Journal of Biostatistics and Biometric Applications
Volume 7| Issue 1

Research Article Open Access

A Java Software for Randomized Phase II Clinical Cancer Trial Designs

Cao S[1], Grayling MJ[2], Jung SH[1*]

[1]Department of Biostatistics and Bioinformatics, Duke University, Durham, NC 27710, USA
[2]Population Health Sciences Institute, Newcastle University, Newcastle upon Tyne, UK

[*]**Corresponding author:** Jung SH, Department of Biostatistics and Bioinformatics, Duke University, Durham, NC 27710, USA. Tel: 9199289696, Email: sinho.jung@duke.edu

Citation: Cao S, Grayling MJ, Jung SH (2022) A Java Software for Randomized Phase II Clinical Cancer Trial Designs. J Biostat Biometric App 7(1): 102

Abstract

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 historicalcontrol. Due to a small sample size and heterogeneity of patient population, the patient 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 may lead to biased conclusions. A randomized phase II trial can resolve such problems by randomizing patients between an experimental arm and a control arm. We proposea software package for designing and analyzing randomized phase II trials.

We develop a user-friendly Java software that will help us find optimal two-stage phase II trial designs. Although the programs accommodate trial designs based onvarious statistical methods and different types of early stopping rules, the main partof our paper is focused on randomized phase II trials based on Fisher's exact test with futility and superiority early stopping values. If users enter input parameter values,the software generates a graphical output displaying all efficient two-stage designs. Minimax, optimal, and admissible designs are highlighted as good designs, but userscan select any of the displayed designs. When the circle representing a design is clicked, all the specifics of the selected design are displayed. Fisher's test is an exact method whose critical values depend on the total number of responders from two arms. So, the computations required to search for optimal randomized multi-stage phase II trial designs based on Fisher's exact test is very heavy. By using efficientalgorithms, our software provides output at almost real time speed.

Keywords: Admissible Design, Futility Stopping, Minimax Design, Optimal Design, Superiority Stopping

Introduction

A phase II cancer clinical trial is to evaluate an experimental therapy using a short- term efficacy outcome, such as tumor response, before proceeding to a large scale phase III clinical trial with a long-term outcome. Traditionally, the single-arm phase II study 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 [9].A historical control is often taken from a previous phase II trial with a small numberof patients. In this case, the patient characteristics can be easily different between the historical study and the new study, so that the experimental therapy and the chosen historical control therapy may not be comparable. More often than not, an appropriate historical control does not exist. Pointing out various pitfalls of single- arm phase II trial designs, Cannistra [1] recommends randomized phase II trials as a viable alternative.

Due to the small sample sizes, we need exact statistical methods to design and analyze phase II trials. Furthermore, exact methods usually require heavy computing. These problems have hindered clinical trialists from adopting randomized designs for phase II cancer clinical trials. Jung[5] proposes a design method for randomized phaseII trials, called MaxTest design, based on a binomial test. As an alternative to thebinomial test, Jung et al.[8] propose to use Fisher's exact test[4] to eliminate nuisance parameters from the null hypothesis. These methods consider an early stopping dueto futility. Cao et al.[2] extends the latter method to allow both superiority and futility interim tests. In this paper, we present a user-friendly Java software that identifies good randomized phase II trial designs and visually displays them. Using the interactive functions, it provide all the detailed information for each selected design. Using efficient algorithms, the software carries out the complicated computing within a short time period.

Methods

Two-stage Design Using Fisher's Exact Test

We consider a two-stage phase II clinical trial with an early stopping rule when the experimental arm, called arm x, is found to have a low efficacy (futility stopping) or a high efficacy (superiority stopping) compared to the control arm, called arm y. Let p_X and p_Y denote the response rates of arms x and y, respectively, and $q_k = 1 - p_k$ for k = x, y. For the odds ratio defined as $\theta = p_X q_Y / (p_Y q_X)$, we want to test $H_0 : \theta = 1$ against $H_1 : \theta = \theta_1 (> 1)$ using two-stage Fisher's exact test. At the design stage, we specify (p_X, p_Y) under H_1 and type I error rate $\alpha^*$and power $1 - \beta^*$. Note thatthe response rate for the control arm, p_Y, is common under H_0 and H_1. Let X_k andY_k denote the number of responders from arms x and y, respectively, during stage $k(= 1, 2)$.

A two-stage balanced randomized phase II trial with early stopping values a_1 forfutility 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 $X_1 - Y_1 \le a_1$, reject arm x and stop the trial.
2. If $X_1 - X_1 \ge b_1$, accept arm x 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 $Z_1 = X_1 + Y_1$, $Z_2 = X_2 + Y_2$, $X = X_1 + X_2$, and $Y = Y_1 + Y_2$.

1.Choose the second stage rejection value a depending on $Z_1 = z_1$ and $Z_2 = z_2$.
2. Reject arm x if $X - Y \le a$.

3. Accept arm x if $X - Y > a$ for further investigation.

Cao et al.[2] consider stopping the trial early if $X_1 - Y_1$ is smaller than the difference in the expected number of responders under H_0, i.e. $a_1 = -1$ or larger than the difference in the expected number of responders under H_1, i.e. $b_1 = [n_1(p_X - p_Y)]+1$, where $[c]$ denote the round down of c. The stage 2 rejection value $a = a(z_1, z_2)$ is calculated conditional on Z_1 and Z_2. With a_1 and b_1 fixed at these values and for aselected integer a, the conditional type I error rate for given Z_1 and Z_2 is given as

$$\alpha(z_1, z_2) = P(X_1 - Y_1 \geq b_1 | z_1, H_0) + P(a_1 < X_1 - Y_1 < b_1, X_1 + X_2 - Y_1 - Y_2 > a | z_1, z_2, H_0)$$

which is calculated based on two independent random variables X_1 and X_2 with hypergeometric distributions for given z_1 and z_2, refer to Cao et al.[2]. Hence, $a(z_1, z_2)$is selected by the smallest integer maintaining the conditional type I error rate below α^*, i.e.

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

Given (n_1, n_2), and using critical values (a_1, b_1, a), conditional and marginal powerare calculated by

$$1 - \beta(z_1, z_2) = P(X_1 - Y_1 \geq b_1 | z_1, H_1) + P(a_1 < X_1 - Y_1 < b_1, X_1 + X_2 - Y_1 - Y_2 > a | z_1, z_2, H_1)$$

and

$$1 - \beta = E\{1 - \beta(Z_1, Z_2) | H_1\}$$

respectively, where the expectation is taken with respect to two independent con- volutions of binomial random variables $Z_l = X_l + Y_l$, for $l = 1, 2$. A design defined by sample sizes (n_1, n_2) and critical values $\{a_1, b_1, a(z_1, z_2), z_l \in [0, 2n_l], l = 1, 2\}$ is called a candidate design if $1 - \beta \geq 1 - \beta^*$.

Minimax, Optimal, and Admissible Designs

Among the candidate designs, the one with the smallest maximal sample size $n = n_1 + n_2$ is called the minimax design.

Simon[9] proposes optimal designs for single-arm two-stage design with futility stopping only by minimizing the expected sample size under H_0. With both futility and superiority early stopping, we have to consider minimizing the expected sample size under both H_0 and H_1. The probabilities of early termination under H_h $(h = 0, 1)$ is calculated by $\text{PET}_h = P(X_1 - Y_1 \leq a_1$ or $X_1 - Y_1 \leq b_1 | H_h)$. This marginal probability is calculated directly from two independent binomial random variables X_1 and Y_1 since the early stopping values (a_1, b_1) do not depend on z_1. The expected sample size under H_h is calculated by

$$\text{EN}_h = n_1 \times \text{PET}_h + n \times (1 - \text{PET}_h)$$

For randomized two-stage designs with both futility and superiority early stopping, Cao[2] et al. propose to minimize the weighted expected sample sizes under H_0 andH_1 using the weights of inverse type I and II error rates to account for the relative importance between type I and II errors. That is, among the candidate two-stage designs, the *optimal* design has the smallest EN_w

$$\text{EN}_w = \frac{\text{EN}_0/\alpha^* + \text{EN}_1/\beta^*}{1/\alpha^* + 1/\beta^*} = \frac{\beta^* \times \text{EN}_0 + \alpha^* \times \text{EN}_1}{\alpha^* + \beta^*}$$

our software also considers minimizing the expected sample size with equal weights

$$\mathrm{EN} = \mathrm{EN}_0/2 + \mathrm{EN}_1/2$$

that was proposed by Chang et al.[3] and Therneau et al.[10] for single-arm phase II trials.

Jung and colleagues[6] and [7] define *admissible* designs for single-arm two-stage designs as those with the smallest weighted average between the maximal sample size and the expected sample size for any weights. Applying their concept to randomizedtwo-stage trials, a candidate two-stage randomized phase II trial is *admissible* if it minimizes $w \times n + (1-w) \times \mathrm{EN}_0$ 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.

Algorithm of Software

We propose efficient algorithms to find good designs (n, n_1) for given input param-eters $(\alpha^*, \beta^*, p_X, p_Y)$. For each n, we first find the optimal n_1 (i.e. the one with the smallest expected sample size) in Algorithm 1.

Given (n_1, n), we have

$$\begin{aligned} &P(X_1 - Y_1 > a_1 | H_1) = P(a_1 < X_1 - Y_1 < b_1 | H_1) + P(X_1 - Y_1 \geq b_1 | H_1) \\ &\geq P(a_1 < X_1 - Y_1 < b_1, X - Y > a(z_1, z_2) | H_1) + P(X_1 - Y_1 \geq b_1 | H_1) \\ &= 1 - \beta(z_1, z_2) \end{aligned}$$

for any $z_1 (\in [0, 2n_1])$ and $z_2 (\in [0, 2n_2])$, so that the marginal power $1-\beta$ has anupper bound of $P(X_1 - Y_1 > a_1 | H_1)$ which is free of (z_1, z_2). So, for given n, we can skip the procedure for a selected n_1 if $P(X_1 - Y_1 \leq a_1) > \beta^*$ (i.e. $1-\beta < 1-\beta^*$) which does not have to go through the calculations for all combinations of (z_1, z_2). Since the calculation of EN_w does not depend on (z_1, z_2) either, we can calculate EN_w before iterating through z_1 and z_2 either. So, if EN_w is greater than the recorded minimum value of EN_w, we can skip the complicated combinations requiring the iterations through $z_l \in [0, 2n_l]$ for $l = 1, 2$. This computational saving becomes available by choosing the stage 1 stopping values (a_1, b_1) free of z_1. In Algorithm 1, $g_l(z_l | p_X, p_Y)$ denote the probability mass function of the convolution $Z_l = X_l + Y_l$ for two independent random variables, $X_l \sim b(n_l, p_X)$ and $Y_l \sim b(n_l, p_Y)$.

Furthermore, Algorithm 2 finds a reasonable range of n that leads to good designs. The minimum of the range is the value of n for the Minimax design. To find this value, we start iterating at the required sample size for a single-stage design as proposed by Jung et al. [8], denoted as n_{single}. Specifically, we first decrease n by 1 starting from n_{single} until we find the Minimax design, then increase n by 1 starting from n_{single} until a two-stage design satisfies either $EN_w > n_{single}$, or $n > N_{max}$, which isa given upper limit of the sample size. If a two-stage design's expected sample size isgreater than the required sample size for a single-stage design (i.e. $EN_w > n_{single}$), this design is not useful as it does not save the expected sample size by allowing foran early stopping. Therefore, we can stop the searching procedure early if we find such a design, even if we have not reached the upper limit of the sample size.

Algorithm 1 Optimal Design For Fixed n

procedure OPTIMAL FIXEDN(n)
  $(n1, power, EN_W) \leftarrow (0, 0, n)$
  for $k = 0$ **to** n **do**
    $B = P(X1 - Y1 \geq a1)$
    $E^{-}N_W = (\beta^* \times EN0 + \alpha^* \times EN1)/(\alpha^* + \beta^*)$
    if $B \leq \beta$ **and** $E^{-}N_W < EN_W$ **then**
      $p\bar{o}wer = 0$
      for $z1 = 0$ **to** $2n1$ **do**
        for $z2 = 0$ **to** $2n2$ **do**
          choose a depending on $z1, z2$
          $p\bar{o}wer$ += Conditional power $\times g1(z1|p_X, p_Y)g2(z2|p_X, p_Y)$
    end for
      end for
      if $p\bar{o}wer \geq 1 - \beta$ **then**
        $(n1, power, \bar{EN}_W) = (k, p\bar{o}wer, \bar{EN}_W)$
      end if
    end if
  end for
  return $(n1, power, EN0)$
end procedure

Algorithm 2 Randomized Two-stage Designs Based on Fisher's Exact Test

procedure TWO STAGE DESIGN($p_X, p_Y, \alpha, \beta, Nmax$)
  $n_{single} \leftarrow$ Optimal sample size of a single stage design
  $Designs \leftarrow$ An empty list of designs
  for $n = n_{single}$ **to** 1 **do**
    $d \leftarrow$ Find Best Design for Fixed n
    if $power < 1 - \beta$ for d **then**
      break
    end if
    Add d to $Designs$
  end for
  for $n = n_{single} + 1$ **to** N_{max} **do**
    $d \leftarrow$ Find Best Design for Fixed n
    Add d to $Designs$
    if $EN_0 > n_{single}$ for d **then**
      break
    end if
  end for
  return $Designs$
end procedure

Result

We demonstrate the results of our project with an example. We found the optimal design under a hypothetical setting. Suppose that the control treatment is known to have a response rate of py = 0.3. We will be interested in the experimental therapy if its response rate is $p_X = 0.5$ or higher. We want to maintain the type I error rate below $\alpha^* = 0.15$ and the power above 1 $\beta^* = 0.8$. Figure 1 is the snapshot of our softwareto specify these input parameter values. We set the upper limit of maximal sample size at $N = 65$. Figure 1 For each maximal sample size n, our software identifies the n_1 value with the smallest expected sample size EN_w, and generates a plot of the minimal EN_w against n as shown in Figure 2 for the design setting. The minimax design comes the leftmost with $n = 52$ and the optimal design comes at the bottomwith $n = 56$ of the figure. The design with $n = 53$ is an admissible design since it lies on the convex hull connecting between the minimax design and the optimal design.

Figure 1: Screenshot of the input values (p_X, p_Y, α^*, β^*, N) = (0.5, 0.3, 0.15, 0.2, 65) andselection of weighted expected sample size

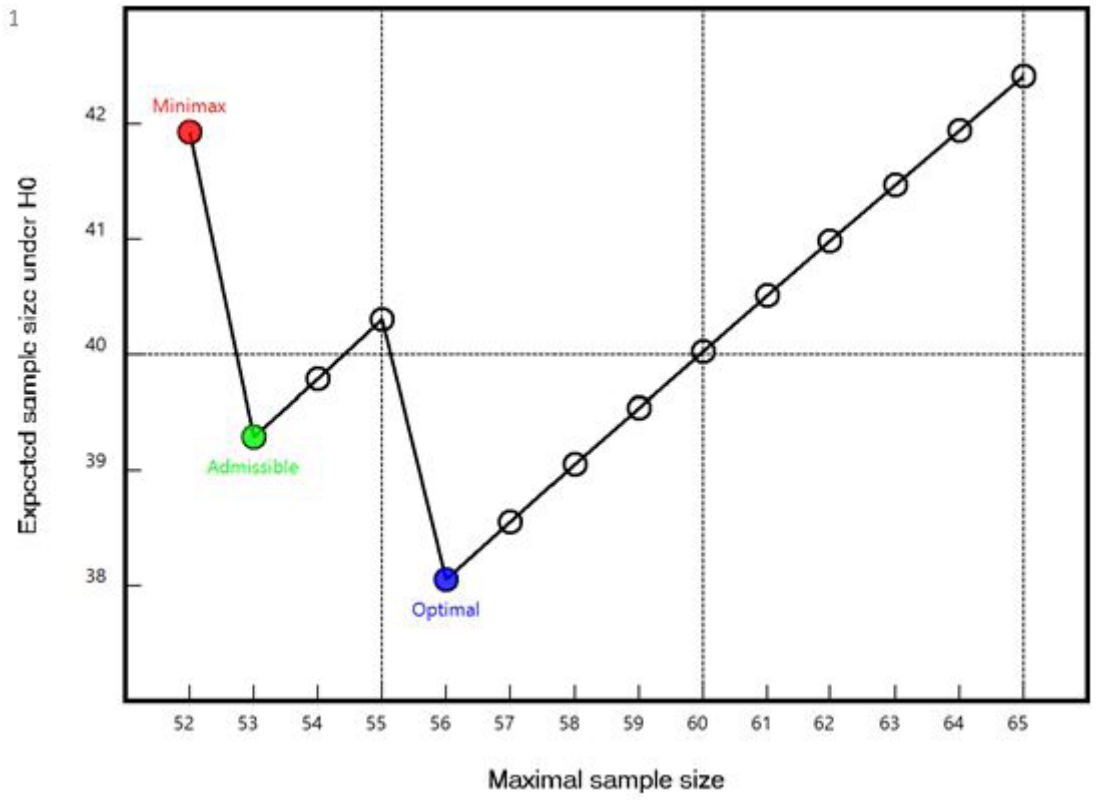

Figure 2: Minimal EN_w against n: output under $(p_X, p_Y, \alpha^*, \beta^*, N)$=$(0.5, 0.3, 0.15, 0.2, 65)$ and weighted sample size

If the circle of a design is clicked, our software displays all the details of the design including $(n, n_1, EN_w, EN_0, EN_1, a_1, b_1)$ and $\{a(z_1, z_2) : 0 \leq z_l \leq 2n_l, l = 1, 2\}$. Table1 gives the details of the optimal design. A table presenting the critical values $a(z_1, z_2)$ will be shown in a new panel when [Show critical value table] button is clicked, but isnot reported in this paper because of its big size to cover all combinations of (z_1, z_2) values.

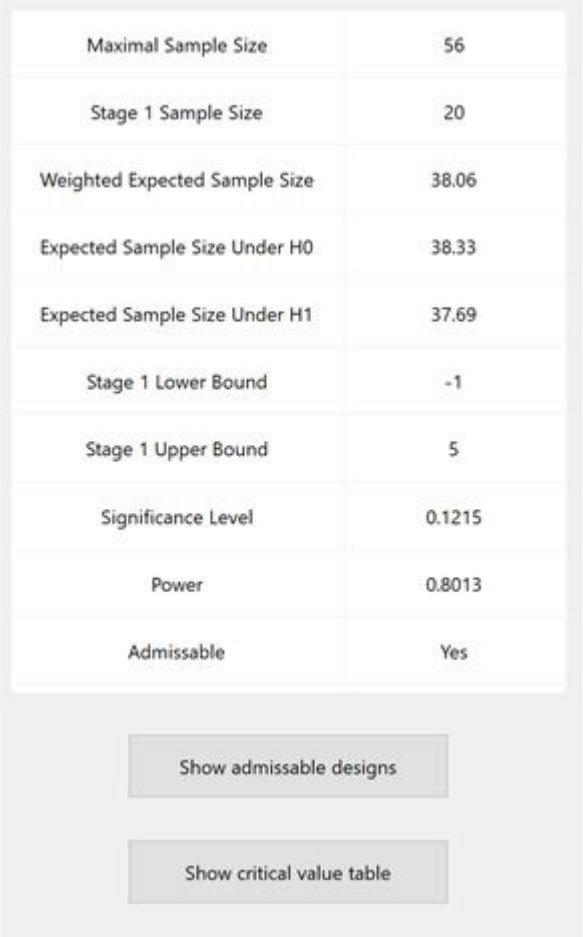

Table 1: Specifications of the optimal design for $(p_X, p_Y, \alpha^*, \beta^*, N)=(0.5, 0.3, 0.15, 0.2, 65)$ and weighted expected sample size

Discussion

We present a user-friendly software for randomized phase II clinical trials based on two-stage Fisher's exact tests. While Fisher's test is a conditional test, we need to cal- culate the marginal distribution of all candidate designs to search for efficient designsat the trial design stage. So, the computation for trial design can be extremely heavy.By using stage 1 stopping values that do not depend on the conditioning randomvariable (i.e. the total number of responders) and efficient computing algorithm, we could skip a lot of heavy computation procedures and shorten the overall computing time. Furthermore, our software generates graphical output for given design settings and interactively displays the details of designs selected by users.

In summary, our software provides optimal designs minimizing the weighted av- erage of EN_0 and EN_1 by the inverse of type I and II error rates as well as the simple mean expected sample size. It also has an option of two-stage designs with futility stopping only that was proposed by Jung and Sargent[8]. In addition to these two-stage designs based on Fisher's exact test, it also provides designs for MaxTest which is based on two-sample binomial test. Our software also supports single-stage designs for Fisher's exact test and MaxTest. Java language was used for computation and graphical display of designs of our software.

Acknowledgements

This research did not receive any specific grant from funding agencies in the public,commercial, or not-for-profit sectors.

References

1. Cannistra SA (2009) Phase II trials in journal of clinical oncology. Journal of Clinical Oncology 27:3073-76.

2. Cao S, Liu L, Jung SH (2021) Randomized phase ii clinical trials using fisher's exact test. Archives of Clinical and Biomedical Research 5, 214-229.

3. Chang MN, Therneau TM, Wieand HS, Cha SS (1987) Designs for group sequential phase ii clinical trials. Biometrics, 43: 865-874.

4. Fisher RA (1935) The logic of inductive inference. Journal of the Royal Statistical Society 98:39-82.

5. Jung SH (2008) Randomized phase II trials with a prospective control. Statistics in Medicine 27:568–83.

6. Jung SH, Carey M, Kim KM (2001) Graphical search for two-stage designs for phase II clinical trials. Controlled Clinical Trials, 22:367-72.

7. Jung SH, Lee T, Kim K, George SL (2004) Admissible two-stage designs for phase II cancer clinical trials. Stat Med. 23:561-69.

8. Jung SH, Sargent DJ (2014) Randomized phase II clinical trials. Journal of biopharmaceutical statistics, 24:802-16.

9. Simon R (1989) Optimal two-stage designs for phase II clinical trials. Control Clin Trials 10:1-10.

10. Therneau TM, Wieand HS, Chang M (1990) Optimal designs for a grouped sequential binomial trial. Biometrics 46:771-81.