 80\%$ when $p > p_1$. The proposed method generally has more power than Simon's design in this case. Although it is not presented, these two methods have similar power for majority of the cases.

4 Discussion

We presented the optimal adaptive two-stage designs for early phase clinical trials. The existing optimal adaptive two-stage designs often suffer from the counter-intuitive feature that the second stage sample size is not a decreasing function of the number of responses from the first stage. This pattern occurred for the existing adaptive two-stage designs, including the BT design and the EK design. By adding the monotonic association of the second stage sample size and the first stage response in the proposed design search, the search design meets the intuitive feature of a design, and it can be practically used in clinical trials. It is expected that the expected sample size under the null hypothesis based on the proposed method is always larger than that based on the EK method, as the proposed method is a modification of the EK method by adding the monotonic property in the design search. It can be seen from Tables 1 and 2 that their sample size difference is often small. Due to the lack of this important feature, current adaptive two-stage designs are not utilized directly in clinical trial designs. In our numerical study, we observe that our proposed

adaptive design has smaller expected sample size under the null as compared to Simon's two-stage optimal design in most cases. In addition, the maximum total sample size of our design $n(S)$ can be slightly larger than the that from Simon's design but for very few S values. The proposed adaptive design is flexible and efficient, therefore, it is recommended for use in practice.

The improvement in sample size under the null may be of a smaller magnitude as compared the proposed design to Simon's design, such savings are still advantageous in the context of where such designs may be implemented and may be viewed as non-trivial. For example, Phase II Oncology studies are often single-center where recruitment is often slow. Consider the number of on-going Phase II clinical trials, the overall sample size saving would be considerable by adopting the proposed design as compared to Simon's design. In such an environment the advantages of the proposed method may be appreciated.

We computed the type I and II error rates at the boundary of the hypotheses, p_0 and p_1 . The monotonicity of the probability of the rejection region should be met to assure that the type I and II error rates occur at the boundary. This property has been shown to be satisfied for binomial distributions with commonly used test statistics [13, 14, 15, 16]. Due to the complexity of the two-stage design, it is a challenge to prove it theoretically that each possible design has this property. However, we checked all the adaptive designs obtained in Table 3 and the table in the Supporting Information, and their actual type I and II error rates occur at the boundary, satisfying the required property.

One restriction used in the adaptive design search is that the maximum total sample size be set as the 110% of the maximum total sample size from Simon's optimal design. This is a reasonable restriction in practice, but this may slightly affect the final resulting design when the maximum sample size in practice is lower than the maximum of $n(S)$, $S = 0, 1, 2, \dots, n_1$. In this case, we would recommend the researcher to run the R software package for the adaptive design. We have developed a program based on that from Englert and Kieser [9], which is available from the first author.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We would like to thank the editor, the associate editor and two referees for their valuable comments and suggestions that improved the article. Shan also thanks the computational support from the National Supercomputing Center for Energy and the Environment (NSCEE) at UNLV. Shan's research is supported by National Institutes of Health grant 5U54GM104944.

References

1. Zheng L, Rosenkranz SL, Taiwo B, Para MF, Eron JJ, Hughes MD. The Design of Single-Arm Clinical Trials of Combination Antiretroviral Regimens for Treatment-Naive HIV-Infected Patients. *AIDS Research and Human Retroviruses*. Dec; 2012 29(4):652–657. <http://dx.doi.org/10.1089/aid.2012.0180> [PubMed: 23228206]

2. Katz PO, Gerson LB, Vela MF. Guidelines for the diagnosis and management of gastroesophageal reflux disease. The American journal of gastroenterology. Mar.2013 108(3) <http://view.ncbi.nlm.nih.gov/pubmed/23419381>.
3. Simon R. Optimal two-stage designs for phase II clinical trials. Controlled clinical trials. Mar; 1989 10(1):1–10. <http://view.ncbi.nlm.nih.gov/pubmed/2702835>. [PubMed: 2702835]
4. Shan G, Hutson AD, Wilding GE. Two-stage k-sample designs for the ordered alternative problem. Pharmaceut Statist. 2012; 11(4):287–294. <http://dx.doi.org/10.1002/pst.1499>. 10.1002/pst.1499
5. Chen TT. Optimal three-stage designs for phase II cancer clinical trials. Statistics in medicine. Dec; 1997 16(23):2701–2711. <http://view.ncbi.nlm.nih.gov/pubmed/9421870>. [PubMed: 9421870]
6. Shan G, Ma C, Hutson AD, Wilding GE. Randomized Two-Stage Phase II Clinical Trial Designs Based on Barnard's Exact Test. Journal of Biopharmaceutical Statistics. Aug; 2013 23(5):1081–1090. <http://dx.doi.org/10.1080/10543406.2013.813525>. 10.1080/10543406.2013.813525 [PubMed: 23957517]
7. Lin Y, Shih WJ. Adaptive two-stage designs for single-arm phase IIA cancer clinical trials. Biometrics. Jun; 2004 60(2):482–490. <http://view.ncbi.nlm.nih.gov/pubmed/15180674>. [PubMed: 15180674]
8. Banerjee A, Tsiatis AA. Adaptive two-stage designs in phase II clinical trials. Statistics in medicine. Oct; 2006 25(19):3382–3395. <http://view.ncbi.nlm.nih.gov/pubmed/16479547>. [PubMed: 16479547]
9. Englert S, Kieser M. Optimal adaptive two-stage designs for phase II cancer clinical trials. Biometrical Journal. Nov; 2013 55(6):955–968. <http://view.ncbi.nlm.nih.gov/pubmed/23868324>. [PubMed: 23868324]
10. Englert S, Kieser M. Improving the flexibility and efficiency of phase II designs for oncology trials. Biometrics. Sep; 2012 68(3):886–892. <http://view.ncbi.nlm.nih.gov/pubmed/22150825>. [PubMed: 22150825]
11. Jin H, Wei Z. A new adaptive design based on Simon's two-stage optimal design for phase II clinical trials. Contemporary clinical trials. Nov; 2012 33(6):1255–1260. <http://view.ncbi.nlm.nih.gov/pubmed/22772088>. [PubMed: 22772088]
12. Wolsey, LA. Integer Programming. 1. Wiley-Interscience; 1998. <http://www.worldcat.org/isbn/0471283665>
13. Röhm J, Mansmann U. Unconditional Non-Asymptotic One-Sided Tests for Independent Binomial Proportions When the Interest Lies in Showing Non-Inferiority and/or Superiority. Biometrical Journal. May; 1999 41(2):149–170. [http://dx.doi.org/10.1002/\(sici\)1521-4036\(199905\)41:2%3C149::aid-bimj149%3E3.0.co;2-e](http://dx.doi.org/10.1002/(sici)1521-4036(199905)41:2%3C149::aid-bimj149%3E3.0.co;2-e). 10.1002/(sici)1521-4036(199905)41:2%3C149::aid-bimj149%3E3.0.co;2-e
14. Berger RL, Sidik K. Exact unconditional tests for a 2 x 2 matched-pairs design. Statistical methods in medical research. Mar; 2003 12(2):91–108. <http://view.ncbi.nlm.nih.gov/pubmed/12665205>. [PubMed: 12665205]
15. Shan, G.; Ma, C. Unconditional tests for comparing two ordered multinomials. Statistical Methods in Medical Research. 2012. In press <http://dx.doi.org/10.1177/0962280212450957>
16. Wang, W.; Shan, G. Exact confidence intervals for the relative risk and the odds ratio. Biometrics. Jul. 2015 <http://view.ncbi.nlm.nih.gov/pubmed/26228945>

Figure 1.

All three optimal designs require 16 patients in the first stage for the design parameters $(\alpha, \beta, p_0, p_1) = (0.05, 0.2, 0.4, 0.6)$. The second stage sample size, $n_2(S)$, is plotted as a function of the number of response from the first stage, S .

Figure 2. Power comparison between the proposed design and Simon's design when the alternative response rates different from that pre-specified. The designs for $(\alpha, \beta, p_0, p_1) = (0.05, 0.2, 0.2, 0.4)$ are considered.

Table 1

Comparison between three optimal designs for expected sample size at $\alpha = 0.05$ given $p_1 - p_0 = 0.2$.

p_0	p_1	Power	$ESS_{Simon}(p_0)$	$ESS_{EK}(p_0)$	$ESS_{New}(p_0)$	$ESS_{Simon}(p_1)$	$ESS_{New}(p_1)$
0.05	0.25	0.8	11.96	11.03	11.21	16.40	13.16
		0.9	16.76	16.40	16.68	28.42	24.41
0.1	0.3	0.8	15.01	14.72	14.85	26.16	21.64
		0.9	22.53	21.70	22.38	33.98	24.85
0.2	0.4	0.8	20.58	19.80	20.48	37.94	28.50
		0.9	30.43	29.02	29.74	51.56	38.43
0.3	0.5	0.8	23.63	23.02	23.45	41.32	35.99
		0.9	34.72	33.31	34.08	60.04	50.00
0.4	0.6	0.8	24.52	24.09	24.39	41.73	40.03
		0.9	35.98	34.48	35.64	62.81	50.20
0.5	0.7	0.8	23.50	23.03	23.33	39.33	36.58
		0.9	34.01	32.95	33.45	58.25	48.57
0.6	0.8	0.8	20.48	19.72	20.28	37.84	31.73
		0.9	29.47	28.15	28.74	50.70	42.91
0.7	0.9	0.8	14.82	14.82	14.82	24.60	24.60
		0.9	21.23	20.42	20.80	34.83	32.73

Table 2

Comparison between three optimal designs for expected sample size at $\alpha = 0.05$ given $p_1 - p_0 = 0.15$.

p_0	p_1	Power	$ESS_{Simon}(p_0)$	$ESS_{EK}(p_0)$	$ESS_{New}(p_0)$	$ESS_{Simon}(p_1)$	$ESS_{New}(p_1)$
0.05	0.2	0.8	17.62	17.45	17.59	26.96	24.46
		0.9	26.66	25.83	25.92	39.85	32.03
0.1	0.25	0.8	24.66	24.41	24.49	39.62	32.55
		0.9	36.82	35.13	36.45	62.65	45.38
0.2	0.35	0.8	35.37	34.05	34.87	63.86	50.53
		0.9	51.45	50.07	50.80	80.30	69.65
0.3	0.45	0.8	41.71	40.61	41.33	72.76	64.97
		0.9	60.77	58.56	59.96	96.92	88.68
0.4	0.55	0.8	44.93	43.20	44.05	76.17	69.24
		0.9	63.96	62.85	63.84	96.89	90.62
0.5	0.65	0.8	43.72	42.15	43.01	75.15	69.89
		0.9	62.29	60.77	61.87	96.60	98.17
0.6	0.75	0.8	39.35	37.86	38.53	62.47	60.87
		0.9	55.60	54.30	54.99	91.28	84.56
0.7	0.85	0.8	30.29	29.16	29.78	53.22	48.78
		0.9	43.40	41.57	42.60	75.25	70.08

Table 3

Proposed optimal adaptive designs for $p_1 = p_0 + 0.2$ at $\alpha = 0.05$. Simon's optimal design is provided as reference, with design parameters (n_1, n, r_1, r) .

Power=80%				Power=90%			
S	$n_2(S)$	$r(S)$	$n(S)$	S	$n_2(S)$	$n(S)$	$r(S)$
$p_0 = 0.05$							
Simon:(9, 17, 0, 2)				Simon:(9, 30, 0, 3)			
New: $n_1 = 8$				New: $n_1 = 9$			
0	0	8	0	0	0	0	9
1	10	18	2	1	21	30	3
2	8	16	2	2	20	29	3
3	0	8	0	3	20	29	3
				4	0	9	0
$p_0 = 0.1$							
Simon:(10, 29, 1, 5)				Simon:(18, 35, 2, 6)			
New: $n_1 = 10$				New: $n_1 = 14$			
1	0	10	0	1	0	14	0
2	19	29	5	2	21	35	6
3	18	28	5	3	20	34	6
4	12	22	4	4	20	34	6
5	0	10	0	5	0	14	0
$p_0 = 0.2$							
Simon:(13, 43, 3, 12)				Simon:(19, 54, 4, 15)			
New: $n_1 = 14$				New: $n_1 = 19$			
3	0	14	0	4	0	19	0
4	23	37	11	5	34	53	15
5	20	34	10	6	34	53	15
6	20	34	10	7	32	51	14
7	17	31	9	8	31	50	14

Power=80%					Power=90%				
<i>S</i>	<i>u</i> (<i>S</i>)	<i>u</i> (<i>S</i>)	<i>r</i> (<i>S</i>)	<i>S</i>	<i>n</i> ₂ (<i>S</i>)	<i>n</i> (<i>S</i>)	<i>n</i> (<i>S</i>)	<i>r</i> (<i>S</i>)	<i>r</i> (<i>S</i>)
8	0	14	0	9	0	19	0		

*p*₀ = 0.3

Simon:(15, 46, 5, 18)					Simon:(24, 63, 8, 24)				
New: <i>n</i> ₁ = 15					New: <i>n</i> ₁ = 22				
5	0	15	0	7	0	22	0		
6	31	46	18	8	38	60	23		
7	31	46	18	9	37	59	23		
8	30	45	18	10	35	57	22		
9	28	43	17	11	35	57	22		
10	0	15	0	12	35	57	22		
				13	34	56	22		
				14	0	22	0		

*p*₀ = 0.4

Simon:(16, 46, 7, 23)					Simon:(25, 66, 11, 32)				
New: <i>n</i> ₁ = 16					New: <i>n</i> ₁ = 25				
7	0	16	0	11	0	25	0		
8	30	46	23	12	41	66	32		
9	30	46	23	13	41	66	32		
10	28	44	22	14	40	65	32		
11	28	44	22	15	37	62	30		
12	26	42	21	16	37	62	30		
13	24	40	21	17	0	25	0		
14	17	33	17						
15	0	16	0						

*p*₀ = 0.5

Simon:(15, 43, 8, 26)					Simon:(24, 61, 13, 36)				
New: <i>n</i> ₁ = 15					New: <i>n</i> ₁ = 21				

Power=80%				Power=90%			
S	$n_2(S)$	$n(S)$	$r(S)$	S	$n_2(S)$	$n(S)$	$r(S)$
8	0	15	0	11	0	21	0
9	28	43	26	12	38	59	35
10	28	43	26	13	38	59	35
11	26	41	25	14	38	59	35
12	24	39	24	15	38	59	35
13	21	36	22	16	36	57	34
14	0	15	0	17	0	21	0

$p_0 = 0.6$

Simon:(11, 43, 7, 30)				Simon:(19, 53, 12, 37)			
New: $n_1 = 14$				New: $n_1 = 19$			
9	0	14	0	12	0	19	0
10	24	38	27	13	33	52	36
11	21	35	25	14	31	50	35
12	20	34	24	15	31	50	35
13	20	34	24	16	31	50	35
14	6	20	16	17	14	33	22
				18	0	19	0

$p_0 = 0.7$

Simon:(6, 27, 4, 22)				Simon:(15, 36, 11, 29)			
New: $n_1 = 6$				New: $n_1 = 16$			
4	0	6	0	12	0	16	0
5	21	27	22	13	20	36	29
6	21	27	22	14	19	35	28
				15	19	35	28
				16	13	29	23