

Chapter 130

Three-Stage Phase II Clinical Trials

Introduction

Phase II clinical trials determine whether a drug or regimen has sufficient activity against disease to warrant more extensive study and development. In a three-stage design, the patients are divided into three groups or stages. At the completion of the first stage, an interim analysis is made to determine if the second stage should be conducted. If the number of patients responding is greater than a certain amount, the second stage is conducted. Otherwise, it is not. A similar interim analysis is conducted at the end of the second stage.

This module finds designs that meet the error rate (alpha and beta) criterion and minimize the expected sample size. The formulation is given in Chen (1997). Extending Chen's work, our algorithm allows the investigation of near-optimal designs that may have other useful properties.

Technical Details

Phase I clinical trials are designed to provide information about the maximum tolerated dose levels of a treatment. They consist of three to six patients at each dose level and provide little information about the effectiveness of the treatment.

Phase II trials obtain initial estimates of the degree of treatment activity. A patient's response may be measured by the decrease in the size of a tumor. For example, a patient may be considered to have responded to treatment if the tumor shrinks by 50% or more. There is no control group in these designs. Rather, the purpose of the trial is to determine if the drug shows enough activity against disease to warrant a full-scale, phase III clinical trial.

Let P_0 be the largest response proportion which, if true, clearly implies that the treatment does not warrant further study. P_0 is sometimes called the response rate of a *poor* treatment. For a new anti-tumor drug, this may be set to 0.10.

Let P_1 be the smallest response proportion which, if true, clearly implies that the treatment does warrant further study. P_1 is sometimes called the response rate of a *good* treatment. For a new anti-tumor drug, this may be set to 0.30.

A statistical test of hypothesis may be conducted to test the null hypothesis that $P \leq P_0$ versus the alternative hypothesis that $P \geq P_1$ (P is the true proportion responding to the treatment in the population). Let α be the probability of rejecting the null hypothesis when it is true. Let β be the probability of rejecting the alternative hypothesis when it is true.

Three-Stage Phase II Clinical Trials

A three-stage phase II design can be represented by six numbers: $R1$, $N1$, $R2$, $N2$, $R3$ and $N3$. $N1$ is the sample size in the first stage. $R1$ is the critical value in the first stage. If $R1$ or fewer responses occur in the $N1$ patients, the drug is rejected. $N2$ is the total sample size of stages one and two. $R2$ is the critical value in the second stage. If $R2$ or fewer responses occur in the $N2$ patients, the drug is rejected. $N3$ is the combined sample size of all three stages. $R3$ is the critical value in the combined sample. If $R3$ or fewer of the $N3$ patients respond, the drug is rejected.

The expected (or average) sample size of this design is

$$E(N_E) = N1 + (1 - PET1)(N2 - N1) + (1 - PET2)(N3 - N2)$$

where $PET1$ is the probability of early termination of the study after stage one and $PET2$ is the probability of early termination after stage two.

The probability of rejecting a drug with success proportion P can be found using the binomial distribution. The formulation is

$$\Pr(\text{reject} | P, N1, R1, R2, N2, R3, N3) = PET1 + PET2 + PET3$$

where

$$PET1 = B(R1 | P, N1)$$

$$PET2 = \sum_{X1=R1+1}^{\min(N1, R2)} b(X1 | P, N1) B(R2 - X1 | P, N2 - N1)$$

$$PET3 = \sum_{X1=R1+1}^{\min(N1, R3)} b(X1 | P, N1) \sum_{X2=R2+1-X1}^{\min(N3-N2, R3-X1)} b(X2 | P, N2 - N1) B(R3 - X1 - X2 | P, N3 - N2)$$

$$b(X | P, N) = \frac{N!}{X!(N-X)!} P^X (1-P)^{N-X}$$

$$B(X | P, N) = \sum_{r=0}^X b(r | P, N)$$

The two error rate constraints are

$$\Pr(\text{reject} | P0, N1, R1, R2, N2, R3, N3) \geq 1 - \alpha$$

and

$$\Pr(\text{reject} | P1, N1, R1, R2, N2, R3, N3) \leq \beta$$

Optimum Design

The optimum design minimizes the average sample size, $E(N)$, while meeting the error rate constraints. This design is found through an exhaustive search of all possible designs. This search may take several minutes to complete.

Designs Other Than Optimal

The optimal design minimizes the average sample size. There are examples where a less-than optimal design may be more desirable. For example, suppose the optimal design were $N1 = 5$, $N2 = 25$, and $N3 = 26$. This design is poor because the bulk of the subjects are tested in the second phase. Most researchers would rather have more balance in the sample sizes of the three stages. For reasons like this, the actual optimal design may be replaced by another, sub-optimal, design.

Design Flexibility

Dealing with sequential designs is complicated. It may be difficult to achieve exactly the number of patients proscribed for each phase. However, it should be remembered that the validity of the probability statements depends on the sample size requirements being met exactly. This is because the interpretation of an error rate probability statement is for repeated studies conducted in exactly the same way. We envision that if many studies of the same drug are conducted using the specific sampling plan when $P = P0$, a proportion α of them will be falsely terminated due to chance occurrences.

The point is, the interpretation of the error rates is for a large number of identical studies in which the sampling plan is identical and as proscribed. If the sampling plan is allowed to vary, this interpretation is invalid. Of course, the degree of possible error in interpretation depends on the degree to which the sampling plan is changed. We recognize that when dealing with human subjects, flexibility must be maintained. However, the researcher must also recognize that when the sampling plan is changed, the exact probability statements can no longer be calculated.

Procedure Options

This section describes the options that are specific to this procedure. These are located on the Design tab. For more information about the options of other tabs, go to the Procedure Window chapter.

Design Tab

The Design tab contains most of the parameters and options that you will be concerned with.

Designs

Designs to Display

This parameter controls which designs are displayed. Since several thousand designs may be considered during the search for the optimum, it is important to limit the number of designs reported on.

The options are:

- **All designs**
All designs considered are output. This option should only be used in special cases in which a small number of designs are tested. Otherwise, hundreds of pages of output will be generated.
- **Only designs that meet alpha & beta constraints**
Only designs that meet the alpha and beta constraints are shown. This allows you to consider many near optimal designs which may be selected on grounds other than expected sample size.
- **Optimum designs only**
Only the optimum design, the minimax design, and the single stage design are displayed.

Three-Stage Phase II Clinical Trials

Power and Alpha

Power

Power is the probability of rejecting the null hypothesis that the proportion responding to the treatment is less than or equal to $P0$ when this hypothesis is false. That is, $\text{Power} = \Pr(\text{rejecting } P \leq P0 | P >= P1)$.

Beta is the probability of not rejecting the hypothesis that the proportion responding to the treatment is less than or equal to $P0$ when this hypothesis is false. That is, $\text{Beta} = \Pr(\text{not rejecting } P \leq P0 | P >= P1)$.

The common range of power is 0.6 to 0.999 (Beta = 0.001 to 0.4). Popular values for power are 0.80 and 0.90 (Beta = 0.1 and 0.2).

Alpha

Alpha is the probability of rejecting the hypothesis that the proportion responding to the treatment is less than or equal to $P0$ when this hypothesis is actually true. That is, $\text{Alpha} = \Pr(\text{Rejecting } P \leq P0 | P \leq P0)$.

The range of Alpha is 0.001 to 0.25. Popular values are 0.05 and 0.10.

Effect Size

P0 (Poor)

This is the response proportion of a poor drug. If the true proportion responding to the treatment is less than this amount, study of the treatment would not be recommended.

This value must be less than $P1$ and greater than zero.

Only one value can be entered.

P1 (Good)

This is the response proportion of a good drug. If the true proportion responding to the treatment is greater than or equal to this amount, study of the treatment would be recommended.

This value must be greater than $P0$ and less than one.

Only one value can be entered.

Search Parameters – N (Combined Sample Size)

Min

N is the combined sample size of the three stages of the design. This parameter sets the minimum value of $N3$ that is used during the search. The optimum value of $N3$ must be between N Min and N Max or it will not be found.

The keyword MIN indicates that the value used is the minimum of the smallest sample size from a single stage design and $MIN2$ where $MIN2$ is calculated using

$$MIN2 = \frac{p_0 + p_1}{2} \left(1 - \frac{p_0 + p_1}{2} \right) \left[\frac{z_{1-\alpha} + z_{1-\beta}}{p_1 - p_0} \right]^2$$

Since it is unlikely that the three stage sample size will be less MIN, this provides a reasonable starting point for a search for N . You can also enter a value like MIN-x where x is a positive integer. This will cause the search to begin x units below the MIN.

This procedure may take a long time to run. Specifying a good starting value significantly reduces the running time.

Examples of valid entries are

2, 10, 20, MIN, MIN-1, MIN-15.

Three-Stage Phase II Clinical Trials

Max

N is the combined sample size of the three stages of the design. This parameter sets the maximum value of N_3 used during the search. The optimum value of N_3 should be between N_{Min} and N_{Max} or it will not be found.

The keyword BEST+ X indicates that the search should try at least X units above the latest optimum value of N_3 . For example, suppose the N_{Min} is set at 10. The search algorithm begins at 10, and then continues by examining 11, 12, and so. Suppose that the search finds a candidate optimum at $N = 13$. To make sure that 13 is the optimum, the search continues on from 13 to $13+X$ (if, for example, $X = 5$, this value is 18). If no new optimum designs are found, the design at $N_3 = 13$ is selected.

When using this option, X should be set large enough to guarantee that the true optimum can be found, but small enough so that the search does not take hours to complete. Our experience indicates that X should be greater than or equal to 8.

Examples of valid entries for this parameter are:

20

30

BEST+8

BEST 8 (the plus sign is optional)

BEST 3

Best 4 (capitalization is not necessary)

Step

This parameter sets the step size in the search for N . Usually, you would enter 1 here. Occasionally, you may want to increase the step size during the initial part of your search to speed up convergence. Once you have determined a likely range, you can tighten up the search boundaries and reset this value to 1.

Search Parameters – R (Rejection Number)

Step

This parameter sets the step size in the search for R . Usually, you would enter 1 here. Occasionally, you may want to increase the step size during the initial part of your search to speed up convergence. Once you have determined a likely range, you can tighten up the search boundaries and reset this value to 1.

Search Parameters – M1 (First Stage Sample Size)

Min

$M1$ is the sample size in the first stage. Thus, $M1=N1$. This option sets a minimum value for $M1$ so that designs with a very small value are not considered. For example, if $N_3=100$, you might want to only consider designs with at least $M1=25$. You can enter either an actual amount (no percent sign) or a percentage of N_3 . If you enter “%” after the value, the value is taken to mean a percentage of N_3 . For example, suppose that $N_3=50$. Then the entry of 30% would result in $M1=15$. However, if you did not enter the “%”, $M1=30$.

Note that it up to you to enter meaningful values. For example, you would not want to enter values for $M1\text{-Min}$, $M2\text{-Min}$, and $M3\text{-Min}$ that total to more than 100% because no designs would be considered.

Step

This parameter sets the step size in the search for $M1$. Usually, you would enter 1 here. Occasionally, you may want to increase the step size during the initial part of your search to speed up convergence. Once you have determined a likely range, you can tighten up the search boundaries and reset this value to 1.

Three-Stage Phase II Clinical Trials

Search Parameters – M2 (Second Stage Sample Size)**Min**

$M2$ is the number of subjects added in the second stage. Thus, $M2 = N2 - N1$ and $M1+M2+M3=N3$. This option sets a minimum value for $M2$ so that designs with a very small value are not considered. For example, if $N3=100$, you might want to only consider designs with at least $M2=25$. You can enter either an actual amount (no percent sign) or a percentage of $N3$. If you enter “%” after the value, the value is taken to mean a percentage of $N3$. For example, suppose that $N3=50$. Then the entry of 30% would result in $M2=15$. However, if you did not enter the “%”, $M2=30$.

Note that it up to you to enter meaningful values. For example, you would not want to enter values for $M1$ -Min, $M2$ -Min, and $M3$ -Min that total to more than 100% because no designs would be considered.

Step

This parameter sets the step size in the search for $M2$. Usually, you would enter a 1 here. Occasionally, you may want to increase the step size during the initial part of your search to speed up convergence. Once you have determined a likely range, you can tighten up the search boundaries and reset this value to 1.

Search Parameters – M3 (Third Stage Sample Size)**Min**

$M3$ is the number of subjects added in the third stage. Thus, $M3 = N3 - N2$ and $M1+M2+M3=N3$. This option sets a minimum value for $M3$ so that designs with a very small value are not considered. For example, if $N3=100$, you might want to only consider designs with at least $M3=25$. You can enter either an actual amount (no percent sign) or a percentage of $N3$. If you enter “%” after the value, the value is taken to mean a percentage of $N3$. For example, suppose that $N3=50$. Then the entry of 30% would result in $M3=15$. However, if you did not enter the “%”, $M3=30$.

Note that it up to you to enter meaningful values. For example, you would not want to enter values for $M1$ -Min, $M2$ -Min, and $M3$ -Min that total to more than 100% because no designs would be considered.

Example 1 – Calculating the Power and Validation using Chen

Chen (1997) provides the minimax and optimum design for the case $\alpha = 0.05$, $\beta = 0.20$, $P_0 = 0.05$, and $P_1 = 0.25$. The optimum design is 0/8, 1/13, and 2/19. The minimax design is 0/12, 1/15, and 2/16.

Setup

This section presents the values of each of the parameters needed to run this example. First, from the PASS Home window, load the **Three-Stage Phase II Clinical Trials** procedure window by expanding **Proportions**, then **One Proportion**, then clicking on **Multi-Stage Trials**, and then clicking on **Three-Stage Phase II Clinical Trials**. You may then make the appropriate entries as listed below, or open **Example 1** by going to the **File** menu and choosing **Open Example Template**.

| <u>Option</u> | <u>Value</u> |
|--------------------------|----------------------|
| Design Tab | |
| Designs to Display | Optimum designs only |
| Power | 0.80 |
| Alpha | 0.05 |
| P0 (Poor) | 0.05 |
| P1 (Good) | 0.25 |
| N Min | Min |
| N Max | Best 2 |

Note that the search may take several minutes to run, depending on the speed of your computer.

Annotated Output

Click the Calculate button to perform the calculations and generate the following output.

Numeric Results

| Possible Designs For $P_0=0.050$, $P_1=0.250$, $\alpha=0.050$, $\beta=0.200$ | | | | | | | | |
|---|---------|---------|---------|---------|--------|-------|-------|--------------|
| Stage 1 | Stage 2 | Stage 3 | Stage 1 | Overall | | | | Constraints |
| R1/N1 | R2/N2 | R3/N3 | Ave N | Pet P0 | Pet P0 | Alpha | Beta | Satisfied |
| 2/16 | 2/16 | 2/16 | 16.00 | 0.000 | 0.000 | 0.043 | 0.197 | Single Stage |
| 0/12 | 1/15 | 2/16 | 13.55 | 0.833 | 0.869 | 0.043 | 0.199 | Minimax |
| 0/8 | 1/13 | 2/19 | 10.41 | 0.663 | 0.880 | 0.049 | 0.195 | Optimum |

Report Definitions
 N1 is the sample size in the first stage.
 R1 is the drug rejection number in the first stage.
 N2 is the sample size in the first and second stages.
 R2 is the drug rejection number in the second stage.
 N3 is the combined sample size of all three stages.
 R3 is the drug rejection number in the third stage.
 Stage 1 PET P0 is the probability of early termination at the first stage.
 Stage 2 PET P0 is the probability of early termination at the second stage.
 Ave N is the average sample size if this design is repeated many times.
 Alpha is the probability of rejecting that $P \leq P_0$ when this is true.
 Beta is the probability of rejecting that $P \geq P_1$ when this is true.
 P0 is the response proportion of a poor drug.
 P1 is the response proportion of a good drug.

This report shows three designs. The first is the smallest single stage design. The second is the minimax solution. This is the design with the smallest total sample size (N). The third is the optimum design—the one that minimizes the average sample size.

Note that **PASS** matches the results of Chen.