Chapter 125

Two-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 two-stage design, the patients are divided into two 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.

This module finds designs that meet the error rate (alpha and beta) criterion and minimize the expected sample size. The algorithm is discussed in Simon (1989). Extending Simon'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 *P0* be the largest response proportion which, if true, clearly implies that the treatment does not warrant further study. *P0* is sometimes called the response rate of a poor treatment. For a new anti-tumor drug, this may be set to 0.10.

Let *P1* be the smallest response proportion which, if true, clearly implies that the treatment does warrant further study. *P1* 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 $P \leq P0$ versus the alternative hypothesis that $P \geq P1$ (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.

A phase II design can be represented by four numbers: *N1*, *R1*, *N*, and *R*. *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. *N* is the combined sample size for both the first and second stages. *R* is the critical value in the combined sample. If *R* or fewer of the *N* patients respond, the drug is rejected.

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

$$E(N_E) = N1 + (1 - PET)(N - N1)$$

where *PET* is the probability of early termination of the study.

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

$$\Pr(reject|P, N1, R1, R, N) = B(R1|P, N1) + \sum_{X=R1+1}^{\min(N1,R)} b(X|P, N1)B(R - X|P, N - N1)$$

where

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

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

The two error rate constraints are

$$\Pr(reject|P0, N1, R1, R, N) \geq 1 - \alpha$$

and

$$\Pr(reject|P1, N1, R1, R, N) \leq \beta$$

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 *N*1 = 5 and *N* = 25. This design is poor because only 5 patients are obtained during the first stage, but 20 are needed during the second stage. Most researchers would rather have more balance in the sample sizes of the two stages. Because of this, the actual optimal design may be rejected on other grounds.

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 *N1*, *R1*, *N*, *R* 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 scientist 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 specifies 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.

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, Power = Pr(rejecting $P \leq P0 | P \geq P1$).

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, Beta = Pr(not rejecting $P \leq P0 | P \geq P1$).

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, Alpha = Pr(Rejecting $P \leq P0 \mid P \leq P0$).

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 two stages of the design. This parameter sets the minimum value of *N* that is used during the search. The optimum value of *N* 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 *MIN*2 where *MIN*2 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 two stage sample size will be less MIN, this provides a reasonable starting point for a search for *N*. **However, experience has shown that you should use a small number such as 2 to insure that you obtain the optimum.**

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.

The problem here is that 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.

Max

N is the combined sample size of the two stages of the design. This parameter sets the maximum value of *N* used during the search. The optimum value of *N* 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*. 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 $N = 13$ 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 is that *X* should be greater than or equal to 8.

Examples of valid entries are 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)

Min

R is the treatment rejection number for the combined samples. If the total number of patients responding to the treatment is not greater than *R*, the treatment is deemed unworthy of further study. R Min sets the lower boundary for *R* during the search for the optimum design. The optimum design must have an *R* value between R Min and R Max.

The recommended value for this parameter is zero. Its range is from zero to N.

Max

R is the rejection number for the combined samples. If the total number of patients responding to the treatment is not greater than *R*, the treatment is deemed unworthy of further study.

R Max sets the upper boundary for *R* during the search for the optimum design. The optimum design must have an *R* value between R Min and R Max.

Since the upper value is *N* and *N* is also a varying parameter, you can set this parameter to *MAX* or *MAX-X* (replacing *X* with an appropriate integer like 1, 2, or 3). This causes the maximum value of *R* to be set to the current value of *N-X* during each iteration of the search.

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 – N1 (First Stage Sample Size)

Min

N1 is the sample size of the first stage. This value sets the minimum value of *N1* that is used during the search. The optimum value must be between N1 Min and N1 Max or it will not be found.

Although, in theory, the sample first stage design may have only 1, 2, or 3 patients, you may want to ignore such designs from consideration by setting this value to 4 or 5.

The actual range of this parameter is from 1 to *N*.

Max

N1 is the sample size of the first stage of the design. This parameter sets the maximum value of *N1* used during the search. The optimum value of *N1* should be between N Min and N Max or it will not be found. Although, in theory, the sample first stage design may have *N*-3, *N*-2, or *N*-1 patients, you may want to ignore such designs from consideration by setting this value to a smaller number.

Since the upper value is *N*-1 and *N* is also a varying parameter, you can set this parameter to MAX or MAX-X (replacing *X* with an appropriate value like 1 or 2). This causes the maximum value of *N1* to be set to the current value of *N*-*X*.

Examples:

10

20

MAX

Max-2

Max-4

Step

This parameter sets the step size in the search for *N1*. 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 – R1 (First Stage Rejection Number)

Min

R1 is the drug rejection number for the first stage. If the number of patients responding to the treatment in the first stage is not greater than *R1*, the treatment is deemed unworthy of further study. This parameter sets the lower boundary for *R1* during the search for the optimum design. The optimum design must have an *R1* value between R Min and R Max.

The recommended value for this parameter is zero. Its range is from zero to *N1*.

Max

R1 is the rejection number for the first stage. If the number of patients responding to the treatment in the first stage is not greater than *R1*, the treatment is deemed unworthy of further study. This parameter sets the upper boundary for *R1* in the search for the optimum design.

Since the upper value is *N1* and *N1* is a varying parameter, you can set this parameter to *MAX* or MAX-X (replacing *X* with an integer like 1,2, or 3). This causes the maximum value of *R1* to be set to the current value of *N1*-*X*.

The valid range of *R1* is between zero and *N1*.

Step

This parameter sets the step size in the search for *R1*. 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.

Example 1 – Calculating the Power

Suppose a design is wanted for the case Alpha = 0.05, Beta = 0.20, *P0* = 0.05, and *P1* = 0.25. This would be set up as follows:

Setup

This section presents the values of each of the parameters needed to run this example. First, from the PASS Home window, load the **Two-Stage Phase II Clinical Trials** procedure window by expanding **Proportions**, then **One Proportion**, then clicking on **Multi-Stage Trials**, and then clicking on **Two-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**.

Option	Value
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-1**
N Max	**Best+8**
N Step	**1**
R Min	**0**
R Max	**Max-3**
R Step	**1**
N1 Min	**1**
N1 Max	**Max-4**
N1 Step	**1**
R1 Min	**0**
R1 Max	**Max-1**
R1 Step	**1**

Annotated Output

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

Numeric Results

Possible Designs For P0=0.050, P1=0.250, Alpha=0.050, Beta=0.200

N1	R1	PET	N	R	Ave N	Alpha	Beta	Constraints Satisfied
16	2	0.000	16	2	16.00	0.043	0.197	Single Stage
12	0	0.540	16	2	13.84	0.043	0.199	Minimax
9	0	0.630	17	2	11.96	0.047	0.188	Optimum

Report Definitions
N1 is the sample size in the first stage.
R1 is the drug rejection number in the first stage.
PET is the probability of early termination of the study.
N is the combined sample size of both stages.
R is the combined drug rejection number after both stages.
Ave N is the average sample size if this design is repeated many times.
Alpha is the probability of rejecting that P ≤ P0 when this is true.
Beta is the probability of rejecting that P ≥ P1 when this is true.
P0 is the response proportion of a poor drug.
P1 is the response proportion of a good drug.

Summary Statements
The optimal two-stage design to test the null hypothesis that P ≤ 0.050 versus the alternative that P ≥ 0.250 has an expected sample size of 11.96 and a probability of early termination of 0.630. If the drug is actually not effective, there is a 0.047 probability of concluding that it is (the target for this value was 0.050). If the drug is actually effective, there is a 0.188 probability of concluding that it is not (the target for this value was 0.200). After testing the drug on 9 patients in the first stage, the trial will be terminated if 0 respond. If the trial goes on to the second stage, a total of 17 patients will be studied. If the total number responding is less than or equal to 2, the drug is rejected.

This report shows three designs. The first is the smallest single stage design. The second is the Minmax 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.

Example 2 – Validation using Simon

Simon (1989) page 4 in his Table 1 presents designs for several scenarios. The first row of the table sets *P0* to 0.05, *P1* to 0.25, Alpha to 0.10, and Beta to 0.10. The optimal design is *N1* = 9, *R1* = 0, *N* = 24, and *R* = 2. The minimax design is *N1* = 13, *R1* = 0, *N* = 20, and *R* = 2. We will now run this example through ***PASS***.

Setup

This section presents the values of each of the parameters needed to run this example. First, from the PASS Home window, load the **Two-Stage Phase II Clinical Trials** procedure window by expanding **Proportions**, then **One Proportion**, then clicking on **Multi-Stage Trials**, and then clicking on **Two-Stage Phase II Clinical Trials**. You may then make the appropriate entries as listed below, or open **Example 2** by going to the **File** menu and choosing **Open Example Template**.

Option	Value
Design Tab	
Which Designs	**Optimum designs only**
Power	**0.90**
Alpha	**0.10**
P0 (Poor)	**0.05**
P1 (Good)	**0.25**
N Min	**Min-1**
N Max	**Best+8**
N Step	**1**
R Min	**0**
R Max	**Max-3**
R Step	**1**
N1 Min	**1**
N1 Max	**Max-4**
N1 Step	**1**
R1 Min	**0**
R1 Max	**Max-1**
R1 Step	**1**

Output

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

Numeric Results

Possible Designs For P0=0.050, P1=0.250, Alpha=0.050, Beta=0.200

N1	R1	PET	N	R	Ave N	Alpha	Beta	Constraints Satisfied
20	2	0.000	20	2	20.00	0.075	0.091	Single Stage
13	0	0.513	20	2	16.41	0.074	0.097	Minimax
9	0	0.630	24	2	14.55	0.093	0.097	Optimum

PASS calculates exactly the same optimal design and minimax design.