



Research Article

Randomized Phase II Clinical Trials using Fisher's Exact Test

Shiwei Cao, Lu Liu, and Sin-Ho Jung*

Department of Biostatistics and Bioinformatics, Duke University, USA

***Corresponding author:** Sin-Ho Jung, Department of Biostatistics and Bioinformatics, Duke University, 2424 Erwin Road, 11070 Hock Plaza, Suite 1102, DUMC Box 2721, Durham, NC 27710, USA

Received: 09 March 2021; **Accepted:** 16 March 2021; **Published:** 09 April 2021

Citation: Shiwei Cao, Lu Liu, and Sin-Ho Jung. Randomized Phase II Clinical Trials using Fisher's Exact Test. Archives of Clinical and Biomedical Research 5 (2021): 214-229.

Abstract

A phase II trial is to evaluate an experimental therapy using an early efficacy outcome, such as tumor shrinkage, before proceeding to a large-scale phase III trial. Traditionally, a typical phase II trial has been conducted using a single-arm design recruiting patients only to the experimental therapy to be compared with a historical control. Due to a small sample size and heterogeneity of patient population, the characteristics of the patients in a new phase II trial is often different from that of the selected historical control, so that the single-arm phase II trial results in false positive or false negative conclusions. A randomized phase II trial can resolve such problems by randomizing patients between an experimental arm and a control arm. In this paper, we propose randomized phase II trial designs based on 2-stage Fisher's exact tests allowing for both superiority and futility early stopping options, so that we can save number of patients when experimental therapy is definitely efficacious as well as when it is futile. We propose a weighted expected sample size as a new criterion to define optimal two-stage designs.

Keywords: Admissible design; Futility stopping; Minimax design; Optimal design; Superiority stopping; Weighted average sample size

1. Introduction

Phase II clinical cancer trials are to evaluate the short-term efficacy, such as tumor response, of experimental therapies and to determine to continue to a randomized phase III trial with a long-term outcome, such as overall survival. The most popular

primary outcome is the overall response, meaning partial response and complete response. Traditionally, the single-arm phase II trial was the principal mechanism to recruit a small number of patients only to the experimental arm to be compared to a historical control as proposed by Simon [1]. A single-arm trial is not appropriate if no reliable historical control exists. Even when a historical control exists, its patient population may be somewhat different from that of a new phase II trial or assessment of tumor size may differ between the historical control and a new trial. In this case, an experimental therapy and the chosen historical control may not be comparable. This happens especially when a historical control is taken from a previous phase II trial with a small number of patients. Raising various pitfalls of single-arm phase II trials, Cannistra [2] recommends randomized phase II trials as a viable alternative.

We consider randomized phase II trials with a parallel control arm. Let p_x , p_y denote the response probabilities of an experimental arm (arm x) and a control arm (arm y), respectively. We want to test the null hypothesis $H_0: p_x \leq p_y$ against the alternative hypothesis $H_1: p_x > p_y$. Sample sizes of phase II trials are not large enough to apply asymptotic theories, so that usually exact methods are used in designing and analyzing phase II trials. Jung [3] proposes a design method for randomized phase II trials, called MaxTest design, based on a binomial test. In the binomial test, the null distribution of the test statistic depends on the nuisance parameter p_y . Therefore, if p_y is mis-specified, the test may not control the type I error rate accurately. Jung [3] addresses this problem by controlling the type I error rate at $p_y = 0.5$. However, this will result in too a conservative type I error control if the p_y is different from 0.5.

Jung et al. [4] propose to use Fisher's exact test [5] as an alternative to the binomial test. In a Fisher's exact test, the null distribution of the test statistic is free of p_y by conditioning on the total number of responders which is a sufficient statistic for the nuisance parameter. Jung et al. [4] show that the design method based on Fisher's exact test has a higher power than the binomial design method in a wide range of p_y values. These randomized phase II designs consider two-stage designs with futility early stopping only. One may also want to stop early if the experimental arm is more efficacious than the control arm. In this paper, extending Jung et al. [4], we investigate two-stage randomized phase II trial designs based on Fisher's exact tests allowing for both superiority and futility early stopping. We propose weighted expected sample size as a new criterion to define optimal and admissible two-stage designs [6]. We briefly review Jung et al. [4] method in Section 2 and extend it to two-stage design with both futility and superiority interim tests in Section 3. In Section 4, we conclude this paper with some discussions.

2. Two-Stage Design with A Futility Interim Test: Review

Jung et al. [4] consider a two-stage phase II clinical trial with an early stopping rule if the experimental arm has a low efficacy. Suppose that n_1 patients are randomized to each arm during stage 1 ($= 1, 2$). Let $n = n_1 + n_2$ denote the maximal sample size

per arm, so that the maximal number of patients required for the randomized trial is $2n$. Also, let X_1 and Y_1 denote the number of responders observed from the experimental arm and the control arm, respectively, during stage 1.

Let $q_x = 1 - p_x$, $q_y = 1 - p_y$, then the odds ratio is defined as $\theta = p_x q_y / (p_y q_x)$. Fisher's exact test is conditioning on $Z_1 = X_1 + Y_1$, which is the total number of responders during stage 1. A two-stage randomized phase II trial for testing $H_0 : \theta = 1$ against $H_1 : \theta > 1$ using a futility interim test is conducted as follows:

Stage 1: Randomize n_1 patients to each arm, and observe number of responders X_1 and Y_1 .

1. If $X_1 - Y_1 \leq a_1$, then reject the experimental arm and stop the trial.
2. Otherwise, proceed to stage 2.

Stage 2: Randomize n_2 patients to each arm and observe number of responders X_2 and Y_2 . Let $X = X_1 + X_2$ and $Y = Y_1 + Y_2$ denote the cumulative number of patient for experimental and control arms, respectively.

1. Reject the experimental arm if $X - Y \leq a$.
2. Otherwise, accept the experimental arm for further investigation.

For two-stage Fisher's exact tests, both a_1 and a may change depending on the total number of responders $Z_1 = X_1 + Y_1$ and $Z_2 = X_2 + Y_2$.

2.1 Conditional Test:

Given Z_1 , X_1 has a hypergeometric distribution whose probability mass function is given as:

$$f_1(x_1 | z_1, \theta) = \frac{\binom{n_1}{x_1} \binom{n_1}{z_1 - x_1} \theta^{x_1}}{\sum_{i=m_{1-}}^{m_{1+}} \binom{n_1}{i} \binom{n_1}{z_1 - i} \theta^i}$$

for $m_{1-} \leq x_1 \leq m_{1+}$, where $m_{1-} = \max(0, z_1 - n_1)$ and $m_{1+} = \min(z_1, n_1)$. Note that the conditional distribution is free of p_x or p_y under $H_0 : p_x = p_y$ (or $\theta = 1$). Hence, a test conditioning on the values of z_1 and z_2 can control the type I error rate without specifying p_x and p_y . At the design of a 2-stage randomized phase II trial, we specify the expected response rates p_x and p_y for the experimental arm and the control arm, respectively, together with type I error rate α^* , and power $1 - \beta^*$. Given input parameters $(p_x, p_y, \alpha^*, 1 - \beta^*)$, a two-stage design, defined by (n_1, n) and $\{a_1(z_1), a(z_1, z_2), 0 \leq z_1 \leq 2n_1, 0 \leq z_2 \leq 2n_2\}$, is identified as follows.

2.2 How to choose a_1 and a given (n_1, n) :

Suppose that sample sizes (n, n_1) are given and (z_1, z_2) are observed from the trial. For a two-stage design defined by (n_1, n_2) and $\{a_1(z_1), a(z_1, z_2) : 0 \leq z_l \leq 2n_l, l = 1, 2\}$, the conditional type I error rate and the conditional power function are calculated as:

$$\alpha(z_1, z_2) = P(X_1 - Y_1 > a_1, X_1 + X_2 - Y_1 - Y_2 > a | z_1, z_2, \theta = 1)$$

$$= \sum_{x_1=m_{1-}}^{m_{1+}} \sum_{x_2=m_{2-}}^{m_{2+}} f_1(x_1|z_1, 1) f_2(x_2|z_2, 1) I\{x_1 > (z_1 + a_1)/2, x_1 + x_2 > (z_1 + z_2 + a)/2\}$$

and

$$1 - \beta(z_1, z_2) = P(X_1 - Y_1 > a_1, X_1 + X_2 - Y_1 - Y_2 > a | z_1, z_2, \theta)$$

$$= \sum_{x_1=m_{1-}}^{m_{1+}} \sum_{x_2=m_{2-}}^{m_{2+}} f_1(x_1|z_1, \theta) f_2(x_2|z_2, \theta) I\{x_1 > (z_1 + a_1)/2, x_1 + x_2 > (z_1 + z_2 + a)/2\}$$

respectively, where $I(\cdot)$ is the indicator function. From $x_1 + y_1 = z_1$ and $x_1 - y_1 = a_1$, we have $2x_1 = z_1 + a_1$. So, in the indicator above functions, $z_1 + a_1$ is an even number and $(z_1 + a_1)/2$ is an integer. Similarly, $(z_1 + z_2 + a)/2$ is an integer.

We would like to choose a reasonable combination of (a_1, a) depending on the values of z_1 and z_2 by controlling the type I error rate below a specified α^* level while maintaining the power as high as possible. Jung et al. [4] choose $a_1 = -1$ to stop the trial early when the number of responders in the experimental arm is smaller than the number of responders in the control arm. This rule is motivated by the single-arm two-stage designs stopping early when the number of responders from the experimental therapy is smaller than the expected number of responders under the null hypothesis [1, 7]. Note that, with a_1 fixed at -1 , the type I error rate $\alpha(z_1, z_2)$ decreases in a . Hence, we choose the smallest integer $a = a(z_1, z_2)$ for which the type I error rate is controlled below α^* , i.e.

$$a = \min\{a : \alpha(z_1, z_2) \leq \alpha^*\} \quad (1)$$

for each combination of $z_1 \in [0, 2n_1]$ and $z_2 \in [0, 2n_2]$.

2.3 How to choose n and n_1 :

Now, given (n_1, n) , we have a rule to choose rejection values (a_1, a) conditioning on (z_1, z_2) , so that, by selecting (n_1, n) , we completely define a two-stage design. At first, we choose a value of n large enough for a reasonable power, say $n \geq 20$. Then, the stage 1 sample size n_1 is selected in $[1, n_1]$. Given (n_1, n) , we choose $a_1 = -1$ and $a = a(z_1, z_2)$ so that the conditional type I error rate is controlled below α^* for any $z_1 \in [0, 2n_1]$ and $z_2 \in [0, 2n_2]$ as in (1). Hence, the marginal type I error rate is controlled below α^* as well.

Given n_l , Z_l has a probability distribution function:

$$g_l(z_l | p_x, p_y) = \sum_{x_l=m_{l-}}^{m_{l+}} \binom{n_l}{x_l} p_x^{x_l} q_x^{n_l-x_l} \binom{n_l}{z_l-x_l} p_y^{z_l-x_l} q_y^{n_l-z_l+x_l}$$

for $z_1 = 0, \dots, 2n_1$. By taking expectations over (z_1, z_2) , the marginal type I error rate and power are given as

$$\alpha = \sum_{z_1=0}^{2n_1} \alpha(z_1, z_2) g_1(z_1 | p_y, p_y) g_2(z_2 | p_y, p_y)$$

and

$$1 - \beta = 1 - \sum_{z_1=0}^{2n_1} \beta(z_1, z_2) g_1(z_1 | p_x, p_y) g_2(z_2 | p_x, p_y)$$

respectively. For a specified type I error rate α^* and power $1 - \beta^*$, a two-stage design defined by (n_1, n) and $\{a_1, a(z_1, z_2), 0 \leq z_1 \leq 2n_1, l = 1, 2\}$ is a candidate design if $1 - \beta \geq 1 - \beta^*$.

2.4 Minimax, optimal, and admissible designs:

Among the candidate designs, the one with the smallest maximal sample size n is called the minimax design.

A two-stage design with a futility stopping is to save the number of patients when H_0 is true. For a candidate two-stage design defined by (n_1, n) and $\{a_1, a(z_1, z_2), 0 \leq z_1 \leq 2n_1, l = 1, 2\}$, the probability of early termination (PET_0) under $H_0: p_x = p_y$ for a specified p_y for sample size calculation is obtained by

$$PET_0 = P(X_1 - Y_1 \leq a_1 | p_x = p_y) = \sum_{z_1=0}^{2n_1} g_1(z_1 | p_x = p_y) \sum_{x=m_1-}^{m_1+} f_1(x | z_1, 1) I\{x_1 \leq (z_1 + a_1)/2\}$$

Under H_0 , the sample size per arm is n_1 with probability PET_0 and n with probability $1 - PET_0$, so that the expected sample size is given by

$$EN_0 = n_1 \times PET_0 + n \times (1 - PET_0)$$

Given $(\alpha^*, 1 - \beta^*)$, the two-stage design with the smallest EN_0 is called the optimal design.

Jung and his colleagues [6, 7] define admissible designs for single-arm two-stage designs. Applying their concept to randomized trials, a candidate two-stage randomized phase II trial is admissible if it minimizes $w \times n + (1 - w) \times EN_0$ for any $w \in [0, 1]$. Since $w = 0$ corresponds to the optimal design and $w = 1$ corresponds to the minimax design, both minimax and optimal designs are admissible.

3. Two-Stage Design with Both Superiority And Futility Interim Tests

For a randomized phase II trial, we may want to save the number of patients when the experimental therapy is definitely efficacious as well as when it is inefficacious. To this end, we consider two-stage phase II clinical trial designs with an early stopping for both futility and superiority of the experimental arm compared to the control arm. A two-stage randomized phase II trial with early stopping values a_1 for futility and b_1 for superiority is carried out as follows.

Stage 1: Randomize n_1 patients to each arm, and observe numbers of responders X_1 and Y_1 .

1. If $X_1 - Y_1 \leq a_1$, reject the experimental arm and stop the trial.

2. If $X_1 - Y_1 \geq b_1$, accept the experimental arm and stop the trial.
3. Otherwise (i.e. $a_1 < X_1 - Y_1 < b_1$), proceed to stage 2.

Stage 2: Randomize n_2 patients to each arm, and observe numbers of responders X_2 and Y_2 . Let $Z_1 = X_1 + Y_1$, $Z_2 = X_2 + Y_2$.

1. Choose the second stage critical value a depending on z_1 and z_2 .
2. Reject the experimental arm if $X - Y \leq a$, where $X = X_1 + X_2$ and $Y = Y_1 + Y_2$.
3. Accept the experimental arm if $X - Y > a$ for further investigation.

For two-stage Fisher's test, the critical values $a_1 = a_1(z_1)$, $b_1 = b_1(z_1)$, and $a = a(z_1, z_2)$ will depend on the total numbers of responders from the two stages, z_1 and z_2 . Given input parameters $(p_x, p_y, \alpha^*, 1-\beta^*)$, a two-stage designs, defined by sample sizes (n_1, n_2) and critical values, $\{(a_1(z_1), a(z_1, z_2)), 0 \leq z_1 \leq 2n_1, 0 \leq z_2 \leq 2n_2\}$, is identified as follows.

3.1 How to find a_1 , b_1 , and a : Suppose that (n, n_1) are given. And, (z_1, z_2) are observed from the trial. For a two-stage design defined by (n_1, n_2) and $\{a_1(z_1), b_1(z_1), a(z_1, z_2) : 0 \leq z_l \leq 2n_l, l = 1, 2\}$, the conditional type I error rate and the conditional power function are calculated by

$$\begin{aligned} \alpha(z_1, z_2) &= P(X_1 - Y_1 \geq b_1 | z_1, \theta = 1) + P(a_1 < X_1 - Y_1 < b_1, X - Y > a | z_1, z_2, \theta = 1) \\ &= \sum_{x_1=m_{1-}}^{m_{1+}} f_1(x_1 | z_1, 1) I\{x_1 \geq (z_1 + b_1)/2\} \\ &\quad + \sum_{x_1=m_{1-}}^{m_{1+}} \sum_{x_2=m_{2-}}^{m_{2+}} f_1(x_1 | z_1, 1) f_2(x_2 | z_2, 1) \\ &\quad \times I\{(z_1 + a_1)/2 < x_1 < (z_1 + b_1)/2, x_1 + x_2 > (z_1 + z_2 + a)/2\} \end{aligned}$$

and

$$\begin{aligned} 1 - \beta(z_1, z_2) &= P(X_1 - Y_1 \geq b_1 | z_1, \theta) + P(a_1 < X_1 - Y_1 < b_1, X - Y > a | z_1, z_2, \theta) \\ &= \sum_{x_1=m_{1-}}^{m_{1+}} f_1(x_1 | z_1, \theta) I\{x_1 \geq (z_1 + b_1)/2\} \\ &\quad + \sum_{x_1=m_{1-}}^{m_{1+}} \sum_{x_2=m_{2-}}^{m_{2+}} f_1(x_1 | z_1, \theta) f_2(x_2 | z_2, \theta) \\ &\quad \times I\{(z_1 + a_1)/2 < x_1 < (z_1 + b_1)/2, x_1 + x_2 > (z_1 + z_2 + a)/2\} \end{aligned}$$

Given (n_1, n_2) , we want to choose a reasonable combination of (a_1, b_1, a) conditioning on the values of z_1 and z_2 that controls the type I error rate below a given α^* while maintaining the power as high as possible. It takes too much computing time to search for optimal values of (a_1, b_1, a) for every combination of (z_1, z_2) . Jung[8] claims that reasonable early stopping values a_1 and b_1 are determined around the expected differences under H_0 and H_1 , respectively, based on the results of two-stage binomial designs by Jung[3]. As a result, we choose $a_1 = -1$ and $b_1 =$

$[n_1(p_x - p_y)] + 1$ which do not depend on z_1 or z_2 , where $[c]$ denotes the round down of c . With a_1 and b_1 fixed at these values, we choose a as in the designs with a futility interim test only. Since the type I error rate $\alpha(z_1, z_2)$ decreases in a , we choose $a = a(z_1, z_2)$ to control the conditional type I error rate below α^* by

$$a = \min\{a : \alpha(z_1, z_2) \leq \alpha^*\} \quad (2)$$

for each combination of $z_1 \in [0, 2n_1]$ and $z_2 \in [0, 2n_2]$.

3.2 How to choose n and n_1 :

We increase n from a small number, like 20, until it has a high enough power. For a given n , we choose n_1 in $[1, n - 1]$ for a two-stage design. Given (n_1, n) , we $\{a(z_1, z_2), 0 \leq z_1 \leq 2n_1, 0 \leq z_2 \leq 2n_2\}$, select $a_1 = -1$ and $b_1 = [n_1(p_x - p_y)] + 1$, and $a = a(z_1, z_2)$ is chosen so that the conditional type I error does not exceed the specified α^* as in (2). So, for the selected sample sizes (n_1, n) , and critical values $a_1 = -1$, $b_1 = [n_1(p_x - p_y)] + 1$, and $\{a(z_1, z_2), 0 \leq z_l \leq 2n_l, l = 1, 2\}$, the marginal type I error rate the marginal power are calculated as

$$\alpha = \sum_{z_1=0}^{2n_1} \alpha(z_1, z_2) g_1(z_1 | p_y, p_y) g_2(z_2 | p_y, p_y)$$

and

$$1 - \beta = 1 - \sum_{z_1=0}^{2n_1} \beta(z_1, z_2) g_1(z_1 | p_x, p_y) g_2(z_2 | p_x, p_y)$$

respectively. Since the critical values are chosen so that the conditional type I error rates do not exceed α^* for all (z_1, z_2) outcomes, the marginal type I error rate is maintained below α^* too.

Given $(\alpha^*, 1 - \beta^*)$, a two-stage design defined by (n_1, n) , $a_1 = -1$, $b_1 = [n_1(p_x - p_y)] + 1$, and $\{a(z_1, z_2), 0 \leq z_l \leq 2n_l, l = 1, 2\}$ is a candidate design if $1 - \beta \geq 1 - \beta^*$.

3.3 Minimax, Optimal, and Admissible Two-Stage Designs: Among candidate two-stage designs, the design with the smallest maximal sample n is called the minimax design.

Recall that a two-stage design with both futility and superiority interim analyses is to save the number of patients when the experimental therapy is inefficacious (i.e. when H_0 is true) as well as when it is efficacious (i.e. when H_1 is true) compared to the control therapy. So, an optimal two-stage design will have small, expected sample sizes under both H_0 and H_1 . Since we cannot minimize the both expected sample sizes at the same time, we will minimize a weighted average of them. Under H_h ($h = 0, 1$), the probability of early termination is calculated by

$$\begin{aligned}
\text{PET}_0 &= P(X_1 - Y_1 \leq a_1 \text{ or } X_1 - Y_1 \geq b_1 | \theta = 1) \\
&= \sum_{z_1=0}^{2n_1} g_1(z_1 | p_y, p_y) \sum_{x=m_{1-}}^{m_{1+}} f_1(x_1 | z_1, 1) I\{x_1 \leq (z_1 + a_1)/2 \text{ or } x_1 \geq (z_1 + b_1)/2\} \\
\text{PET}_1 &= P(X_1 - Y_1 \leq a_1 \text{ or } X_1 - Y_1 \geq b_1 | \theta) \\
&= \sum_{z_1=0}^{2n_1} g_1(z_1 | p_x, p_y) \sum_{x=m_{1-}}^{m_{1+}} f_1(x_1 | z_1, \theta) I\{x_1 \leq (z_1 + a_1)/2 \text{ or } x_1 \geq (z_1 + b_1)/2\}
\end{aligned}$$

Hence, the expected sample size under H_h ($h = 0, 1$) is given by $EN_h = n_1 \times \text{PET}_h + n \times (1 - \text{PET}_h)$. We propose to minimize a weighted average of the expected sample sizes by the relative importance of type I error and type II error rates specified at the design stage. When we design a trial, we would select a small α^* if the control of false decision under H_0 is important and a small β^* if that under H_1 is important. So accounting for their relative importance, we consider a weighted average between the two expected sample sizes by

$$EN = \frac{EN_0/\alpha^* + EN_1/\beta^*}{1/\alpha^* + 1/\beta^*} = \frac{\beta^* EN_0 + \alpha^* EN_1}{\alpha^* + \beta^*}$$

The two-stage design with the smallest EN is called the optimal design. If $\alpha^* = \beta^*$, then our weighted expected sample size is identical to the simple average of the expected sample sizes that was used for single-arm two-stage trial designs [9, 10] regardless of the relative size between α^* and β^* .

Both minimizing maximal sample size and minimizing weighted mean expected sample sizes are good strategies to define good two-stage designs, but these two strategies do not get along sometimes. In other words, we may have to sacrifice expected sample size to minimize maximal sample size or we may have to sacrifice maximal sample size to minimize expected sample size. This happens since we cannot minimize both maximal sample size and expected sample size at the same time. In this case, we may consider minimizing a linear combination of maximal sample size and expected sample size as an effort to identify a good compromise between the minimax design and the optimal design. Using the concept of [6] and [7], we call a candidate two-stage randomized phase II trial an admissible design if it minimizes $w \times n + (1 - w) \times EN$ for any $w \in [0, 1]$. Since $w = 0$ corresponds to the optimal design and $w = 1$ corresponds to the minimax design, both minimax and optimal designs are admissible.

4 Numerical studies:

Tables 1 to 6 report the minimax and optimal designs and their type I error rates, powers, and weighted averages of expected sample sizes for various combinations of (p_x, p_y) under $(\alpha^*, \beta^*) = (0.1, 0.85), (0.1, 0.8), (0.15, 0.85), (0.15, 0.8), (0.2, 0.85), \text{ and } (0.2, 0.8)$. As stated earlier, $a_1 = -1$ and $b_1 = [n_1(p_x - p_y)] + 1$ for all the designs, but $a = a(z_1, z_2)$ will be dependent on (z_1, z_2) , so that we cannot report all the conditional second stage critical values in these tables. As expected, maximal sample size n decreases in α^* and increases in $1 - \beta^*$. For some design settings, the minimax design is identical to the optimal design, e.g. when $(\alpha^*, 1 - \beta^*, p_x, p_y) = (0.15, 0.8, 0.2, 0.05)$. Under each design setting,

the minimax design tends to conduct the interim analysis later (i.e. n_1/n is larger) than the optimal design. We have this result because the minimax design delays the interim analysis time to mimic the corresponding single-stage design, while the optimal design tries to conduct the interim analysis early to lower the weighted average between n_1 and n . Under each design setting, the maximal sample size of the minimax design is mostly identical to the sample size of the single-stage design as reported in Jung [3].

ρ_y	ρ_x	Maximal Design				Optimal Design			
		(n, n_1)	α	$1 - \beta$	EN	(n, n_1)	α	$1 - \beta$	EN
0.05	0.15	(95, 51)	0.0509	0.8002	77.99	(97, 43)	0.0527	0.8001	76.02
0.2	(52, 49)	0.0316	0.8006	50.83	(53, 25)	0.0348	0.8014	42.53	
0.25	(34, 22)	0.0218	0.8006	29.89	(36, 12)	0.0222	0.8042	29.06	
0.1	0.25	(68, 41)	0.0584	0.8025	56.99	(70, 31)	0.061	0.8001	53.48
	0.3	(43, 22)	0.05	0.8003	34.87	(44, 16)	0.0533	0.8001	33.87
0.15	0.3	(80, 57)	0.0674	0.8004	69.72	(84, 35)	0.0727	0.8001	62.77
	0.35	(50, 26)	0.0626	0.8007	40.23	(51, 24)	0.0652	0.8006	39.26
0.2	0.35	(91, 54)	0.0737	0.8007	74.51	(100, 35)	0.0793	0.8001	70.55
	0.4	(56, 31)	0.0675	0.8011	45.36	(59, 21)	0.071	0.8011	43.04
0.25	0.4	(99, 66)	0.0772	0.8001	83.2	(108, 40)	0.0813	0.8009	77.01
	0.45	(59, 36)	0.0694	0.8036	48.94	(62, 26)	0.0732	0.8014	46.32
0.3	0.45	(106, 106)	0.0777	0.8014	106	(118, 40)	0.0844	0.8006	81.63
	0.5	(60, 49)	0.0695	0.8	54.9	(60, 49)	0.0695	0.8	54.9
0.35	0.5	(110, 62)	0.0796	0.8007	87.23	(119, 47)	0.0848	0.8015	84.99
	0.55	(65, 34)	0.0773	0.8	50.28	(68, 28)	0.0799	0.8001	49.19
0.4	0.55	(111, 60)	0.0825	0.8009	87.06	(115, 55)	0.0843	0.8012	86.22
	0.6	(65, 35)	0.0789	0.8002	50.41	(68, 29)	0.0826	0.8001	49.15
0.45	0.6	(111, 61)	0.0864	0.8017	87.23	(115, 55)	0.0865	0.8012	86.11
	0.65	(65, 34)	0.0806	0.8	50.12	(68, 28)	0.0834	0.8001	48.97
0.5	0.65	(110, 62)	0.0871	0.8007	86.99	(119, 47)	0.0869	0.8015	84.57
	0.7	(60, 49)	0.0726	0.8	54.82	(60, 49)	0.0726	0.8	54.82
0.55	0.7	(106, 106)	0.0737	0.8014	106	(112, 47)	0.0859	0.8009	81.1
	0.75	(59, 36)	0.0751	0.8036	48.66	(66, 23)	0.0858	0.8002	45.51
0.6	0.75	(99, 66)	0.08	0.8001	82.8	(108, 40)	0.087	0.8009	75.96
	0.8	(56, 30)	0.0777	0.8008	44.67	(59, 21)	0.0816	0.8011	42.05
0.65	0.8	(91, 54)	0.0796	0.8007	73.92	(100, 35)	0.0871	0.8001	69.01
	0.85	(50, 26)	0.0739	0.8007	39.54	(55, 18)	0.0851	0.8029	37.87
0.7	0.85	(80, 57)	0.077	0.8004	69.24	(90, 29)	0.088	0.8006	61.2
	0.9	(43, 22)	0.0748	0.8003	33.92	(44, 16)	0.0779	0.8001	32.38
0.75	0.9	(68, 40)	0.0756	0.8022	55.89	(75, 21)	0.0856	0.8018	51.37
	0.95	(34, 22)	0.0677	0.8006	29.05	(36, 12)	0.0788	0.8042	26.46
0.8	0.95	(52, 49)	0.0652	0.8006	50.71	(57, 17)	0.0902	0.8002	39.65
0.85	0.95	(95, 51)	0.0772	0.8002	75.92	(102, 31)	0.0858	0.8018	71.53

Table 1: Minimax and Optimal designs for $(\alpha^*, 1-\beta^*) = (0.1, 0.8)$

ρ_v	ρ_x	Maximal Design				Optimal design			
		(n, n_1)	α	$1 - \beta$	EN	(n, n_1)	α	$1 - \beta$	EN
0.05	0.15	(109, 66)	0.0545	0.8502	91.16	(114, 47)	0.0573	0.8505	86.46
0.2	(61, 28)	0.0383	0.8528	49.52	(62, 26)	0.0401	0.8527	48.05	
0.25	(39, 22)	0.0252	0.8513	33.21	(40, 19)	0.0244	0.8521	32.51	
0.1	0.25	(79, 41)	0.0615	0.8519	63.66	(83, 32)	0.0668	0.8503	60.94
	0.3	(50, 24)	0.0556	0.8504	39.3	(53, 20)	0.0562	0.8545	39.12
0.15	0.3	(94, 56)	0.071	0.8502	77.28	(101, 39)	0.0754	0.8507	72.42
	0.35	(58, 31)	0.0644	0.8515	47	(61, 25)	0.0682	0.8507	44.88
0.2	0.35	(107, 61)	0.0755	0.8503	86.57	(115, 44)	0.0799	0.8511	82.01
	0.4	(64, 36)	0.0687	0.8504	52.13	(71, 24)	0.0755	0.8502	49.69
0.25	0.4	(117, 86)	0.0774	0.85	102.3	(124, 49)	0.0818	0.8506	89.17
	0.45	(71, 35)	0.0748	0.85	54.07	(74, 29)	0.0774	0.8519	53.14
0.3	0.45	(122, 74)	0.0785	0.8506	100.02	(135, 52)	0.0849	0.85	94.27
	0.5	(72, 45)	0.0741	0.85	59.25	(75, 35)	0.0777	0.8502	55.89
0.35	0.5	(131, 71)	0.0822	0.8501	102.17	(136, 59)	0.0847	0.8504	97.97
	0.55	(73, 42)	0.0745	0.8502	58.95	(74, 35)	0.0765	0.8514	56.95
0.4	0.55	(132, 71)	0.0821	0.8503	102.5	(138, 57)	0.0862	0.8503	98.52
	0.6	(73, 44)	0.0795	0.8502	59.33	(83, 30)	0.0842	0.8508	56.89
0.45	0.6	(132, 71)	0.0849	0.8503	102.42	(138, 57)	0.0892	0.8503	98.39
	0.65	(73, 42)	0.0823	0.8502	58.85	(83, 30)	0.0866	0.8526	56.81
0.5	0.65	(131, 71)	0.0853	0.8501	101.92	(136, 59)	0.0908	0.8504	97.57
	0.7	(72, 45)	0.083	0.85	59.06	(80, 28)	0.0821	0.8505	55.43
0.55	0.7	(122, 74)	0.0793	0.8506	99.74	(135, 52)	0.0864	0.85	93.47
	0.75	(71, 35)	0.0817	0.85	53.56	(74, 29)	0.0835	0.8519	52.42
0.6	0.75	(117, 86)	0.0823	0.85	102.05	(124, 49)	0.0856	0.8506	88.19
	0.8	(64, 35)	0.0726	0.8508	51.54	(71, 24)	0.0868	0.8502	48.49
0.65	0.8	(107, 60)	0.0797	0.8503	85.77	(115, 44)	0.0873	0.8511	80.62
	0.85	(58, 31)	0.0766	0.8515	46.4	(61, 25)	0.0837	0.8507	43.65
0.7	0.85	(94, 56)	0.0785	0.8502	76.58	(101, 39)	0.0871	0.8507	70.61
	0.9	(50, 24)	0.0778	0.8504	38.19	(53, 19)	0.0837	0.8543	37.45
0.75	0.9	(79, 40)	0.0773	0.8513	62.45	(83, 32)	0.0863	0.8503	58.76
	0.95	(39, 22)	0.0693	0.8513	32.15	(41, 11)	0.0777	0.85	30.12
0.8	0.95	(61, 27)	0.0745	0.8531	47.66	(62, 26)	0.0846	0.8527	45.46
0.85	0.95	(109, 66)	0.0773	0.8502	89.51	(114, 47)	0.0855	0.8505	82.88

Table 2: Minimax and Optimal designs for $(\alpha^*, 1 - \beta^*) = (0.1, 0.85)$

	p_x	Minimax Design				Optimal design			
		(n, n_1)	α	$1 - \beta$	EN	(n, n_1)	α	$1 - \beta$	EN
0.05	0.15	(78, 41)	0.0827	0.8001	64.27	(79, 35)	0.0849	0.8001	61.85
0.2	(44, 18)	0.0637	0.8031	34.81	(44, 18)	0.0637	0.8031	34.81	
0.25	(29, 14)	0.0475	0.8088	23.96	(29, 14)	0.0475	0.8088	23.96	
0.1	0.25	(55, 40)	0.0881	0.8	48.31	(57, 25)	0.0947	0.8003	43.31
	0.3	(36, 17)	0.0771	0.8014	28.92	(37, 14)	0.0837	0.8026	27.86
0.15	0.3	(65, 42)	0.1042	0.8001	55.05	(73, 26)	0.1144	0.8024	50.78
	0.35	(41, 22)	0.0925	0.8001	33.23	(46, 15)	0.1064	0.8002	31.97
0.2	0.35	(73, 50)	0.109	0.8001	62.47	(80, 30)	0.1191	0.8004	56.7
	0.4	(45, 27)	0.1021	0.8005	37.28	(48, 18)	0.1085	0.8032	34.76
0.25	0.4	(82, 44)	0.1189	0.8002	63.85	(85, 36)	0.1217	0.8007	61.84
	0.45	(47, 27)	0.1041	0.8007	38.17	(54, 20)	0.1187	0.8047	37.45
0.3	0.45	(85, 49)	0.1172	0.8004	67.98	(93, 37)	0.1263	0.8011	65.48
	0.5	(53, 22)	0.1169	0.8008	38.97	(56, 20)	0.1215	0.8013	38.06
0.35	0.5	(86, 54)	0.1187	0.8001	71.07	(98, 35)	0.1275	0.8009	67.46
	0.55	(54, 27)	0.1113	0.8015	41.66	(58, 19)	0.1249	0.8028	38.79
0.4	0.55	(86, 58)	0.1258	0.8001	72.22	(93, 41)	0.128	0.8012	68.05
	0.6	(54, 27)	0.1115	0.8	41.56	(58, 19)	0.1295	0.8009	38.61
0.45	0.6	(86, 58)	0.1298	0.8001	72.17	(93, 41)	0.1315	0.8012	67.95
	0.65	(54, 27)	0.1141	0.8015	41.53	(57, 20)	0.1345	0.8008	38.02
0.5	0.65	(86, 54)	0.13	0.8001	70.93	(98, 35)	0.1326	0.8009	67.07
	0.7	(53, 22)	0.1206	0.8008	38.63	(56, 20)	0.1347	0.8013	37.58
0.55	0.7	(85, 49)	0.1312	0.8004	67.66	(93, 37)	0.131	0.8011	64.87
	0.75	(47, 27)	0.1098	0.8007	37.88	(54, 20)	0.1306	0.8047	36.74
0.6	0.75	(82, 44)	0.1247	0.8002	63.31	(92, 32)	0.132	0.8023	60.92
	0.8	(45, 27)	0.1211	0.8005	36.92	(48, 18)	0.1214	0.8032	33.93
0.65	0.8	(73, 50)	0.1163	0.8001	62.1	(80, 30)	0.1282	0.8004	55.52
	0.85	(41, 22)	0.1125	0.8001	32.66	(46, 15)	0.1295	0.8002	30.54
0.7	0.85	(65, 42)	0.1149	0.8001	54.53	(73, 26)	0.1312	0.8024	49.05
	0.9	(36, 17)	0.111	0.8014	28.03	(37, 14)	0.1215	0.8026	26.45
0.75	0.9	(56, 27)	0.1161	0.802	43.82	(57, 25)	0.1226	0.8003	41.78
	0.95	(29, 13)	0.1093	0.8072	22.51	(31, 9)	0.1246	0.801	21.7
0.8	0.95	(44, 18)	0.1185	0.8031	32.61	(44, 18)	0.1185	0.8031	32.61
0.85	0.95	(78, 41)	0.1133	0.8001	62.53	(85, 27)	0.1322	0.8004	57.77

Table 224: Minimax and Optimal designs for $(\alpha^*, 1-\beta^*) = (0.15, 0.8)$

ρ_y	ρ_x	Minimax Design				Optimal Design			
		(n, n_1)	α	$1 - \beta$	EN	(n, n_1)	α	$1 - \beta$	
0.05	0.15	(91, 49)	0.0876	0.8501	73.09	(95, 40)	0.0912	0.8511	71.26
0.2	(51, 25)	0.068	0.8535	41.18	(53, 19)	0.0695	0.8533	40.3	
0.25	(34, 15)	0.0527	0.8592	27.06	(34, 15)	0.0527	0.8592	27.06	
0.1	0.25	(65, 36)	0.0954	0.8502	53.14	(70, 26)	0.101	0.8523	50.58
	0.3	(41, 24)	0.0822	0.8503	33.98	(45, 15)	0.091	0.8533	32.35
0.15	0.3	(78, 42)	0.1079	0.8501	62.52	(83, 33)	0.1135	0.8506	59.38
	0.35	(49, 22)	0.0992	0.8502	38.03	(51, 20)	0.1037	0.8536	36.96
0.2	0.35	(88, 54)	0.1127	0.8504	73.15	(93, 39)	0.1196	0.8511	67.03
	0.4	(52, 32)	0.1005	0.8502	43.45	(59, 20)	0.1146	0.8509	40.55
0.25	0.4	(94, 62)	0.1145	0.85	79.42	(106, 39)	0.1257	0.8506	72.81
	0.45	(59, 28)	0.1143	0.85	45.01	(67, 20)	0.1215	0.8517	44.04
0.3	0.45	(104, 57)	0.1241	0.8501	81.64	(117, 39)	0.1293	0.851	77.52
	0.5	(60, 34)	0.1105	0.8511	47.82	(60, 34)	0.1105	0.8511	47.82
0.35	0.5	(106, 66)	0.1184	0.8503	86.34	(114, 46)	0.1276	0.8506	79.47
	0.55	(61, 33)	0.1136	0.8513	48.04	(62, 28)	0.1183	0.8516	46.15
0.4	0.55	(107, 65)	0.1214	0.8504	86.47	(115, 46)	0.1312	0.8501	79.67
	0.6	(61, 34)	0.1197	0.8513	48.13	(66, 25)	0.1284	0.851	45.46
0.45	0.6	(107, 65)	0.1251	0.8504	86.41	(115, 46)	0.1351	0.8501	79.55
	0.65	(61, 33)	0.1229	0.8513	47.94	(66, 25)	0.1324	0.8525	45.41
0.5	0.65	(106, 66)	0.1256	0.8503	86.19	(108, 52)	0.1341	0.8507	79.92
	0.7	(60, 34)	0.1232	0.8511	47.62	(65, 25)	0.1325	0.8531	44.97
0.55	0.7	(104, 57)	0.1243	0.8501	81.31	(117, 40)	0.1315	0.8511	76.67
	0.75	(59, 28)	0.1237	0.85	44.61	(67, 20)	0.1285	0.8517	43.19
0.6	0.75	(94, 62)	0.1241	0.85	79.15	(106, 39)	0.1334	0.8506	71.86
	0.8	(52, 32)	0.1069	0.8502	43.15	(60, 19)	0.1304	0.8532	40.24
0.65	0.8	(88, 54)	0.1209	0.8504	72.73	(93, 39)	0.1284	0.8511	66.01
	0.85	(49, 22)	0.1178	0.8502	37.33	(51, 20)	0.1221	0.8536	35.91
0.7	0.85	(78, 42)	0.1207	0.8501	61.81	(83, 33)	0.1283	0.8506	57.97
	0.9	(41, 24)	0.1086	0.8503	33.38	(45, 14)	0.125	0.8513	30.95
0.75	0.9	(65, 36)	0.1143	0.8502	52.31	(70, 26)	0.1271	0.8523	48.71
	0.95	(34, 15)	0.1184	0.8592	25.48	(37, 9)	0.1276	0.8519	25.42
0.8	0.95	(51, 24)	0.1149	0.8534	39.62	(53, 20)	0.1261	0.8545	37.61
0.85	0.95	(91, 49)	0.1207	0.8501	71.18	(95, 40)	0.1269	0.8511	68.23

Table 4: Minimax and Optimal designs for $(\alpha^*, 1-\beta^*) = (0.15, 0.85)$

		Minimax Design				Optimal design			
	p_x	(n, n_1)	α	$1 - \beta$	EN	(n, n_1)	α	$1 - \beta$	
0.05	0.15	(65, 44)	0.11	0.8001	56.83	(68, 28)	0.1199	0.8002	51.66
0.2	(37, 18)	0.0803	0.8007	30.24	(39, 13)	0.0843	0.8008	29.53	
0.25	(25, 13)	0.0484	0.8048	21.23	(26, 10)	0.0514	0.8054	20.57	
0.1	0.25	(47, 23)	0.1242	0.8009	37.32	(49, 20)	0.1306	0.8012	35.91
	0.3	(30, 17)	0.1046	0.8007	25.19	(33, 10)	0.1169	0.8014	23.49
0.15	0.3	(54, 36)	0.1364	0.8001	46.17	(63, 20)	0.1567	0.8	41.9
	0.35	(34, 19)	0.1244	0.8002	27.46	(36, 15)	0.133	0.8048	26.43
0.2	0.35	(63, 31)	0.1544	0.8006	47.81	(67, 26)	0.161	0.8006	46.54
	0.4	(39, 19)	0.1456	0.8026	29.86	(42, 15)	0.1521	0.8028	29.03
0.25	0.4	(67, 38)	0.1548	0.8001	52.87	(77, 26)	0.1686	0.8	50.62
	0.45	(40, 24)	0.1393	0.8005	32.52	(46, 15)	0.1594	0.8002	30.55
0.3	0.45	(68, 50)	0.153	0.8002	59.41	(75, 32)	0.1661	0.801	52.96
	0.5	(41, 24)	0.1464	0.8003	32.88	(43, 17)	0.1536	0.8007	31.25
0.35	0.5	(69, 50)	0.164	0.8004	59.81	(78, 31)	0.1731	0.8006	53.93
	0.55	(41, 29)	0.1508	0.8002	35.26	(45, 16)	0.1631	0.803	32.15
0.4	0.55	(70, 47)	0.1712	0.8006	59.24	(70, 47)	0.1712	0.8006	59.24
	0.6	(42, 26)	0.1586	0.8037	34.88	(45, 16)	0.1697	0.8015	32.02
0.45	0.6	(70, 47)	0.1749	0.8006	59.21	(70, 47)	0.1749	0.8006	59.21
	0.65	(41, 29)	0.1615	0.8002	35.21	(45, 16)	0.1737	0.803	31.99
0.5	0.65	(69, 50)	0.1757	0.8004	59.72	(78, 31)	0.1797	0.8006	53.65
	0.7	(41, 24)	0.1623	0.8003	32.71	(43, 17)	0.1743	0.8007	30.96
0.55	0.7	(68, 50)	0.1728	0.8002	59.27	(75, 32)	0.1825	0.801	52.5
	0.75	(40, 24)	0.1634	0.8005	32.27	(46, 15)	0.1744	0.8002	29.9
0.6	0.75	(67, 38)	0.1738	0.8001	52.46	(77, 26)	0.1748	0.8	49.74
	0.8	(39, 19)	0.165	0.8026	29.36	(42, 14)	0.1677	0.8016	28.29
0.65	0.8	(63, 31)	0.1646	0.8006	47.14	(67, 26)	0.1745	0.8006	45.55
	0.85	(34, 19)	0.1458	0.8002	26.97	(36, 14)	0.1588	0.8031	25.56
0.7	0.85	(54, 36)	0.1538	0.8001	45.75	(64, 19)	0.1752	0.8013	40.9
	0.9	(30, 17)	0.1432	0.8007	24.66	(33, 9)	0.1624	0.8036	22.16
0.75	0.9	(47, 23)	0.1562	0.8009	36.36	(50, 19)	0.1668	0.8041	34.91
	0.95	(25, 12)	0.1381	0.8056	20.2	(26, 10)	0.1578	0.8054	18.83
0.8	0.95	(37, 18)	0.1512	0.8007	28.83	(39, 13)	0.1649	0.8008	26.99
0.85	0.95	(65, 44)	0.1451	0.8001	55.93	(68, 28)	0.1642	0.8002	49.06

Table 5: Minimax and Optimal designs for $(\alpha^*, 1-\beta^*) = (0.2, 0.8)$

ρ_v	ρ_x	Minimax Design				Optimal design			
		(n, n_1)	α	$1 - \beta$	EN	(n, n_1)	α	$1 - \beta$	EN
0.05	0.15	(78, 46)	0.1139	0.85	65.08	(84, 30)	0.1278	0.8513	60.57
0.2	(44, 23)	0.0923	0.8552	36.72	(48, 13)	0.0995	0.8516	35	
0.25	(29, 15)	0.06	0.8518	23.79	(31, 10)	0.0649	0.8568	23.68	
0.1	0.25	(56, 31)	0.1314	0.8506	45.43	(62, 20)	0.14	0.8506	42.87
	0.3	(35, 18)	0.1112	0.8507	28.4	(37, 15)	0.1183	0.8547	27.63
0.15	0.3	(65, 39)	0.1431	0.8505	52.96	(72, 26)	0.1539	0.8504	50.08
	0.35	(42, 21)	0.1334	0.8516	33.89	(46, 15)	0.1421	0.8541	31.77
0.2	0.35	(74, 46)	0.1534	0.8504	60.57	(80, 33)	0.1617	0.8513	56.74
	0.4	(45, 25)	0.1406	0.8506	35.6	(47, 20)	0.1465	0.8528	34.16
0.25	0.4	(78, 54)	0.1543	0.8509	67.32	(91, 33)	0.1682	0.8513	61.38
	0.45	(46, 30)	0.1422	0.8502	38.37	(54, 20)	0.1604	0.8529	37.33
0.3	0.45	(87, 64)	0.1573	0.85	76.1	(93, 38)	0.1674	0.8502	65.62
	0.5	(52, 24)	0.1576	0.8505	38.64	(57, 20)	0.162	0.8517	38.46
0.35	0.5	(89, 59)	0.1568	0.8505	74.19	(95, 39)	0.1703	0.8506	66.24
	0.55	(54, 23)	0.1617	0.8518	39.46	(54, 23)	0.1617	0.8518	39.46
0.4	0.55	(89, 70)	0.1598	0.8502	79.98	(97, 39)	0.1782	0.8521	66.96
	0.6	(54, 23)	0.1667	0.8503	39.35	(59, 20)	0.1696	0.8525	39.03
0.45	0.6	(89, 70)	0.1635	0.8502	79.96	(97, 39)	0.1822	0.8521	66.86
	0.65	(54, 23)	0.1704	0.8518	39.33	(58, 20)	0.1721	0.8504	38.51
0.5	0.65	(89, 59)	0.1659	0.8505	74.08	(95, 39)	0.1819	0.8506	65.96
	0.7	(52, 24)	0.1705	0.8505	38.4	(57, 20)	0.1721	0.8517	38.09
0.55	0.7	(87, 60)	0.1625	0.8501	73.47	(92, 40)	0.1793	0.8517	64.79
	0.75	(46, 30)	0.1573	0.8502	38.18	(54, 20)	0.1663	0.8529	36.8
0.6	0.75	(78, 54)	0.155	0.8509	67.13	(91, 33)	0.1728	0.8513	60.6
	0.8	(45, 26)	0.1654	0.8506	36.95	(47, 19)	0.1602	0.8508	33.59
0.65	0.8	(74, 46)	0.1667	0.8504	60.17	(80, 33)	0.1754	0.8513	55.87
	0.85	(42, 21)	0.154	0.8516	33.43	(46, 15)	0.1686	0.8541	30.69
0.7	0.85	(65, 39)	0.1567	0.8505	52.42	(72, 26)	0.1726	0.8504	48.82
	0.9	(35, 18)	0.1431	0.8507	27.79	(37, 14)	0.1578	0.8519	26.7
0.75	0.9	(56, 31)	0.1575	0.8506	44.68	(63, 19)	0.1714	0.8515	41.78
	0.95	(29, 14)	0.146	0.8506	22.78	(31, 10)	0.1615	0.8568	21.72
0.8	0.95	(44, 20)	0.1576	0.8512	32.93	(48, 13)	0.1685	0.8516	32.06
0.85	0.95	(78, 46)	0.1542	0.85	63.87	(84, 30)	0.1709	0.8513	57.5

Table 6: Minimax and Optimal designs for $(\alpha^*, 1-\beta^*) = (0.2, 0.85)$

5. Discussion:

We have presented two-stage randomized phase II design based on Fisher exact tests with both futility and superiority early stopping options, so that we can save the number of patients when the experimental arm is futile as well as when it is efficacious compared to a parallel control arm. We usually want about 40 to 60 patients per arm for randomized phase II trials, so that we use a high α level, such as 1-sided $\alpha = 10\%$ to 20% , rather than a standard 2-sided $\alpha = 0.05$ which is a standard level for phase III trials.

In this paper, we propose to choose stopping values of stage 1, a_1 and b_1 , that are free of the total number of responders z_1 to lower the computing burden. We also have tried to optimize the designs by considering a_1 and b_1 dependent on z_1 , but it increased the computation time seriously without lowering the weighted average of expected sample sizes much. Although we present minimax and optimal designs under various design settings in Tables 1-6, readers cannot use them for their own trials since these tables do not report the rejection value of the second stage a since each design will have a long list of rejection values depending on (z_1, z_2) . Furthermore, these tables do not show the admissible designs since there can be multiple admissible designs under each design setting. We have developed a user-friendly computer program to help us find minimax, optimal and admissible designs while giving the list of conditional critical values for each selected design. This program is available to the readers upon request.

We have considered balanced randomization cases. We can easily extend our method for unbalanced randomized trials. Let n_{1k} denote the number of patients for stage 1 ($k = 1, 2$) randomized to arm $k (= x, y)$. For a chosen $\gamma \in (0, 1)$, we will have $n_{1x}/(n_{1x} + n_{1y}) = \gamma$. In this case, we may consider a two-stage design as follows.

Stage 1: Randomize n_{1x} and n_{1y} patients to arms x and y , respectively, and observe number of responders X_1 and Y_1 .

1. If $X_1/n_{1x} - Y_1/n_{1y} \leq a_1$, reject the experimental arm and stop the trial.
2. If $X_1/n_{1x} - Y_1/n_{1y} \geq b_1$, accept the experimental arm and stop the trial.
3. Otherwise (i.e. $a_1 < X_1/n_{1x} - Y_1/n_{1y} < b_1$), proceed to stage 2.

Stage 2: Randomize n_{2x} and n_{2y} patients to arms x and y and observe number of responders X_2 and Y_2 . Let $Z_1 = X_1 + Y_1$, $Z_2 = X_2 + Y_2$.

1. Choose the first stage critical value a depending on z_1 and z_2 .
2. Reject the experimental arm if $X/(n_{1x} + n_{2x}) - Y/(n_{1y} + n_{2y}) \leq a$, where $X = X_1 + X_2$ and $Y = Y_1 + Y_2$.
3. Accept the experimental arm if $X/(n_{1x} + n_{2x}) - Y/(n_{1y} + n_{2y}) > a$ for further investigation.

For an unbalanced randomization trial ($\gamma \neq 1/2$), we may choose the stage 1 critical values a_1 and b_1 to be the expected values of the sample proportions under H_0 and H_1 , respectively.

References

1. Simon, R.: Optimal two-stage designs for phase II clinical trials. *Controlled Clinical Trials* 10 (1989): 1–10.

2. Cannistra, S.A.: Phase II trials in journal of clinical oncology. *Journal of Clinical Oncology* 27 (2009): 3073–3076.
3. Jung, S.H.: Randomized phase II trials with a prospective control. *Statistics in Medicine* 27 (2008): 568–583.
4. Jung, S.H., Sargent, D.J.: Randomized phase II clinical trials. *Journal of biopharmaceutical statistics* 24 (2014): 802–816.
5. Fisher, R.A.: The logic of inductive inference. *Journal of the Royal Statistical Society* 98 (1935): 39–82.
6. Jung, S.H., Lee, T., Kim, K., George, S.L.: Admissible two-stage designs for phase II cancer clinical trials. *Stat Med* 23 (2004): 561–569.
7. Jung, S.H., Carey, M., Kim, K.M.: Graphical search for two-stage designs for phase II clinical trials. *Controlled Clinical Trials* 22 (2001): 367 – 372.
8. Jung, S.H.: Randomized phase II clinical trials. CRC Press LLC pp. 147–174 (2013).
9. Chang, M.N., Therneau, T.M., Wieand, H.S., Cha, S.S.: Designs for group sequential phase ii clinical trials. *Biometrics* 43 (1987): 865–874.
10. Therneau, T.M., Wieand, H.S., Chang, M.: Optimal designs for a grouped sequential binomial trial. *Biometrics* 46(1990): 771–781.



This article is an open access article distributed under the terms and conditions of the [Creative Commons Attribution \(CC-BY\) license 4.0](https://creativecommons.org/licenses/by/4.0/)