Chapter 761

Group-Sequential T-Tests for Two Means (Simulation)

In this procedure, boundaries are calculated analytically, while simulation is used for the calculation of power (and sample size). A variety of futility boundary options are available.

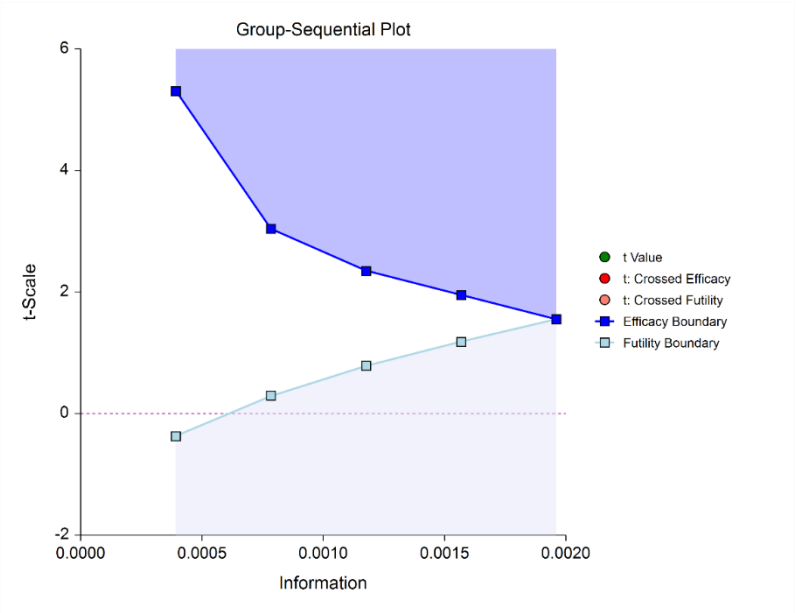
The corresponding analysis and sample size re-estimation procedure, found in NCSS Analysis and Graphics software, is Group-Sequential T-Tests for Two Means.

Introduction

This procedure can be used to determine power, sample size and/or boundaries for group-sequential tests comparing the means of two groups using Welch's t-test. For one- and two-sided tests, efficacy and/or futility boundaries can be generated. The spacing of the stages can be equal or custom specified. Individual stages may also be skipped. Boundaries can be computed based on popular alpha- and beta-spending functions (O'Brien-Fleming Analog, Pocock Analog, Hwang-Shih-DeCani Gamma family, linear) or custom spending functions, or boundaries may be input directly, if desired. Futility boundaries can be binding or non-binding. Corresponding P-Value boundaries are given for each boundary statistic (unless boundaries are input directly). Alpha and/or beta spent at each stage is reported. Plots of boundaries are also produced.

This procedure is used as the planning tool for determining sample size and initial boundaries. Stage data, as it is obtained, can be evaluated using the companion procedure *Group-Sequential T-Tests for Two Means*. The companion procedure also gives the option for sample-size re-estimation and updated boundaries for current-stage information. In that procedure, simulation can be used to evaluate boundary-crossing probabilities given the current stage results.

An example of a group-sequential boundary plot produced in this procedure is shown below.



© NCSS, LLC. All Rights Reserved.

Outline of a Group-Sequential Study

There are three basic phases of a group-sequential (interim analysis) study:

- Design
- Group-Sequential Analysis
- Reporting

Design Phase – Determine the Number of Subjects

To begin the group-sequential testing process, an initial calculation should be made to determine the sample size and target information if the final stage is reached (maximum information). The sample size calculation requires the specification of the following:

- Alpha
- Power
- Test Direction (two-sided or one-sided direction)
- Types of boundaries (efficacy, binding futility, non-binding futility)
- Maximum number of stages
- Proportion of maximum information at each stage
- Spending functions
- Within group standard deviations
- Assumed difference in means

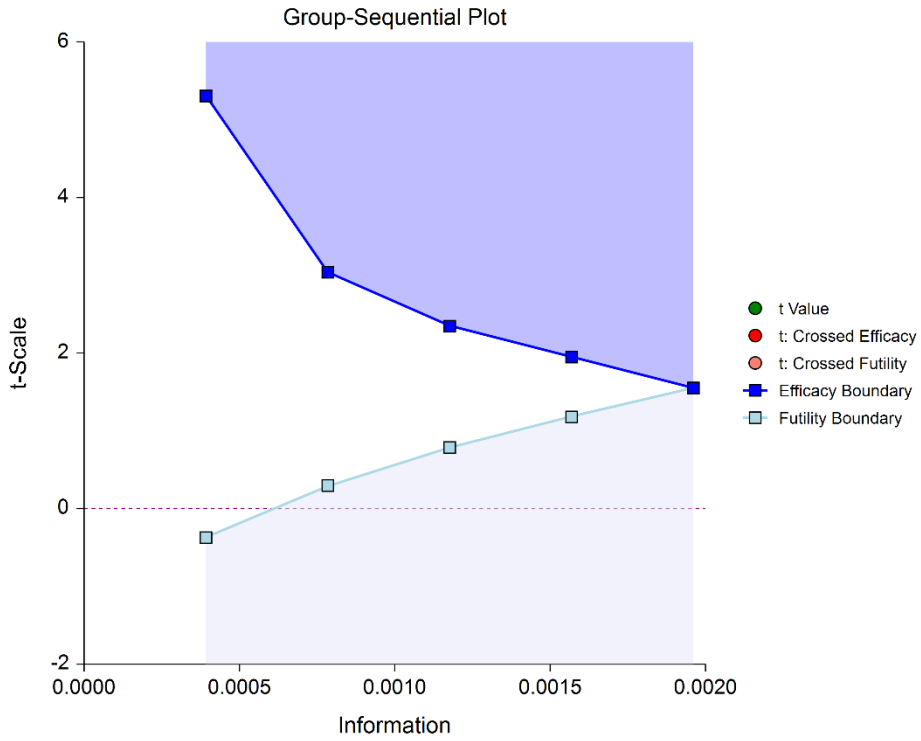
The design phase calculation is performed in this procedure. **PASS** software permits the user to easily try a range of mean differences and standard deviations, as these values are typically not known in advance.

The resulting sample size of the sample size calculation also permits the calculation of the maximum information, which is the total information of the study if the final stage is reached (for calculation details, see the Information section later in this chapter).

Based on the maximum information, the target information and target sample size of each stage may be calculated. In particular, this permits the user to have a target sample size for the first stage.

Although it is likely to change over the course of the group-sequential analysis, a design group-sequential boundary plot can be a useful visual representation of the design:

© NCSS, LLC. All Rights Reserved.



Group-Sequential Analysis Phase

A group sequential analysis consists of a series of stages where a decision to stop or continue is made at each stage. This analysis can be performed using the companion (analysis) procedure to this sample size procedure in **NCSS**.

First Interim Stage

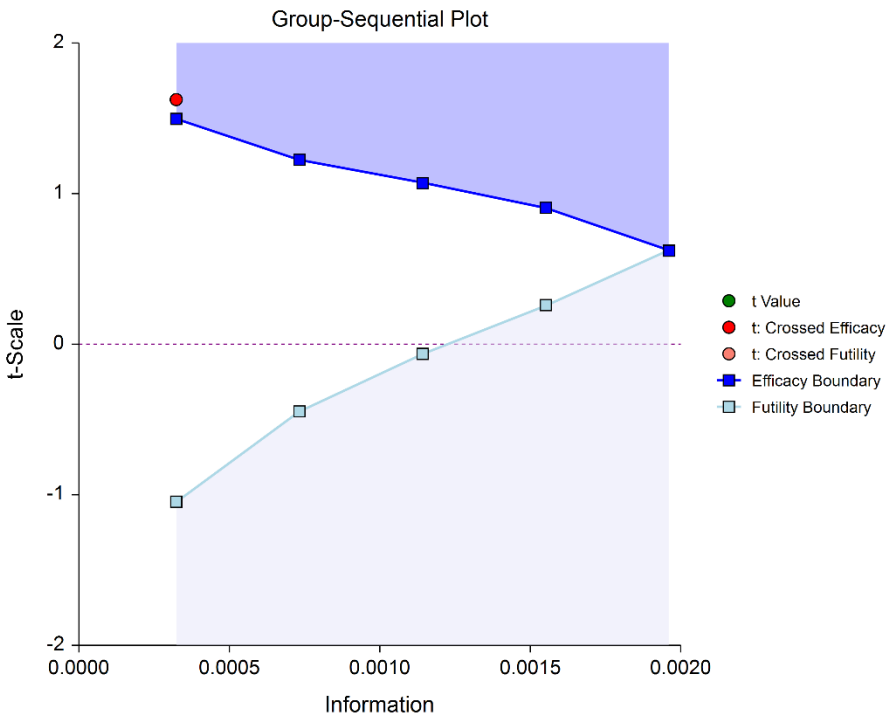
The design phase gives the target number of subjects for the first stage. The study begins, and response data is collected for subjects, moving toward the first-stage target number of subjects, until a decision to perform an analysis on the existing data is made. The analysis at this point is called the first stage.

Unless the number of subjects at the first stage matches the design target for the first stage, the calculated information at the first stage will not exactly match the design information for the first stage. Further, the sample standard deviations will rarely, if ever, match the standard deviations used in the calculation of the information at the design stage, and thus the calculated information at the first stage will differ from the design information. Generally, the calculated information will not differ too greatly from the design information, but regardless, spending function group-sequential analysis is well-suited to make appropriate adjustments for any differences.

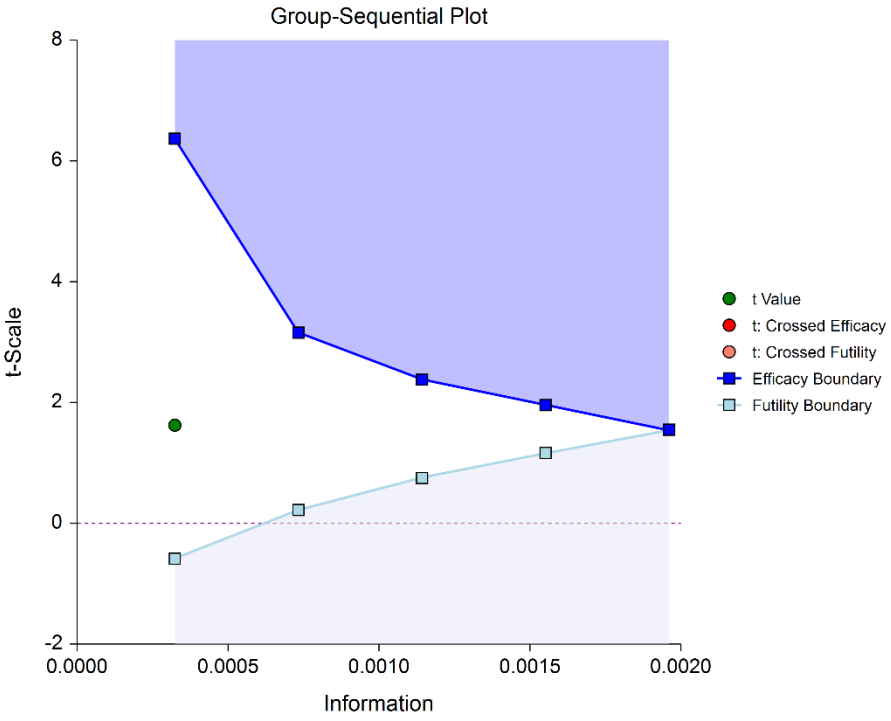
The first stage information is divided by the maximum information to obtain the stage one information proportion (or information fraction). This information proportion is used in conjunction with the spending function(s) to determine the alpha and/or beta spent at that stage. In turn, stage one boundaries, corresponding to the information proportion, are calculated.

A *t*-statistic is calculated from the raw mean difference and sample standard deviations. The stage one *t*-statistic is compared to each of the stage one boundaries. Typically, if one of the boundaries is crossed, the study is stopped (non-binding futility boundaries may be an exception).

© NCSS, LLC. All Rights Reserved.



If none of the boundaries are crossed the study continues to the next stage.



If none of the boundaries are crossed it may also be useful to examine the conditional power or stopping probabilities of future stages, using the **NCSS** procedure. Conditional power and stopping probabilities are based on the user-specified supposed true difference.

© NCSS, LLC. All Rights Reserved.

Second and other interim stages (if reached)

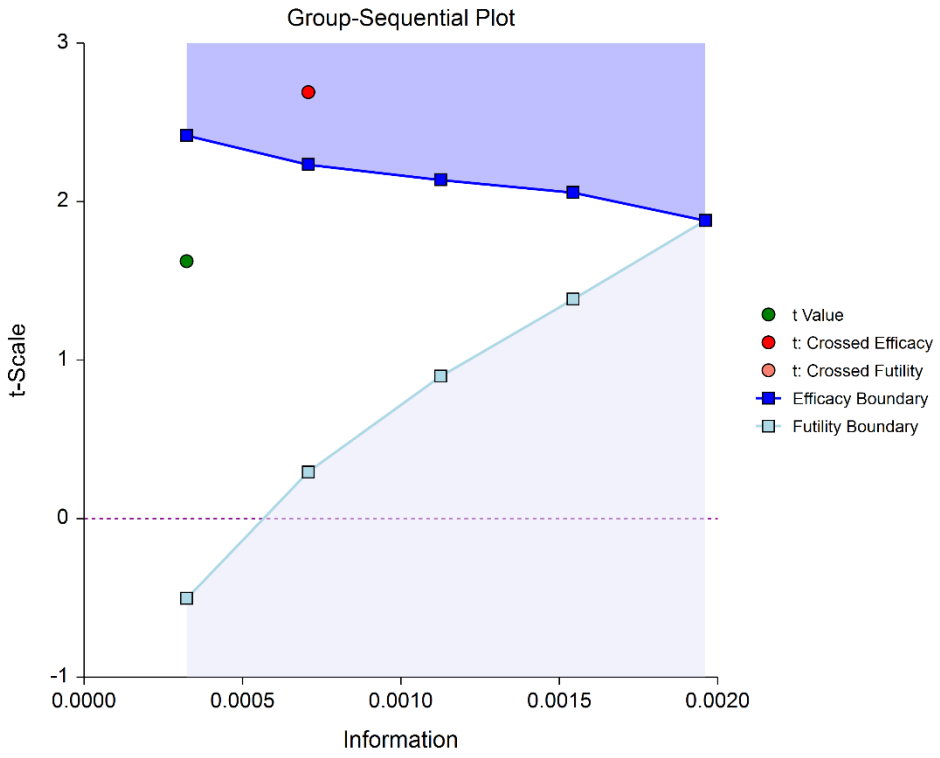
Since the first stage information proportion is not equal to the design information proportion, a designation must be made at this point as to the target information of the second stage. Two options are available in the **NCSS** procedure.

One option is to target the information proportion of the original design. For example, if the original design proportions of a four-stage design are 0.25, 0.50, 0.75, 1.0, and the stage one observed proportion is 0.22, the researcher might still opt to target 0.50 for the second stage, even though that now requires an additional information accumulation of 0.28 (proportion). The third and fourth stage targets would also remain 0.75 and 1.0.

A second option is to adjust the target information proportionally to the remaining proportions. For this option, if the design proportions are 0.25, 0.50, 0.75, 1.0, and 0.22 is observed, the remaining 0.78 is distributed proportionally to the remaining stages. In this example, the remaining target proportions become 0.48, 0.74, 1.0.

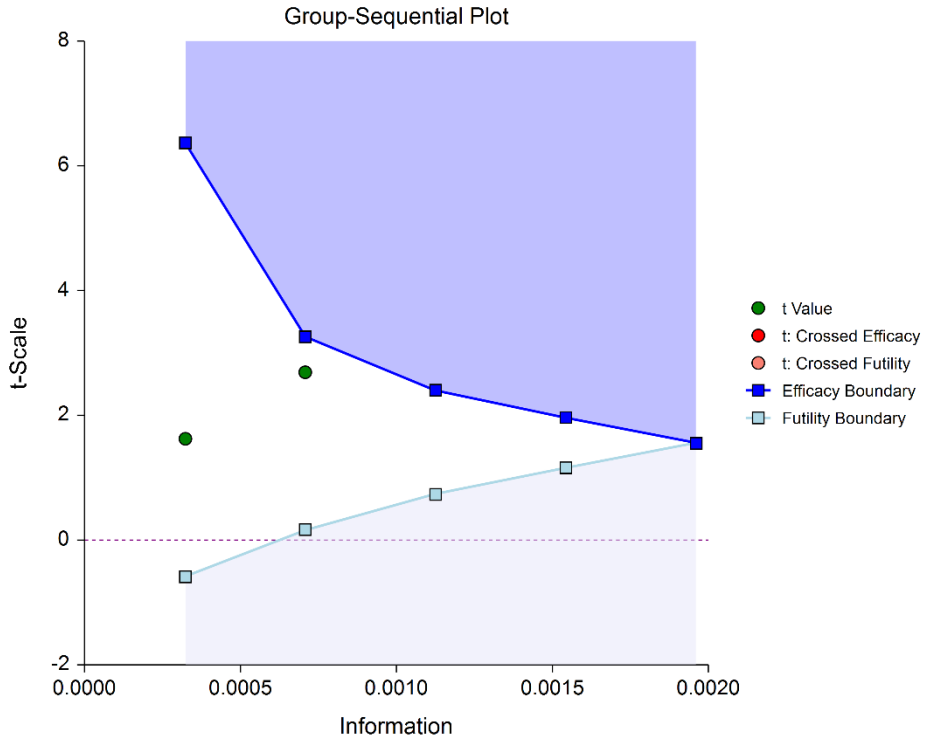
For either option, once the target information is determined for the next stage, revised target sample sizes are given (in the **NCSS** procedure), and the study continues until the decision is made to perform the next interim analysis on the cumulative response data. In the same manner as the first stage, the current stage information proportion is used with the spending function to determine alpha and/or beta spent at the current stage. The current stage boundaries are then computed. The *t*-statistic is calculated and compared to the boundaries, and a decision is made to stop or continue.

If a boundary is crossed, the study is typically stopped.



© NCSS, LLC. All Rights Reserved.

If none of the boundaries are crossed the study continues to the next stage.



Once again, if no boundary is crossed, conditional power and stopping probabilities may be considered based on a choice of a supposed true difference.

The study continues from stage to stage until the study is stopped for the crossing of a boundary, or until the final stage is reached.

Final Stage (if reached)

The final stage (if reached) is similar to all the interim stages, with a couple of exceptions. For all interim analyses the decision is made whether to stop for the crossing of a boundary, or to continue to the next stage. At the final stage, only the decision of efficacy or futility can be made.

Another intricacy of the final stage that does not apply to the interim stages is the calculation of the maximum information. At the final stage, the current information must become the maximum information, since the spending functions require that the proportion of information at the final look must be 1.0. If the current information at the final stage is less than the design maximum information, the scenario is sometimes described as *under-running*. Similarly, if the current information at the final stage is greater than the design maximum information, the result may be termed *over-running*.

For both under-running and over-running, the mechanism for adjustment is the same, and is described in the Technical Details section, under Information and Total Information.

Aside from these two exceptions, the final stage analysis is made in the same way that interim analyses were made. The remaining alpha and beta to be spent are used to calculate the final stage boundaries. If the test is a one-sided test, then the final stage boundary is a single value. The final stage *t*-statistic is computed from the sample means and standard deviations of the complete data from each group. The *t*-statistic is compared to the boundary and a decision of efficacy or futility is made.

© NCSS, LLC. All Rights Reserved.

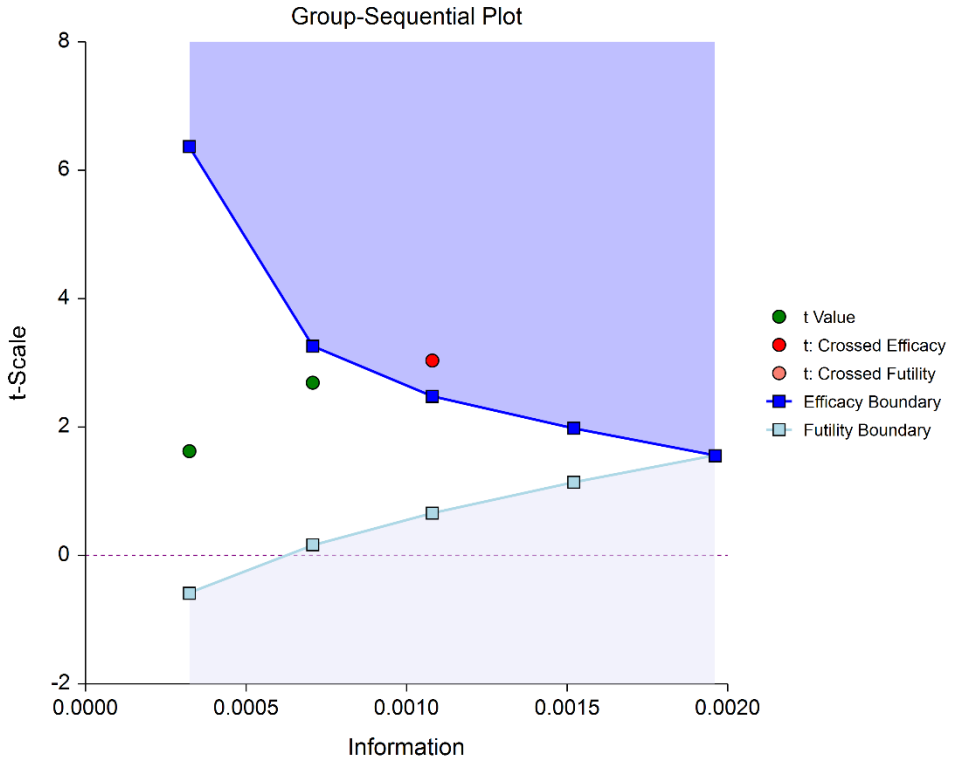
Reporting Phase

Once a group-sequential boundary is crossed and the decision is made to stop, there remains the need to properly summarize and communicate the study results. Some or all of the following may be reported:

- Boundary plot showing the crossed boundary
- Adjusted confidence interval and estimate of the mean difference
- Sample size used

Boundary plot showing the crossed boundary

The boundary plot gives an appropriate visual summary of the process leading to the reported decision of the study.



Adjusted confidence interval and estimate of the mean difference

Due to the bias that is introduced in the group-sequential analysis process, the raw data confidence interval of the difference in means should not be used. An adjusted confidence interval should be used instead.

Sample size used

The sample size at the point the study was stopped should be reported in addition to the sample size that would have been used had the final stage been reached.

© NCSS, LLC. All Rights Reserved.

Technical Details

Many articles and texts have been written about group sequential analysis. Details of many of the relevant topics are discussed below, but this is not intended to be a comprehensive review of group-sequential methods. One of the more influential works in the area of group-sequential analysis is Jennison and Turnbull (2000).

Null and Alternative Hypotheses

For comparing two means, the basic null hypothesis is that the means are equal,

$$H_0: \mu_1 = \mu_2$$

with three common alternative hypotheses,

$$H_a: \mu_1 \neq \mu_2 ,$$

$$H_a: \mu_1 < \mu_2 , \text{ or}$$

$$H_a: \mu_1 > \mu_2 ,$$

one of which is chosen according to the nature of the experiment or study.

These hypotheses may be specified equivalently as

$$H_0: \mu_1 - \mu_2 = 0$$

versus

$$H_a: \mu_1 - \mu_2 \neq 0$$

$$H_a: \mu_1 - \mu_2 < 0$$

$$H_a: \mu_1 - \mu_2 > 0$$

A slightly different set of null and alternative hypotheses are used if the goal of the test is to determine whether μ_1 or μ_2 is greater than or less than the other by a given amount.

The null hypothesis then takes on the form

$$H_0: \mu_1 - \mu_2 = \textit{Hypothesized Difference}$$

and the alternative hypotheses,

$$H_a: \mu_1 - \mu_2 \neq \textit{Hypothesized Difference}$$

$$H_a: \mu_1 - \mu_2 < \textit{Hypothesized Difference}$$

$$H_a: \mu_1 - \mu_2 > \textit{Hypothesized Difference}$$

For testing these hypotheses with a hypothesized difference, a superiority by a margin or non-inferiority test should be used instead.

© NCSS, LLC. All Rights Reserved.

Stages in Group-Sequential Testing

The potential to obtain the benefit from a group-sequential design and analysis occurs when the response data are collected over a period of weeks, months, or years rather than all at once. A typical example is the case where patients are enrolled in a study as they become available, as in many types of clinical trials.

A group-sequential testing stage is a point in the accumulation of the data where an interim analysis occurs, either by design or by necessity. At each stage, a test statistic is computed with all the accumulated data, and it is determined whether a boundary (efficacy or futility) is crossed. When an efficacy (or futility) boundary is crossed, the study is usually concluded, and inference is made. If the final stage is reached, the group-sequential design forces a decision of efficacy or futility at this stage.

For the discussions below, a non-specific interim analysis stage is referenced as *k*, and the final stage is *K*.

Test Statistic

The *t*-statistic for any stage *k* is obtained from all the accumulated data up to and including that stage.

t-Statistic Allowing Unequal Variance (Welch)

Welch's *t*-test is commonly used when no assumption of equal variances for the two population is made. This is commonly known as the Aspin-Welch test, Welch's *t*-test (Welch, 1937), or the Satterthwaite method.

$$t_k = \frac{\bar{X}_{1k} - \bar{X}_{2k}}{\sqrt{\frac{s_{1k}^2}{n_{1k}} + \frac{s_{2k}^2}{n_{2k}}}}$$

with degrees of freedom

$$df = \frac{\left(\frac{s_{1k}^2}{n_{1k}} + \frac{s_{2k}^2}{n_{2k}}\right)^2}{\frac{\left(\frac{s_{1k}^2}{n_{1k}}\right)^2}{n_{1k} - 1} + \frac{\left(\frac{s_{2k}^2}{n_{2k}}\right)^2}{n_{2k} - 1}}$$

Group-Sequential Design Phase

In most group-sequential studies there is a design or planning phase prior to beginning response collection. In this phase, researchers specify the anticipated number and spacing of stages, the types of boundaries that will be used, the desired alpha and power levels, the spending functions, the anticipated standard deviations, and an estimate of the true difference in means.

Based on these input parameters, an initial set of boundaries is produced, an estimate of the total number of needed subjects is determined, and the anticipated total information at the final stage is calculated. This procedure can be used to make these planning phase sample size estimation calculations.

© NCSS, LLC. All Rights Reserved.

Information and Total Information

In the group-sequential design phase, the final stage (*K*) or total (design) information is calculated from the specified standard deviations and the final sample sizes, as

$$I_K^* = \frac{1}{\frac{\sigma_{1K}^2}{n_{1K}} + \frac{\sigma_{2K}^2}{n_{2K}}}$$

The information at any stage *k* may be calculated from the specified standard deviations and the sample sizes, as

$$I_k = \frac{1}{\frac{\sigma_{1k}^2}{n_{1k}} + \frac{\sigma_{2k}^2}{n_{2k}}}$$

The planning σ_1 and σ_2 are used for σ_{1k} and σ_{2k}, since realized values are not available before data is collected. When the analysis is carried out, the sample estimates s_{1k} and s_{2k} will be used in place of σ_{1k} and σ_{2k}. The proportion of the total information (or information fraction) at any stage is

$$p_k = \frac{I_k}{I_K^*}$$

The information fractions are used in conjunction with the spending function(s) to define the alpha and/or beta to be spent at each stage.

To properly use the spending function at the final stage, it is required that $p_K = 1$. However, if the final stage is reached, we see that

$$I_K = \frac{1}{\frac{\sigma_{1K}^2}{n_{1K\,achieved}} + \frac{\sigma_{2K}^2}{n_{2K\,achieved}}} \neq I_K^* = \frac{1}{\frac{\sigma_{1K}^2}{n_{1K}} + \frac{\sigma_{2K}^2}{n_{2K}}}$$

so that

$$p_K = \frac{I_K}{I_K^*} \neq 1$$

When $I_K > I_K^*$, it is called over-running. When $I_K < I_K^*$, it is called under-running. In either case, the spending function is adjusted to accommodate the inequality, by redefining

$$I_K^* = I_K$$

See the discussion in Wassmer and Brannath (2016), pages 78-79, or Jennison and Turnbull (2000), pages 153-154, 162.

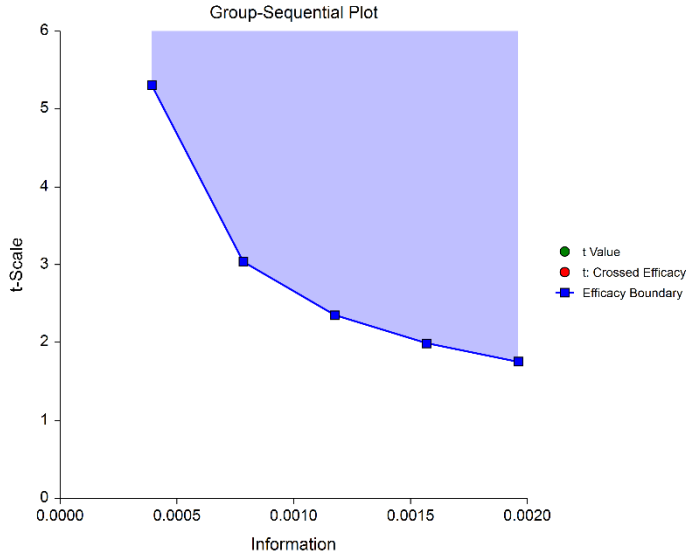
© NCSS, LLC. All Rights Reserved.

Types of Boundaries

A variety of boundary designs are available to reflect the needs of the study design.

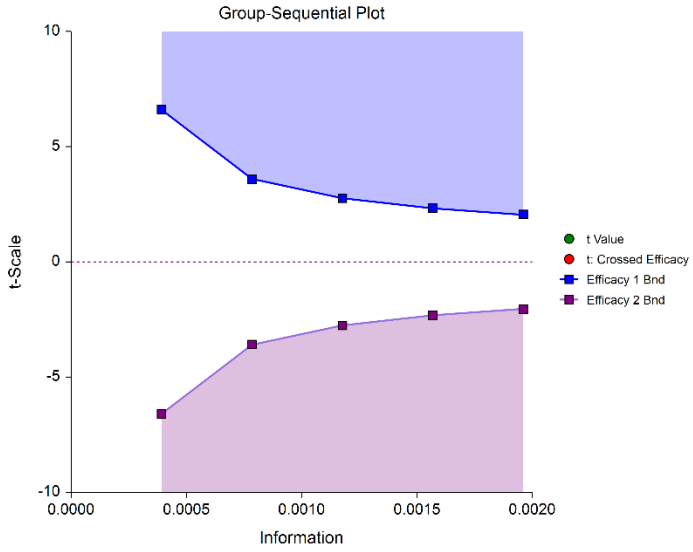
Efficacy Only (One-Sided)

The simplest group-sequential test involves a single set of stage boundaries with early stopping for efficacy.



Efficacy Only (Two-Sided, Symmetric)

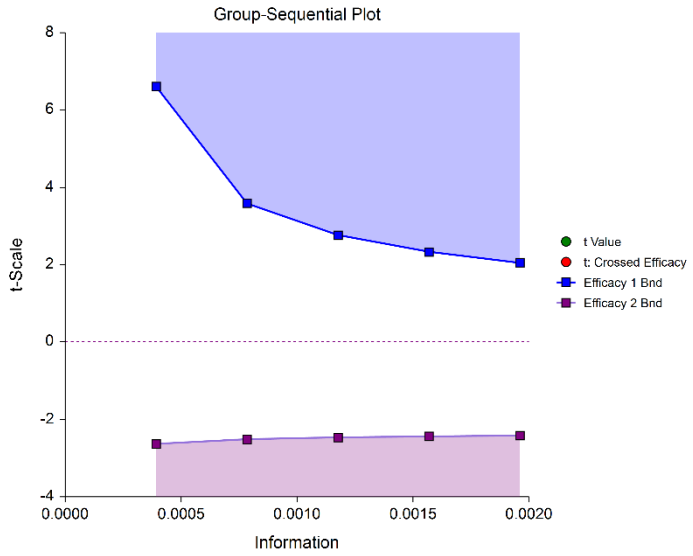
This boundary type would be used if the goal is to compare treatments, and it is not known in advance which treatment should be better.



© NCSS, LLC. All Rights Reserved.

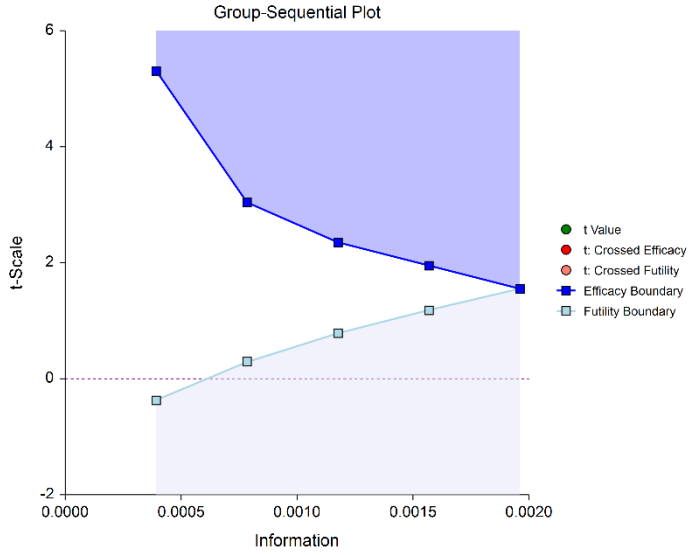
Efficacy 1 and Efficacy 2 / Harm (Two-Sided, Asymmetric)

These boundaries might be used to show efficacy on one side or harm on the other side. This design might be used in place of a one-sided efficacy and futility design if showing harm has additional benefit over stopping early for futility.



Efficacy and Binding Futility (One-Sided)

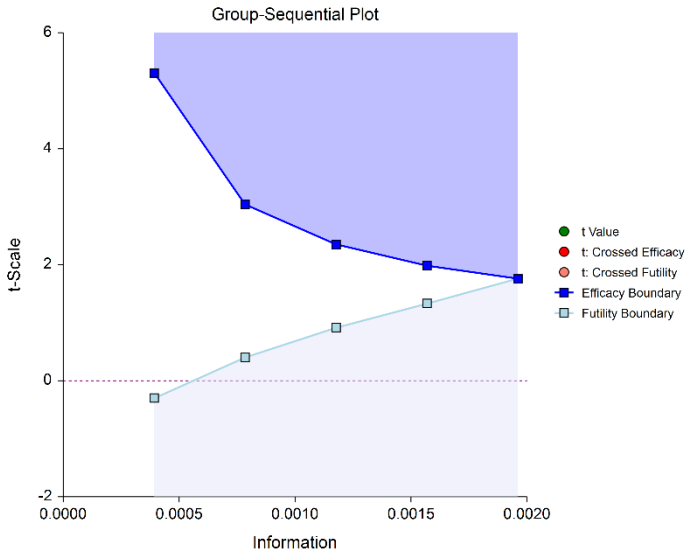
This design allows early stopping for either efficacy or futility. For binding futility designs, the Type I error protection (alpha) is only maintained if the study is strictly required to stop if either boundary is crossed.



© NCSS, LLC. All Rights Reserved.

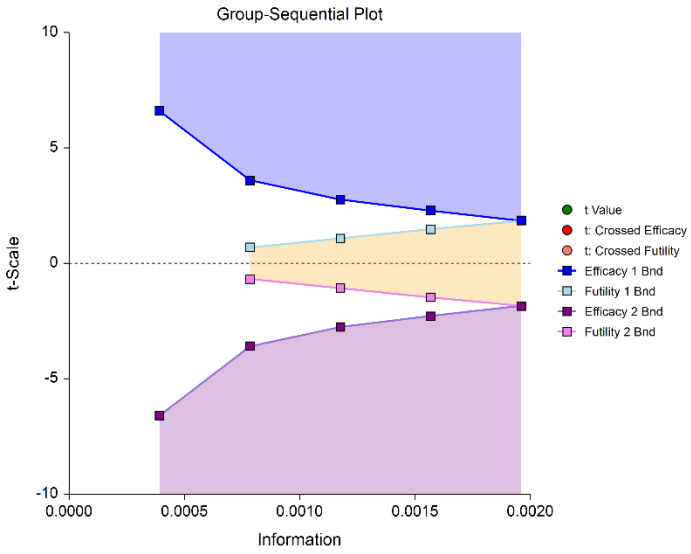
Efficacy and Non-Binding Futility (One-Sided)

This design also allows early stopping for either efficacy or futility. For non-binding futility designs, the Type I error protection (alpha) is maintained, regardless of whether the study continues after crossing a futility boundary. However, the effect is to make the test conservative (alpha is lower than the stated alpha and power is lower than the stated power).



Efficacy and Binding Futility (Two-Sided, Symmetric)

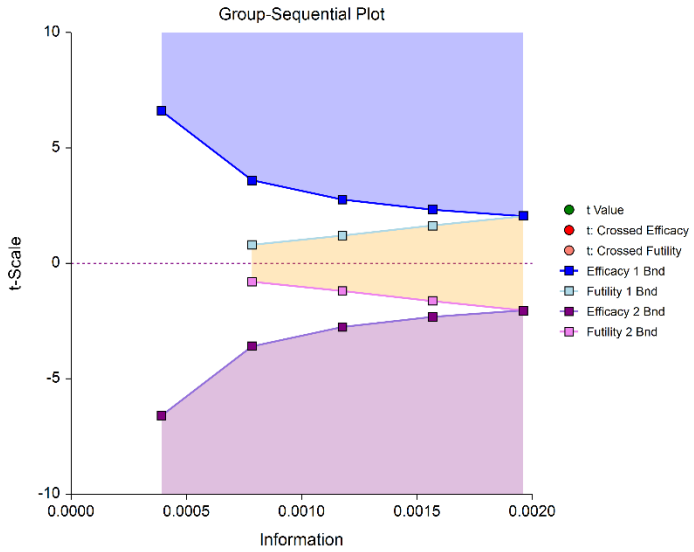
This design allows early stopping for either efficacy or futility on either side. Alpha is preserved only if crossing of futility boundaries strictly leads to early stopping for futility. In early looks of this design, the futility boundaries may overlap. Overlapping futility boundaries may be skipped or left as they are.



© NCSS, LLC. All Rights Reserved.

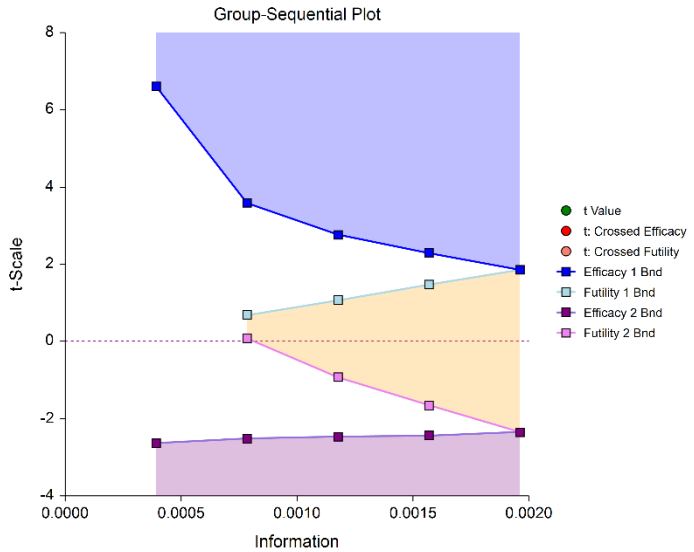
Efficacy and Non-Binding Futility (Two-Sided, Symmetric)

This design allows early stopping for either efficacy or futility on either side. Alpha is preserved even when the study is allowed to continue after crossing a futility boundary. In early looks of this design, the futility boundaries may overlap. Overlapping futility boundaries may be skipped or left as they are.



Efficacy 1, Efficacy 2 / Harm, and Binding Futility (Two-Sided, Asymmetric)

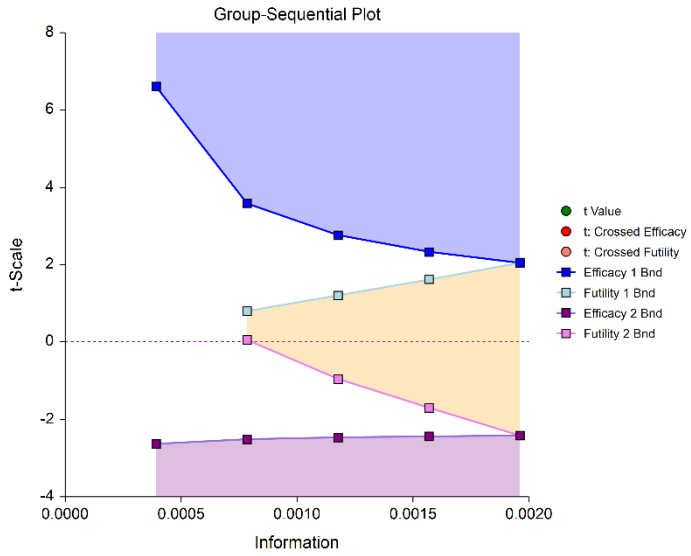
This design allows early stopping for efficacy and efficacy futility, and for harm and harm futility (or efficacy 2 and efficacy 2 futility). Binding futility boundaries require that the study is stopped when a binding futility boundary is crossed. In early looks of this design, the futility boundaries may overlap. Overlapping futility boundaries may be skipped or left as they are.



© NCSS, LLC. All Rights Reserved.

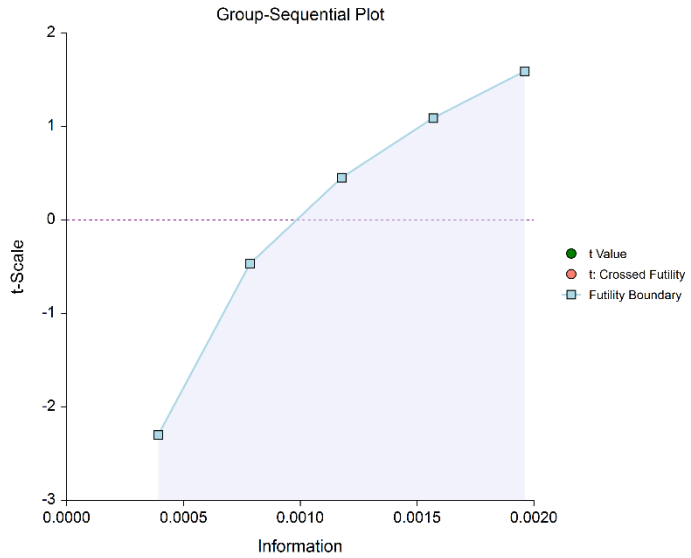
Efficacy 1, Efficacy 2 / Harm, and Non-Binding Futility (Two-Sided, Asymmetric)

This design allows early stopping for efficacy and efficacy futility, and for harm and harm futility (or efficacy 2 and efficacy 2 futility). Non-binding futility boundaries do not require that the study is stopped when a binding futility boundary is crossed, but the study design is conservative. In early looks of this design, the futility boundaries may overlap. Overlapping futility boundaries may be skipped or left as they are.



Futility Only (One-Sided)

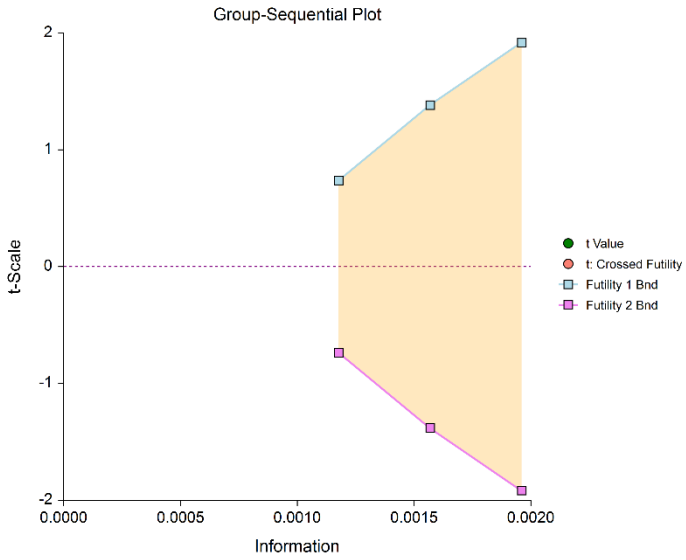
In this design, the interim analyses are used only for futility. Please be aware that, due to computational complexity, these boundaries may take several minutes to compute, particularly when some stages are skipped.



© NCSS, LLC. All Rights Reserved.

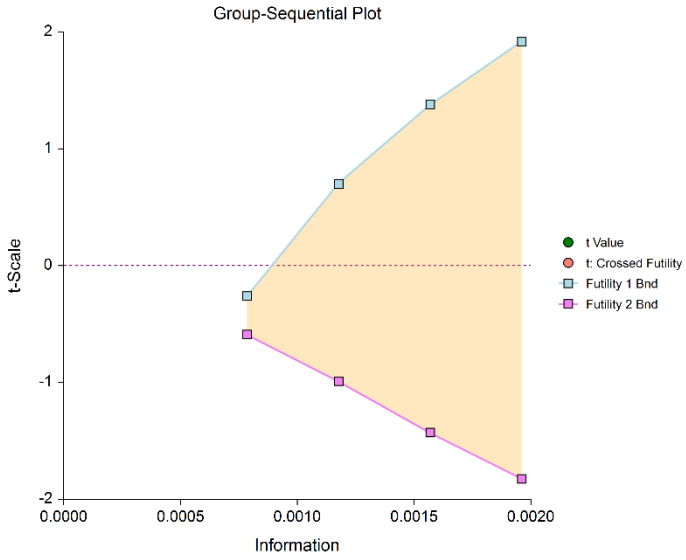
Futility Only (Two-Sided, Symmetric)

In this design, the study is stopped early only for futility. Overlapping futility boundaries may be skipped or left as they are. Please be aware that, due to computational complexity, these boundaries may take several minutes to compute, particularly when overlapping boundaries are removed or some stages are skipped.



Futility Only (Two-Sided, Asymmetric)

In this design, all stages previous to the final stage are used only for futility. Overlapping futility boundaries may be skipped or left as they are. Please be aware that, due to computational complexity, these boundaries may take several minutes to compute, particularly when overlapping boundaries are removed or some stages are skipped.



© NCSS, LLC. All Rights Reserved.

Boundary Calculations

The foundation of the spending function approach used in this procedure is given in Lan & DeMets (1983). This procedure implements the methods given in Reboussin, DeMets, Kim, & Lan (1992) to calculate the boundaries and stopping probabilities of the various group sequential designs. Some adjustments are made to these methods to facilitate the calculation of futility boundaries.

Obtaining t-scale Boundaries from Z-scale Boundaries

The methodology underlying all boundary calculations is based on *Z*-boundary calculations. Following the recommendation of Jennison and Turnbull (2000), p. 80, *Z*-scale boundaries are first produced, using the common methods. The corresponding *p*-value boundaries are computed from the *Z*-value boundaries. These *p*-value boundaries are then converted to *t*-value boundaries, using the appropriate degrees of freedom. As discussed in Jennison and Turnbull (2000), these *t*-value boundaries have been found to be 'remarkably accurate.'

Binding vs. Non-Binding Futility Boundaries

Futility boundaries are used to facilitate the early stopping of studies when early evidence leans to lack of efficacy. When binding futility boundaries are to be used, the calculation of the futility and efficacy boundaries assumes that the study will be strictly stopped at any stage where a futility or efficacy boundary is crossed. If strict adherence is not maintained, then the Type I and Type II error probabilities associated with the boundaries are no longer valid. One (perhaps undesirable) effect of using binding futility boundaries is that the resulting final stage boundary may be lower than the boundary given in the corresponding fixed-sample design.

When non-binding futility boundaries are calculated, the efficacy boundaries are first calculated ignoring futility boundaries completely. This is done so that alpha may be maintained whether or not a study continues after crossing a futility boundary. One (perhaps undesirable) effect of using non-binding futility boundaries is that the overall group-sequential test becomes conservative (alpha is lower than the stated alpha and power is lower than the stated power).

Spending Functions

Spending functions are used to distribute portions of alpha (or beta) to the stages according to the proportion of accumulated information at each look.

Spending Function Characteristics

- Spending functions give a value of zero when the proportion of accumulated information is zero.

$$\alpha(0) = 0 \text{ (for alpha-spending)}$$

$$\beta(0) = 0 \text{ (for beta-spending)}$$

- Spending functions are increasing functions.

© NCSS, LLC. All Rights Reserved.

- Spending functions give a value of alpha (or beta) when the proportion of accumulated information is one.

$$\alpha(1) = \alpha \text{ (for alpha-spending)}$$

$$\beta(1) = \beta \text{ (for beta-spending)}$$

Using spending functions in group-sequential analyses is very flexible in that neither the information proportions nor the number of stages need be specified in advance to maintain Type I and Type II error protection.

Spending Functions Available in this Procedure

The following spending functions are shown as alpha-spending functions. The corresponding beta-spending function is given by replacing α with β.

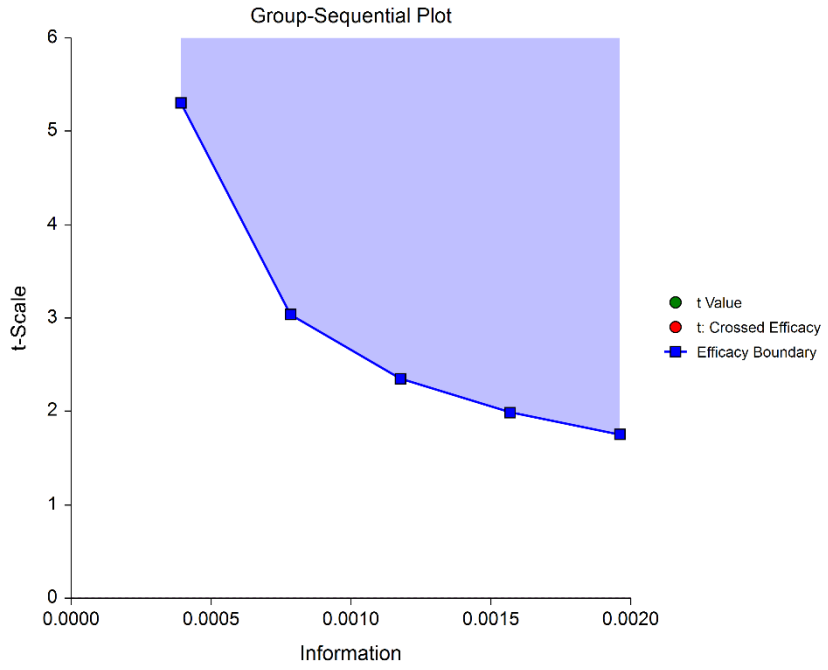
O'Brien-Fleming Analog

The O'Brien Fleming Analog (Lan & DeMets, 1983) roughly mimics the O'Brien-Fleming (non-spending function) design, with the key attribute that only a small proportion of alpha is spent early. Its popularity comes from it proportioning enough alpha to the final stage that the final stage boundary is not too different from the fixed-sample (non-group-sequential) boundary.

$$\alpha(0) = 0$$

$$\alpha(p_k) = 2 - 2\Phi\left(\frac{Z_{1-\alpha/2}}{\sqrt{p_k}}\right)$$

$$\alpha(1) = \alpha$$



© NCSS, LLC. All Rights Reserved.

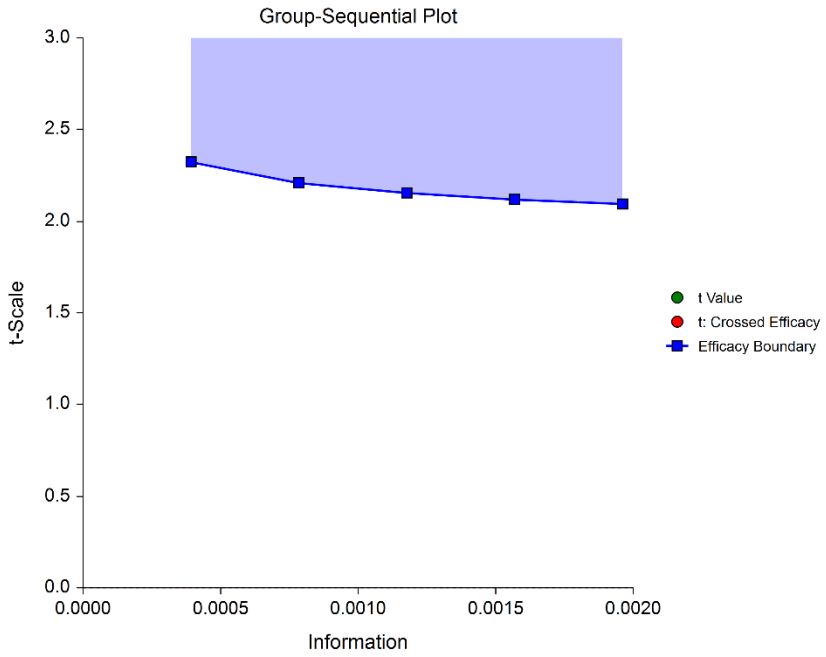
Pocock Analog

The Pocock Analog (Lan & DeMets, 1983) roughly mimics the Pocock (non-spending function) design, with the key attribute that alpha is spent roughly equally across all stages.

$$\alpha(0) = 0$$

$$\alpha(p_k) = \alpha \ln(1 + (e - 1)p_k)$$

$$\alpha(1) = \alpha$$



© NCSS, LLC. All Rights Reserved.

Power Family

The power family of spending functions has a ρ parameter that gives flexibility in the spending function shape.

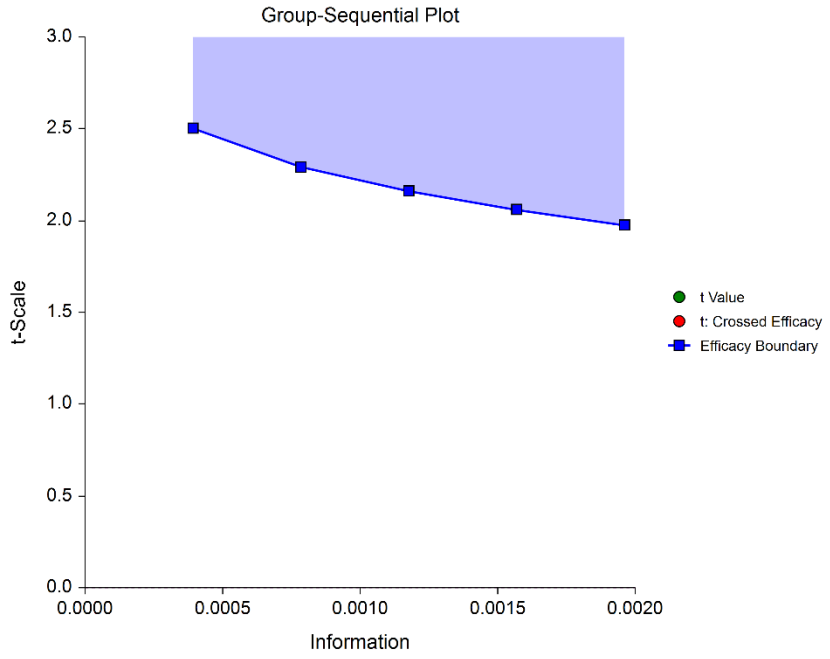
$$\alpha(0) = 0$$

$$\alpha(p_k) = p_k^{\rho}, \ \ \rho > 0$$

$$\alpha(1) = \alpha$$

A power family spending function with a ρ of 1 is similar to a Pocock design, while a power family spending function with a ρ of 3 is more similar to an O'Brien-Fleming design.

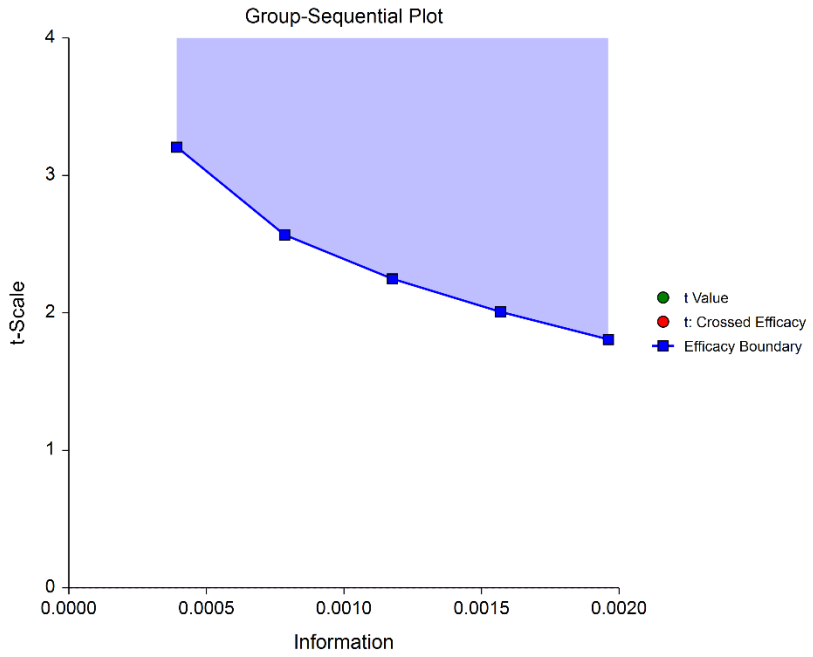
$\rho = 1$



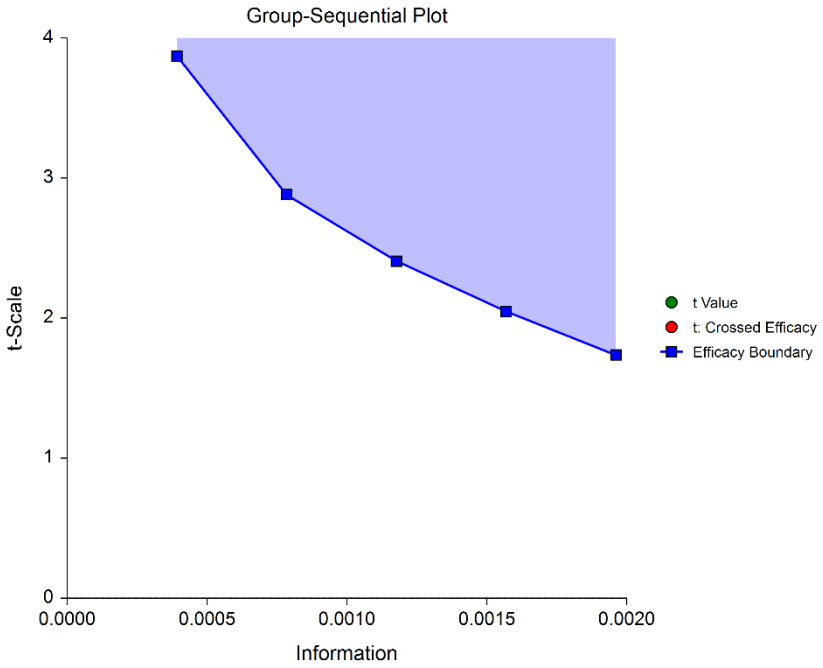
© NCSS, LLC. All Rights Reserved.

Group-Sequential T-Tests for Two Means (Simulation)

$\rho = 2$



$\rho = 3$



© NCSS, LLC. All Rights Reserved.

Hwang-Shih-DeCani (Gamma Family)

The Hwang-Shih-DeCani gamma family of spending function has a γ parameter that allows for a variety of spending functions.

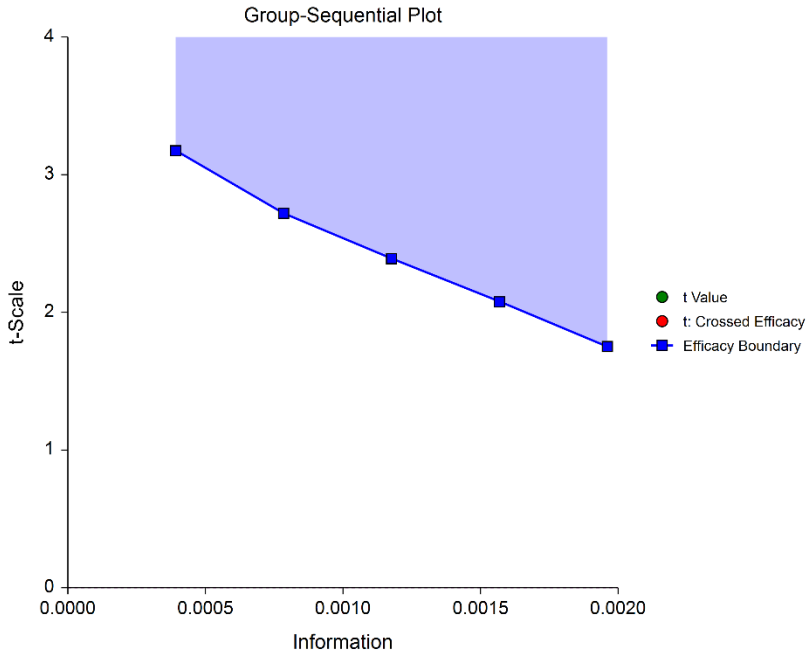
$$\alpha(0) = 0$$

$$\alpha(p_k) = \alpha\left(\frac{1 - e^{-\gamma p_k}}{1 - e^{-\gamma}}\right), \ \gamma \neq 0$$

$$\alpha(p_k) = \alpha p_k, \ \gamma = 0$$

$$\alpha(1) = \alpha$$

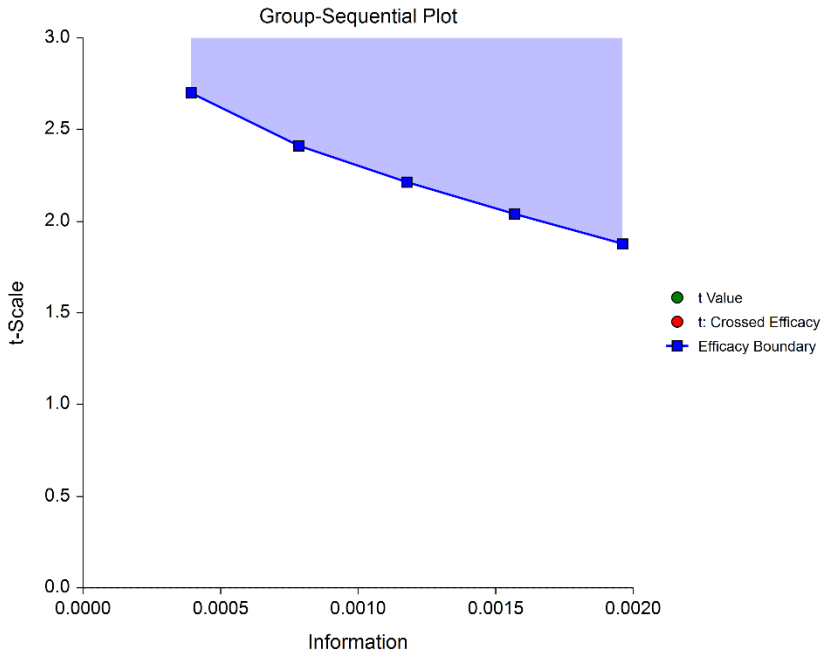
$\gamma = -3$



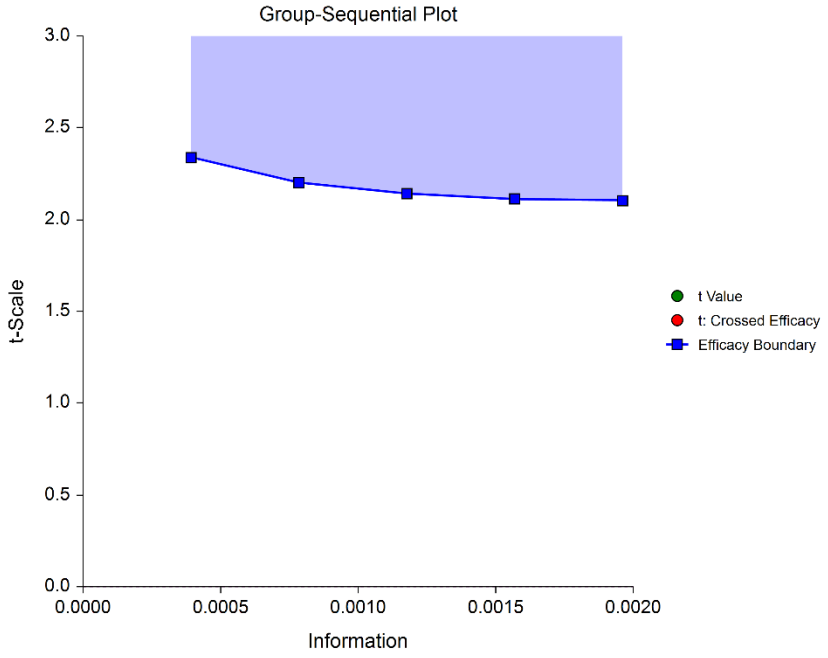
© NCSS, LLC. All Rights Reserved.

Group-Sequential T-Tests for Two Means (Simulation)

$\gamma = -1$

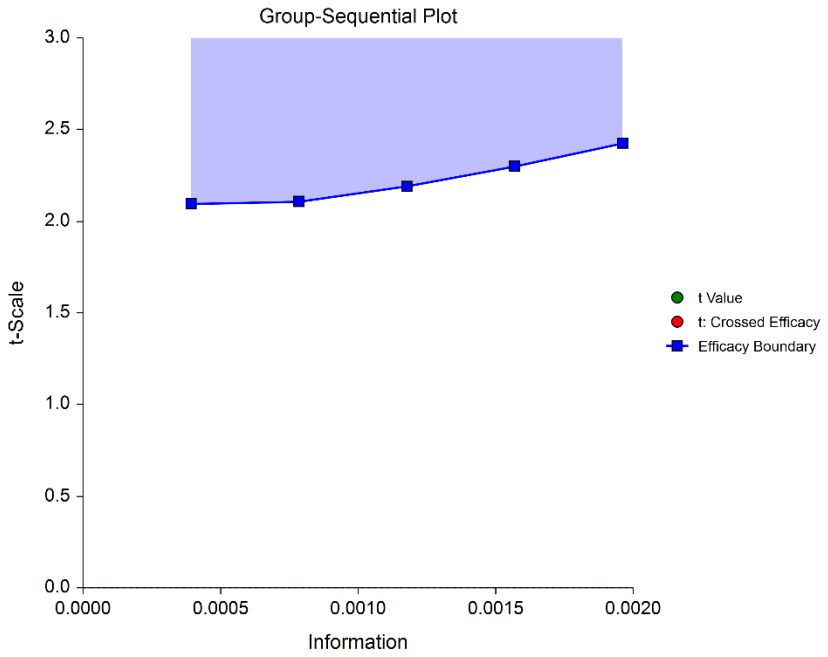


$\gamma = 1$



© NCSS, LLC. All Rights Reserved.

$\gamma = 3$



Using Simulation to obtain Boundary Crossing Probabilities

In addition to providing an overall estimate of power, it can be useful to researchers to know the probability of crossing each of the group-sequential boundaries, given a specified assumed value for the means. The following steps are used to estimate these probabilities using simulation:

1. Determine the target (cumulative) sample sizes for each stage, including the final stage. Fractional sample sizes are rounded up to the next integer.
2. For each simulation, obtain a simulated data set with the final stage sample sizes. Simulated values are generated from Normal distributions with user-specified means and planning (for Stage 0) or current standard deviations.
3. Determine whether simulation *t*-values are 'held out' after crossing a boundary, or whether simulation *t*-values are 'left in' (compared to boundaries at all future stages, regardless of whether a boundary was crossed at a previous stage).
    a. If simulation *t*-values are 'held out' after crossing a boundary, it is determined for each simulation which boundary was crossed first (except in the case of non-binding futility boundaries).
    b. If simulation *t*-values are 'left in' after crossing a boundary, it is determined for each simulation all the boundaries where the *t*-value is across the boundary.
4. The proportion of simulations crossing each boundary provides an estimate of the probability of crossing each boundary, given the specified assumed means.

© NCSS, LLC. All Rights Reserved.

5. Overall power and alpha calculations are also based on the specification of 'held out' or 'left in'.
    a. When Hold Out is selected, power and alpha are calculated as the sum of all efficacy boundary proportions.
    b. When Leave In is selected, power and alpha are calculated as the efficacy boundary proportion of the final stage.

Non-binding Futility Boundaries

When non-binding futility boundaries are used, the study may continue when a futility boundary is crossed. The simulation proportions will have a slightly different interpretation when this is the case.

© NCSS, LLC. All Rights Reserved.

Example 1 – Sample Size and Initial Boundaries for a Group-Sequential T-Test

[Please note that this example has a calculation time of several minutes to an hour or more.]

A blood pressure treatment study is conducted to determine whether a new drug results in lower patient blood pressure than a current standard. The response for each patient is a resting systolic blood pressure. A one-sided test with alpha equal to 0.025 is used.

The new drug is assigned to Group 1, and the standard is assigned to Group 2, so that the null and alternative hypotheses are

$$H_0: \mu_1 - \mu_2 = 0 \quad (H_0: \mu_{New} = \mu_{Std})$$

versus

$$H_a: \mu_1 - \mu_2 < 0 \quad (H_a: \mu_{New} < \mu_{Std})$$

The design calls for five equally spaced stages if the final stage is reached. A power of 0.90 is needed. Standard deviations are expected to be 25 for each group. The assumed mean for the standard treatment is 124. Researchers wish to examine the sample sizes needed for new treatment means of 108, 112, and 116. Both efficacy and non-binding futility boundaries are intended. The efficacy (alpha-spending) spending function used is the O'Brien-Fleming analog. The Hwang-Shih-DeCani (Gamma) beta-spending function with gamma parameter 1.5 is used for futility.

Setup

If the procedure window is not already open, use the PASS Home window to open it. The parameters for this example are listed below and are stored in the **Example 1** settings file. To load these settings to the procedure window, click **Open Example Settings File** in the Help Center or File menu.

Design Tab	
Solve For	**Sample Size**
Power	**0.90**
Alpha	**0.025**
Group Allocation	**Equal (N1 = N2)**
μ1	**108 112 116**
μ2	**124**
σ1	**25**
σ2	**σ1**
Maximum Number of Stages (K)	**5**
Info. Proportion at each Stage	**Equally incremented**
Boundaries Used	**One-sided Efficacy with Futility**
Hypothesis Direction	**Ha: μ1 - μ2 < 0**
Boundary Specification	**Spending Function Calculation**
Alpha Spending Function	**O'Brien-Fleming Analog**
Skipped Efficacy Stages	**<Empty>**

© NCSS, LLC. All Rights Reserved.

Design Beta **0.10**
Beta Spending Function..... **Hwang-Shih-DeCani (γ)**
γ..... **1.5**
Skipped Futility Stages ..... **<Empty>**
Binding or Non-Binding Futility..... **Non-Binding**

Options Tab

Number of Simulations..... **10000**
Random Seed..... **6431010** (for Reproducibility)
After Boundary Crossing..... **Hold out**

Boundary Reports Tab

All Reports ..... **Checked**

Boundary Plots Tab

t-Statistic vs Information..... **Checked**
t-Statistic vs Stage ..... **Checked**
t-Statistic vs N..... **Checked**

Summary Reports Tab

All Reports ..... **Checked**

Summary Plots Tab

All plots ..... **Checked** (only 2D will be used)

© NCSS, LLC. All Rights Reserved.

Output

Click the Calculate button to perform the calculations and generate the following output. This run will take several minutes to an hour (the number of simulations could be lowered, but the variation in results will be greater). The simulation results will differ slightly for each separate run.

Run Summary Report – Scenario 1

This report can be used to confirm that the input was processed as intended.

Run Summary Report - Scenario 1

Item	Value
Solve For	Sample Size
Maximum Number of Stages (Design)	5
Current Stage	0
Alternative Hypothesis	μ1 - μ2 < 0 (one-sided)
t-Statistic	t-Statistic Allowing Unequal Variance (Welch)
Alpha Spending Function	O'Brien-Fleming Analog
Beta Spending Function	Hwang-Shih-DeCani (γ = 1.5)
Futility Boundaries	Non-Binding
Target Alpha	0.025
Alpha (from simulations)	0.0252
μ1	108
μ2	124
σ1	25
σ2	25
N1 (if final stage reached)	55
N2 (if final stage reached)	55
Target Power	0.9
Power (from simulations)	0.9054
Maximum Information	0.044

t-Value Boundaries

This section gives the planning stage t-statistic boundaries, numerically. These values are reflected in the group-sequential boundary plot.

t-Value Boundaries

Maximum Information: 0.044
Alternative Hypothesis: μ1 - μ2 < 0 (one-sided)
Futility Boundaries: Non-Binding

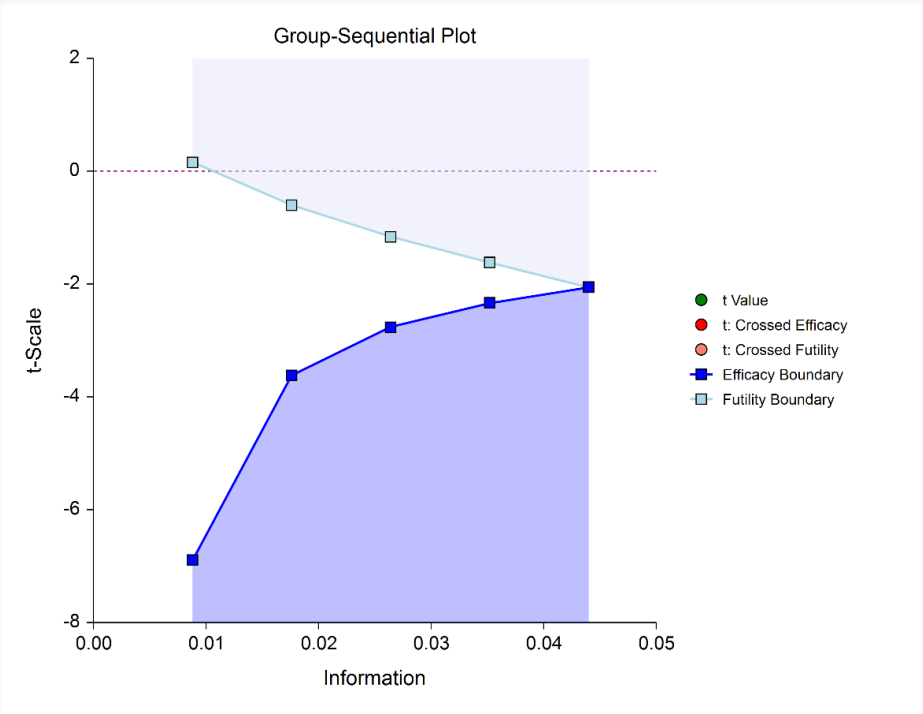
	Boundaries		
Stage	Efficacy	Futility	Information Proportion
1	-6.8910	0.1554	0.2
2	-3.6192	-0.6031	0.4
3	-2.7686	-1.1648	0.6
4	-2.3321	-1.6179	0.8
5	-2.0554	-2.0554	1.0

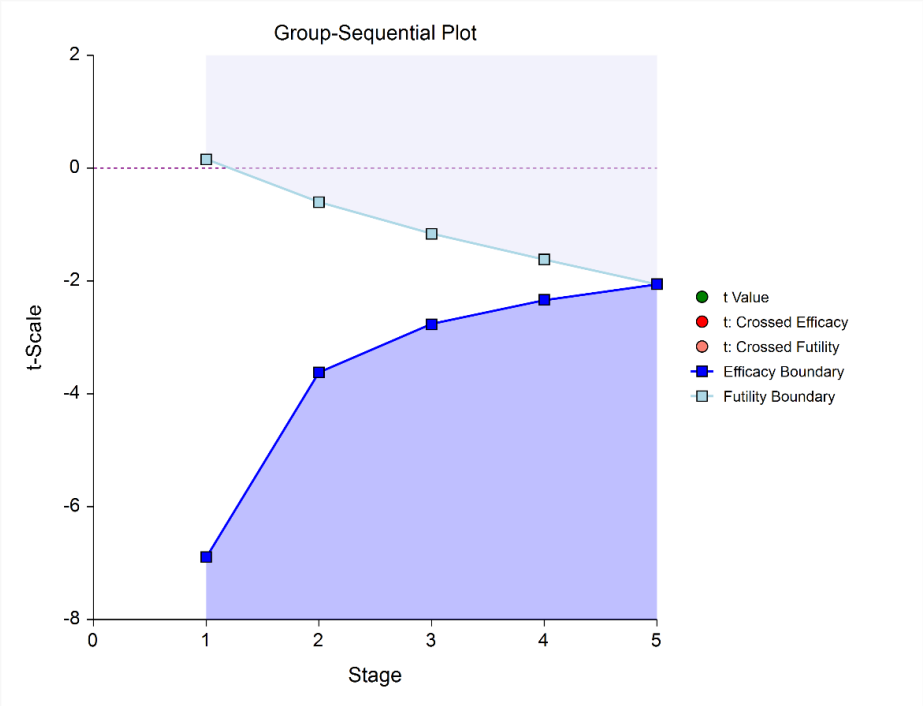
© NCSS, LLC. All Rights Reserved.

Boundary Plot(s)

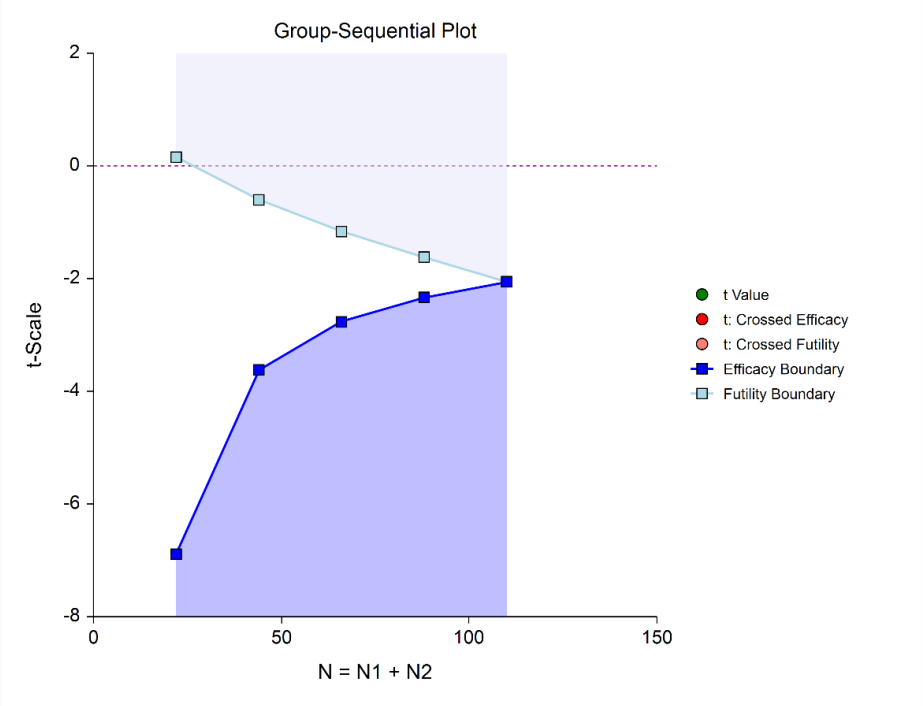
This plot shows the efficacy and futility t-statistic planning boundaries. It is anticipated that these boundaries will adjust to the actual information proportions as the data for each stage is realized.

Boundary Plot(s)





© NCSS, LLC. All Rights Reserved.



P-Value Boundaries

This section reflects the conversion of the Z-value boundaries to the corresponding P-value boundaries.

P-Value Boundaries

Maximum Information: 0.044
Alternative Hypothesis: μ1 - μ2 < 0 (one-sided)
Futility Boundaries: Non-Binding
P-value boundaries are one-sided values.

	Boundaries		
Stage	Efficacy	Futility	Information Proportion
1	0.00000	0.56095	0.2
2	0.00039	0.27484	0.4
3	0.00368	0.12421	0.6
4	0.01102	0.05468	0.8
5	0.02113	0.02113	1.0

© NCSS, LLC. All Rights Reserved.

Underlying Z Scale Boundaries

This section reflects the Z-value boundaries from which the t-value boundaries were derived.

Underlying Z Scale Boundaries

Maximum Information: 0.044
Alternative Hypothesis: μ1 - μ2 < 0 (one-sided)
Futility Boundaries: Non-Binding

	Boundaries		
Stage	Efficacy	Futility	Information Proportion
1	-4.8769	0.1534	0.2
2	-3.3569	-0.5982	0.4
3	-2.6803	-1.1542	0.6
4	-2.2898	-1.6011	0.8
5	-2.0310	-2.0310	1.0

Information Report

This section gives the target information for each stage, as well as the sample sizes and standard deviations used to calculate those informations.

Information Report

Maximum Information: 0.044
Alternative Hypothesis: μ1 - μ2 < 0 (one-sided)
Alpha: 0.025

			Target Sample Size				
Stage	Target Information Proportion	Target Information	N1	N2	S1	S2	df
1	0.2	0.0088	11	11	25	25	20
2	0.4	0.0176	22	22	25	25	42
3	0.6	0.0264	33	33	25	25	64
4	0.8	0.0352	44	44	25	25	86
5	1.0	0.0440	55	55	25	25	108

© NCSS, LLC. All Rights Reserved.

Alpha Spending

This section shows how alpha is anticipated to be spent across the stages.

Alpha Spending

Target Final Stage Alpha: 0.025
Spending Function: O'Brien-Fleming Analog

Stage	Information Proportion	Alpha Spent this Stage	Cumulative Alpha Spent	Nominal (Boundary) Alpha	Percentage Alpha Spent this Stage	Cumulative Percentage Alpha Spent
1 *	0.2	0.0000	0.0000	0.000001	0.0	0.0
2 *	0.4	0.0004	0.0004	0.000394	1.6	1.6
3 *	0.6	0.0034	0.0038	0.003678	13.7	15.2
4 *	0.8	0.0084	0.0122	0.011017	33.6	48.8
5 *	1.0	0.0128	0.0250	0.021128	51.2	100.0

* projected

Beta Spending for Futility

This section shows how beta is anticipated to be spent across the stages.

Beta Spending for Futility

Target Cumulative Beta at Final Stage: 0.1
Spending Function for Futility: Hwang-Shih-DeCani (γ = 1.5)

Stage	Information Proportion	Beta Spent this Stage	Cumulative Beta Spent	Nominal (Boundary) Beta	Percentage Beta Spent this Stage	Cumulative Percentage Beta Spent
1 *	0.2	0.0334	0.0334	0.560952	33.4	33.4
2 *	0.4	0.0247	0.0581	0.274840	24.7	58.1
3 *	0.6	0.0183	0.0764	0.124207	18.3	76.4
4 *	0.8	0.0136	0.0900	0.054676	13.6	90.0
5 *	1.0	0.0100	0.1000	0.021128	10.0	100.0

* projected

© NCSS, LLC. All Rights Reserved.

Boundary Probabilities for δ = -16

Using simulation based on the specified means and standard deviations, this section gives the estimated probabilities of crossing each of the boundaries. Values given here will vary for each simulation.

Boundary Probabilities for δ = -16

Number of Simulations:	10000
Random Seed:	6431010 (User-Entered)
Futility Boundaries:	Non-Binding
After Efficacy Boundary Crossing:	Hold Out
After Non-Binding Futility Boundary Crossing:	Leave In
Alternative Hypothesis:	μ1 - μ2 < 0 (one-sided)
μ1:	108
μ2:	124
δ:	-16
σ1:	25
σ2:	25

			Efficacy		Futility	
Stage	**N1**	**N2**	**Boundary**	**Probability**	**Boundary**	**Probability**
1	*11	*11	-6.8910	0.0000	0.1554	0.0447
2	*22	*22	-3.6192	0.0878	-0.6031	0.0628
3	*33	*33	-2.7686	0.3618	-1.1648	0.0772
4	*44	*44	-2.3321	0.3079	-1.6179	0.0847
5	*55	*55	-2.0554	0.1479	-2.0554	0.1003

* Simulation sample size (Non-integer sample sizes were rounded to the next highest integer.)

Average N1: 40.7561
Average N2: 40.7561

© NCSS, LLC. All Rights Reserved.

Boundary Probabilities for δ = 0 (Alpha)

This section estimates the probabilities of crossing each boundary if the difference for the remaining stages is assumed to be zero (the means are assumed to be the same).

Boundary Probabilities for δ = 0 (Alpha)

Number of Simulations:	10000
Random Seed:	6431010 (User-Entered)
Futility Boundaries:	Non-Binding
After Efficacy Boundary Crossing:	Hold Out
After Non-Binding Futility Boundary Crossing:	Leave In
Alternative Hypothesis:	μ1 - μ2 < 0 (one-sided)
μ1:	124
μ2:	124
δ:	0
σ1:	25
σ2:	25

			Efficacy		Futility	
Stage	**N1**	**N2**	**Boundary**	**Probability**	**Boundary**	**Probability**
1	*11	*11	-6.8910	0.0000	0.1554	0.4348
2	*22	*22	-3.6192	0.0005	-0.6031	0.7255
3	*33	*33	-2.7686	0.0031	-1.1648	0.8734
4	*44	*44	-2.3321	0.0086	-1.6179	0.9444
5	*55	*55	-2.0554	0.0130	-2.0554	0.9779

* Simulation sample size (Non-integer sample sizes were rounded to the next highest integer.)

Average N1: 54.8207
Average N2: 54.8207

Scenario 2

All of the same boundary reports are given for Scenario 2, corresponding to a μ1 value of 112.

Scenario 3

All of the same boundary reports are given for Scenario 3, corresponding to a μ1 value of 116.

© NCSS, LLC. All Rights Reserved.

Power and Sample Size Summary

Power and Sample Size Summary

Solve For:	Sample Size
Maximum Number of Stages:	5
Alternative Hypothesis:	μ1 - μ2 < 0 (one-sided)
Alpha Spending Function:	O'Brien-Fleming Analog
Beta Spending Function:	Hwang-Shih-DeCani (γ = 1.5)
Number of Simulations:	10000
Random Seed:	6431010 (User-Entered)

Target Power	Sim Power	N1	N2	N	μ1	μ2	σ1	σ2	Target Alpha	Sim Alpha
0.9	0.9054	55	55	110	108	124	25	25	0.025	0.025
0.9	0.9078	96	96	192	112	124	25	25	0.025	0.027
0.9	0.9078	217	217	434	116	124	25	25	0.025	0.026

Target Power	The desired power value (or values) entered in the procedure.
Sim Power	The proportion of simulation t-values that cross an efficacy boundary. Because "After Boundary Crossing" is set to "Hold out," it is the sum of the individual boundary crossing proportions.
N1 and N2	The anticipated number of individuals in each group if the final stage is reached.
N	The total sample size if the final stage is reached. N = N1 + N2.
μ1	The assumed mean of population 1 for power calculation simulations.
μ2	The assumed mean of population 2 for power calculation simulations. μ2 is also the assumed mean of populations 1 and 2 for alpha calculation simulations.
σ1 and σ2	The assumed population standard deviations of groups 1 and 2. σ1 and σ2 are used for simulating values as well as for the calculation of maximum information.
Target Alpha	The alpha used in the computation of the boundaries. The desired overall probability of a Type 1 error.
Sim Alpha	The proportion of null simulation t-values that cross an efficacy boundary. Because "After Boundary Crossing" is set to "Hold out," it is the sum of the individual boundary crossing proportions.

Summary Statements

A parallel, two-group, group-sequential design with a maximum of 5 stages will be used to test whether the Group 1 mean is less than the Group 2 mean (H0: μ1 - μ2 ≥ 0 versus H1: μ1 - μ2 < 0). The comparison will be made at each stage using a two-sample, unequal-variance (Welch's) t-test, with efficacy and futility boundary values calculated from the designated spending functions. The target cumulative Type I error rate (α) at the final stage is 0.025. The standard deviation for Group 1 is assumed to be 25 and the standard deviation for Group 2 is assumed to be 25. To detect a Group 1 mean of 108 and a Group 2 mean of 124 (difference of -16), with 90% power, the number of needed subjects at the final stage will be 55 in Group 1 and 55 in Group 2.

References

Jennison, C. and Turnbull, B.W. 2000. Group Sequential Methods with Applications to Clinical Trials. Chapman and Hall/CRC. Boca Raton.

Lan, K.K.G. and DeMets, D.L. 1983. 'Discrete sequential boundaries for clinical trials.' Biometrika, 70, pages 659-663.

Reboussin, D.M., DeMets, D.L., Kim, K., and Lan, K.K.G. 1992. 'Programs for computing group sequential boundaries using the Lan-DeMets Method.' Technical Report 60, Department of Biostatistics, University of Wisconsin-Madison.

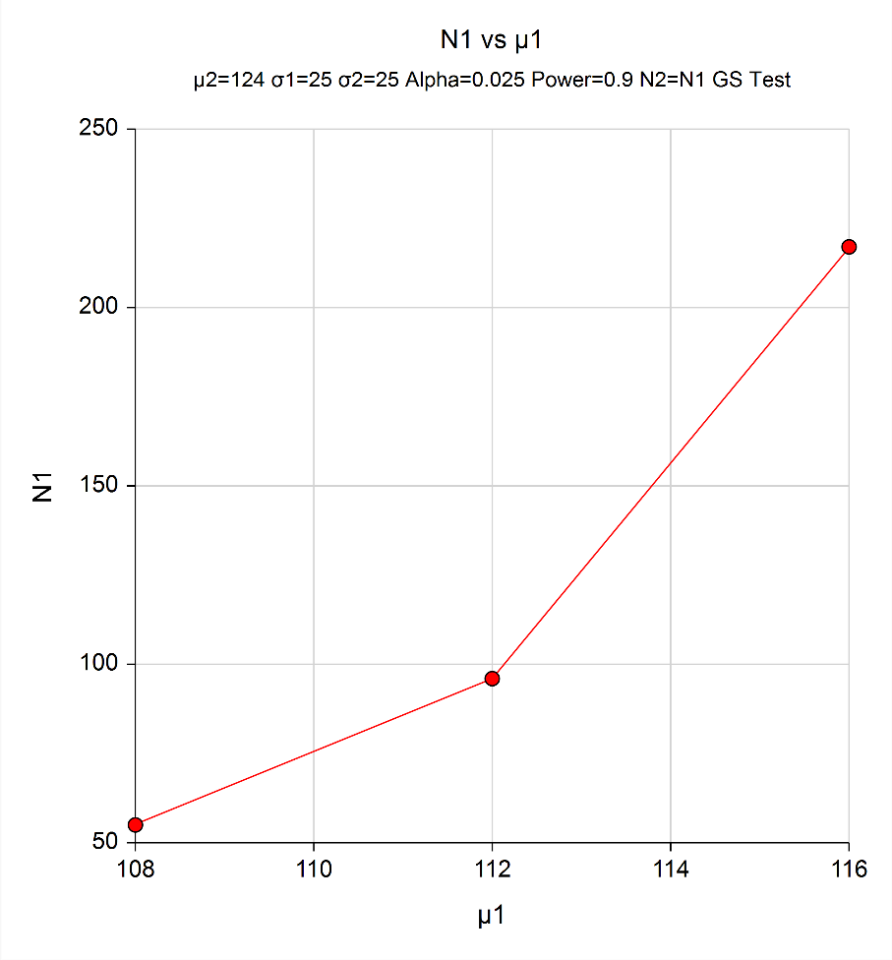
This report shows the values of each of the parameters, one scenario per row. The values may vary slightly due to the variation in simulations. The details for each of the rows of this report are given in the earlier boundary reports.

The values from this table are exhibited in the plot below.

© NCSS, LLC. All Rights Reserved.

Plots Section for Power and Sample Size Summary

Plots



© NCSS, LLC. All Rights Reserved.

Example 2 – Skipping Stage Boundaries

[Please note that this example has a calculation time of several minutes to an hour or more.]

Suppose that the scenario is exactly as in Example 1, except that the first two futility boundaries are to be skipped.

Setup

If the procedure window is not already open, use the PASS Home window to open it. The parameters for this example are listed below and are stored in the **Example 2** settings file. To load these settings to the procedure window, click **Open Example Settings File** in the Help Center or File menu.

Design Tab	
Solve For	**Sample Size**
Power	**0.90**
Alpha	**0.025**
Group Allocation	**Equal (N1 = N2)**
μ1	**108 112 116**
μ2	**124**
σ1	**25**
σ2	**σ1**
Maximum Number of Stages (K)	**5**
Info. Proportion at each Stage	**Equally incremented**
Boundaries Used	**One-sided Efficacy with Futility**
Hypothesis Direction	**Ha: μ1 - μ2 < 0**
Boundary Specification	**Spending Function Calculation**
Alpha Spending Function	**O'Brien-Fleming Analog**
Skipped Efficacy Stages	**<Empty>**
Design Beta	**0.10**
Beta Spending Function	**Hwang-Shih-DeCani (γ)**
γ	**1.5**
Skipped Futility Stages	**1 2**
Binding or Non-Binding Futility	**Non-Binding**

Options Tab	
Number of Simulations	**10000**
Random Seed	**6006159** (for Reproducibility)
After Boundary Crossing	**Hold out**

Boundary Reports Tab	
All Reports	**Checked**

Boundary Plots Tab	
t-Statistic vs Information	**Checked**
t-Statistic vs Stage	**Checked**
t-Statistic vs N	**Checked**

© NCSS, LLC. All Rights Reserved.

Summary Reports Tab

All Reports ..... **Checked**

Summary Plots Tab

All plots ..... **Checked** (only 2D will be used)

Output

Click the Calculate button to perform the calculations and generate the following output. This run will take several minutes to an hour (the number of simulations could be lowered, but the variation in results will be greater). The simulation results will differ slightly for each separate run.

Run Summary Report

Run Summary Report - Scenario 1

Item	Value
Solve For	Sample Size
Maximum Number of Stages (Design)	5
Skipped Futility Stage(s)	1 2
Current Stage	0
Alternative Hypothesis	μ1 - μ2 < 0 (one-sided)
t-Statistic	t-Statistic Allowing Unequal Variance (Welch)
Alpha Spending Function	O'Brien-Fleming Analog
Beta Spending Function	Hwang-Shih-DeCani (γ = 1.5)
Futility Boundaries	Non-Binding
Target Alpha	0.025
Alpha (from simulations)	0.0245
μ1	108
μ2	124
σ1	25
σ2	25
N1 (if final stage reached)	54
N2 (if final stage reached)	54
Target Power	0.9
Power (from simulations)	0.9007
Maximum Information	0.0432

© NCSS, LLC. All Rights Reserved.

t-Value Boundaries

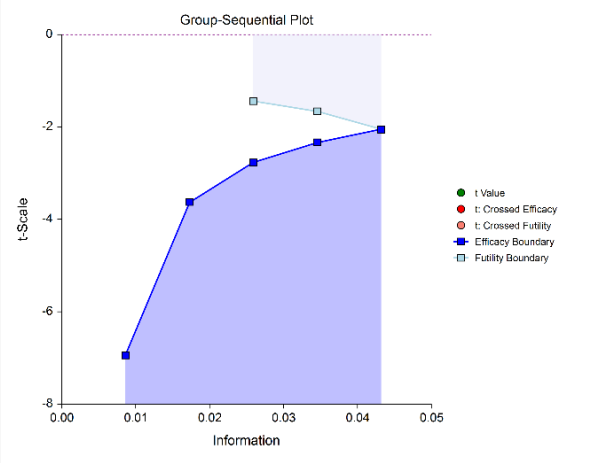
t-Value Boundaries

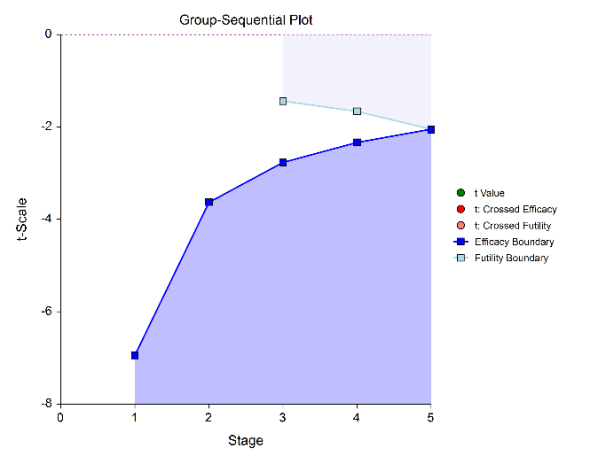
Maximum Information: 0.0432
Alternative Hypothesis: μ1 - μ2 < 0 (one-sided)
Futility Boundaries: Non-Binding

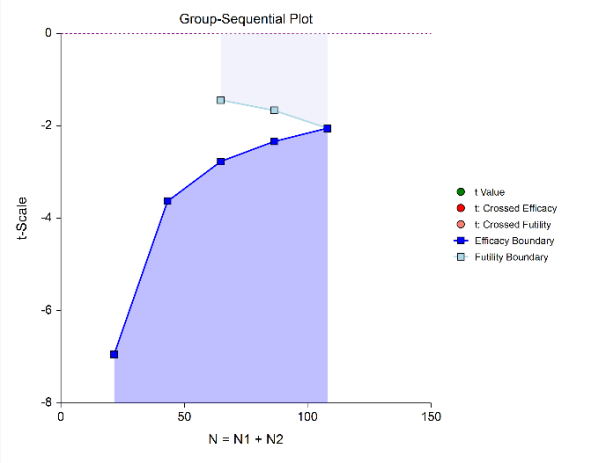
	Boundaries		
Stage	Efficacy	Futility	Information Proportion
1	-6.9449		0.2
2	-3.6246		0.4
3	-2.7703	-1.4406	0.6
4	-2.3329	-1.6625	0.8
5	-2.0558	-2.0558	1.0

Boundary Plot(s)

Boundary Plot(s)







© NCSS, LLC. All Rights Reserved.

P-Value Boundaries

P-Value Boundaries

Maximum Information: 0.0432
Alternative Hypothesis: μ1 - μ2 < 0 (one-sided)
Futility Boundaries: Non-Binding
P-value boundaries are one-sided values.

	Boundaries		
Stage	Efficacy	Futility	Information Proportion
1	0.00000		0.2
2	0.00039		0.4
3	0.00368	0.07733	0.6
4	0.01102	0.05006	0.8
5	0.02113	0.02113	1.0

Underlying Z Scale Boundaries

Underlying Z Scale Boundaries

Maximum Information: 0.0432
Alternative Hypothesis: μ1 - μ2 < 0 (one-sided)
Futility Boundaries: Non-Binding

	Boundaries		
Stage	Efficacy	Futility	Information Proportion
1	-4.8769		0.2
2	-3.3569		0.4
3	-2.6803	-1.4232	0.6
4	-2.2898	-1.6443	0.8
5	-2.0310	-2.0310	1.0

Information Report

Information Report

Maximum Information: 0.0432
Alternative Hypothesis: μ1 - μ2 < 0 (one-sided)
Alpha: 0.025

			Target Sample Size				
Stage	Target Information Proportion	Target Information	N1	N2	S1	S2	df
1	0.2	0.0086	10.8	10.8	25	25	19.6
2	0.4	0.0173	21.6	21.6	25	25	41.2
3	0.6	0.0259	32.4	32.4	25	25	62.8
4	0.8	0.0346	43.2	43.2	25	25	84.4
5	1.0	0.0432	54.0	54.0	25	25	106.0

© NCSS, LLC. All Rights Reserved.

Alpha Spending

Alpha Spending

Target Final Stage Alpha: 0.025
Spending Function: O'Brien-Fleming Analog

Stage	Information Proportion	Alpha Spent this Stage	Cumulative Alpha Spent	Nominal (Boundary) Alpha	Percentage Alpha Spent this Stage	Cumulative Percentage Alpha Spent
1 *	0.2	0.0000	0.0000	0.000001	0.0	0.0
2 *	0.4	0.0004	0.0004	0.000394	1.6	1.6
3 *	0.6	0.0034	0.0038	0.003678	13.7	15.2
4 *	0.8	0.0084	0.0122	0.011017	33.6	48.8
5 *	1.0	0.0128	0.0250	0.021128	51.2	100.0

* projected

Beta Spending for Futility

Beta Spending for Futility

Target Cumulative Beta at Final Stage: 0.1
Spending Function for Futility: Hwang-Shih-DeCani (γ = 1.5)

Stage	Information Proportion	Beta Spent this Stage	Cumulative Beta Spent	Nominal (Boundary) Beta	Percentage Beta Spent this Stage	Cumulative Percentage Beta Spent
1 *	0.2	0.0000	0.0000	1.000000	0.0	0.0
2 *	0.4	0.0000	0.0000	1.000000	0.0	0.0
3 *	0.6	0.0764	0.0764	0.077334	76.4	76.4
4 *	0.8	0.0136	0.0900	0.050056	13.6	90.0
5 *	1.0	0.0100	0.1000	0.021128	10.0	100.0

* projected

© NCSS, LLC. All Rights Reserved.

Boundary Probabilities for δ = -16

Boundary Probabilities for δ = -16

Number of Simulations:	10000
Random Seed:	6006159 (User-Entered)
Futility Boundaries:	Non-Binding
After Efficacy Boundary Crossing:	Hold Out
After Non-Binding Futility Boundary Crossing:	Leave In
Alternative Hypothesis:	μ1 - μ2 < 0 (one-sided)
μ1:	108
μ2:	124
δ:	-16
σ1:	25
σ2:	25

			Efficacy		Futility	
Stage	**N1**	**N2**	**Boundary**	**Probability**	**Boundary**	**Probability**
1	*10.8	*10.8	-6.9449	0.0002		0.0000
2	*21.6	*21.6	-3.6246	0.0869		0.0000
3	*32.4	*32.4	-2.7703	0.3566	-1.4406	0.1218
4	*43.2	*43.2	-2.3329	0.3108	-1.6625	0.0938
5	*54.0	*54.0	-2.0558	0.1462	-2.0558	0.1043

* Simulation sample size (Non-integer sample sizes were rounded to the next highest integer.)

Average N1: 40.1166
Average N2: 40.1166

Boundary Probabilities for δ = 0 (Alpha)

Boundary Probabilities for δ = 0 (Alpha)

Number of Simulations:	10000
Random Seed:	6006159 (User-Entered)
Futility Boundaries:	Non-Binding
After Efficacy Boundary Crossing:	Hold Out
After Non-Binding Futility Boundary Crossing:	Leave In
Alternative Hypothesis:	μ1 - μ2 < 0 (one-sided)
μ1:	124
μ2:	124
δ:	0
σ1:	25
σ2:	25

			Efficacy		Futility	
Stage	**N1**	**N2**	**Boundary**	**Probability**	**Boundary**	**Probability**
1	*10.8	*10.8	-6.9449	0.0000		0.0000
2	*21.6	*21.6	-3.6246	0.0005		0.0000
3	*32.4	*32.4	-2.7703	0.0027	-1.4406	0.9246
4	*43.2	*43.2	-2.3329	0.0086	-1.6625	0.9489
5	*54.0	*54.0	-2.0558	0.0127	-2.0558	0.9794

* Simulation sample size (Non-integer sample sizes were rounded to the next highest integer.)

Average N1: 53.8326
Average N2: 53.8326

© NCSS, LLC. All Rights Reserved.

Scenario 2

All of the same boundary reports are given for Scenario 2, corresponding to a μ1 value of 112.

Scenario 3

All of the same boundary reports are given for Scenario 3, corresponding to a μ1 value of 116.

Power and Sample Size Summary

Power and Sample Size Summary

Solve For:	Sample Size
Maximum Number of Stages:	5
Skipped Futility Stage(s):	1 2
Alternative Hypothesis:	μ1 - μ2 < 0 (one-sided)
Alpha Spending Function:	O'Brien-Fleming Analog
Beta Spending Function:	Hwang-Shih-DeCani (γ = 1.5)
Number of Simulations:	10
Random Seed:	6006159 (User-Entered)

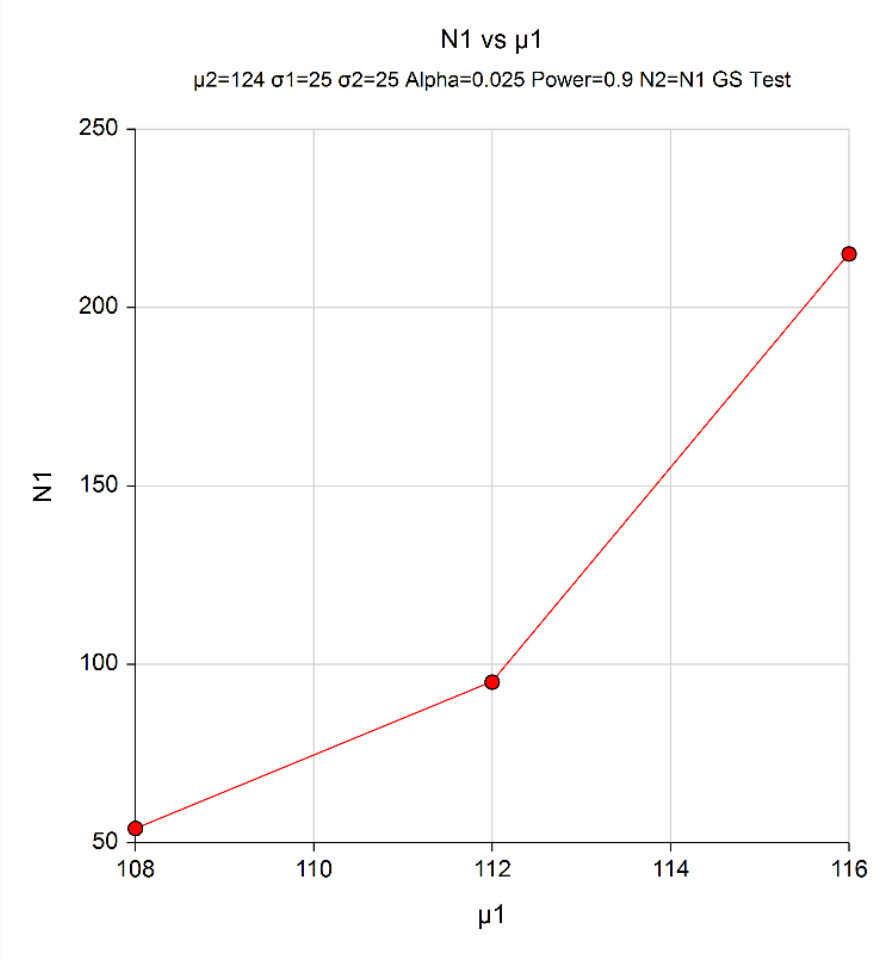
Target Power	Sim Power	N1	N2	N	μ1	μ2	σ1	σ2	Target Alpha	Sim Alpha
0.9	0.9007	54	54	108	108	124	25	25	0.025	0.025
0.9	0.9008	95	95	190	112	124	25	25	0.025	0.027
0.9	0.9075	215	215	430	116	124	25	25	0.025	0.027

This report shows little change in overall sample size, with the difference possibly due to simulation variability.

© NCSS, LLC. All Rights Reserved.

Plots Section for Power and Sample Size Summary

Plots



© NCSS, LLC. All Rights Reserved.

Example 3 – Finding Power

[Please note that this example has a calculation time of several minutes to an hour or more.]

Suppose that the scenario is similar to the setup of Example 1, except that now we will solve for power for various sample sizes of 60, 80, 100, 120, and 140 per group.

Setup

If the procedure window is not already open, use the PASS Home window to open it. The parameters for this example are listed below and are stored in the **Example 3** settings file. To load these settings to the procedure window, click **Open Example Settings File** in the Help Center or File menu.

Design Tab	
Solve For	**Power**
Alpha	**0.025**
Group Allocation	**Equal (N1 = N2)**
Sample Size Per Group	**60 80 100 120 140**
μ1	**108 112 116**
μ2	**124**
σ1	**25**
σ2	**σ1**
Maximum Number of Stages (K)	**5**
Info. Proportion at each Stage	**Equally incremented**
Boundaries Used	**One-sided Efficacy with Futility**
Hypothesis Direction	**Ha: μ1 - μ2 < 0**
Boundary Specification	**Spending Function Calculation**
Alpha Spending Function	**O'Brien-Fleming Analog**
Skipped Efficacy Stages	**<Empty>**
Design Beta	**0.10**
Beta Spending Function	**Hwang-Shih-DeCani (γ)**
γ	**1.5**
Skipped Futility Stages	**<Empty>**
Binding or Non-Binding Futility	**Non-Binding**

Options Tab	
Number of Simulations	**10000**
Random Seed	**6336849** (for Reproducibility)
After Boundary Crossing	**Hold out**

Boundary Reports Tab	
All Reports	**Checked**

Boundary Plots Tab	
t-Statistic vs Information	**Checked**
t-Statistic vs Stage	**Checked**
t-Statistic vs N	**Checked**

© NCSS, LLC. All Rights Reserved.

Summary Reports Tab

All Reports ..... **Checked**

Summary Plots Tab

All plots ..... **Checked**

Output

Click the Calculate button to perform the calculations and generate the following output. This run will take several minutes to an hour (the number of simulations could be lowered, but the variation in results will be greater). The simulation results will differ slightly for each separate run.

Scenario Reports

All of the scenario reports for each of the 15 scenarios are generated in the output, but they are not shown here.

Power and Sample Size Summary

This summary is shown near the bottom of the output. The navigation pane may be useful to find it quickly.

Power and Sample Size Summary

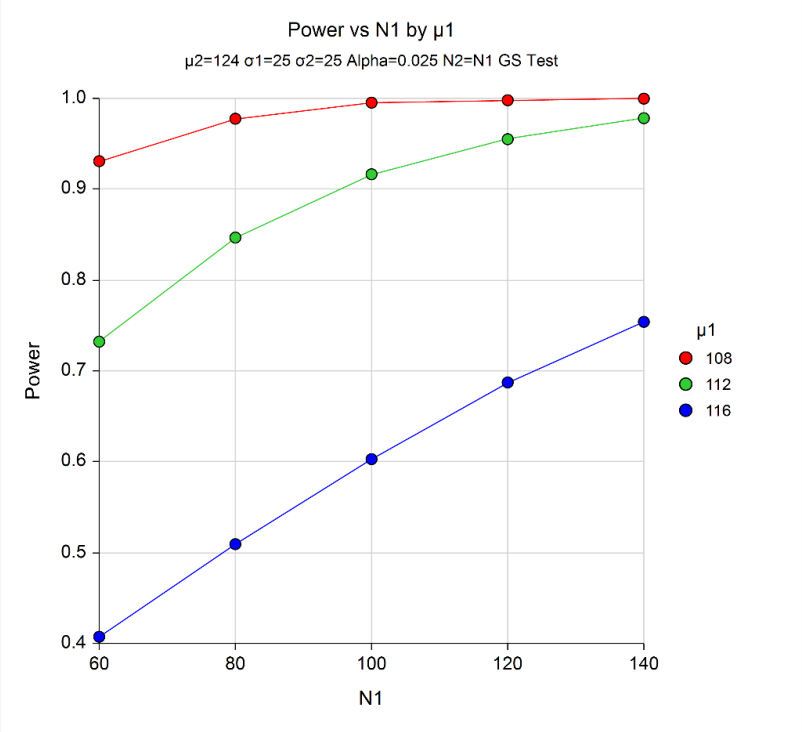
Solve For:	Power
Maximum Number of Stages:	5
Alternative Hypothesis:	μ1 - μ2 < 0 (one-sided)
Alpha Spending Function:	O'Brien-Fleming Analog
Beta Spending Function:	Hwang-Shih-DeCani (γ = 1.5)
Number of Simulations:	10000
Random Seed:	6336849 (User-Entered)

Sim Power	N1	N2	N	μ1	μ2	σ1	σ2	Target Alpha	Sim Alpha
0.9309	60	60	120	108	124	25	25	0.025	0.023
0.9776	80	80	160	108	124	25	25	0.025	0.024
0.9953	100	100	200	108	124	25	25	0.025	0.023
0.9980	120	120	240	108	124	25	25	0.025	0.022
0.9997	140	140	280	108	124	25	25	0.025	0.026
0.7322	60	60	120	112	124	25	25	0.025	0.025
0.8468	80	80	160	112	124	25	25	0.025	0.024
0.9165	100	100	200	112	124	25	25	0.025	0.026
0.9551	120	120	240	112	124	25	25	0.025	0.023
0.9782	140	140	280	112	124	25	25	0.025	0.023
0.4074	60	60	120	116	124	25	25	0.025	0.029
0.5093	80	80	160	116	124	25	25	0.025	0.027
0.6029	100	100	200	116	124	25	25	0.025	0.023
0.6873	120	120	240	116	124	25	25	0.025	0.028
0.7540	140	140	280	116	124	25	25	0.025	0.026

© NCSS, LLC. All Rights Reserved.

Plots Section for Power and Sample Size Summary

Plots



The power curve plot shows the effect of sample size and mean difference on the power for each scenario.

© NCSS, LLC. All Rights Reserved.

Example 4 – Finding Power with Binding Futility Boundaries

[Please note that this example has a calculation time of several minutes to an hour or more.]

Following the setup of Example 3, we wish to see the effect on power of changing from non-binding futility boundaries to binding futility boundaries.

Setup

If the procedure window is not already open, use the PASS Home window to open it. The parameters for this example are listed below and are stored in the **Example 4** settings file. To load these settings to the procedure window, click **Open Example Settings File** in the Help Center or File menu.

Design Tab	
Solve For	**Power**
Alpha	**0.025**
Group Allocation	**Equal (N1 = N2)**
Sample Size Per Group	**60 80 100 120 140**
μ1	**108 112 116**
μ2	**124**
σ1	**25**
σ2	**σ1**
Maximum Number of Stages (K)	**5**
Info. Proportion at each Stage	**Equally incremented**
Boundaries Used	**One-sided Efficacy with Futility**
Hypothesis Direction	**Ha: μ1 - μ2 < 0**
Boundary Specification	**Spending Function Calculation**
Alpha Spending Function	**O'Brien-Fleming Analog**
Skipped Efficacy Stages	**<Empty>**
Design Beta	**0.10**
Beta Spending Function	**Hwang-Shih-DeCani (γ)**
γ	**1.5**
Skipped Futility Stages	**<Empty>**
Binding or Non-Binding Futility	**Binding**

Options Tab	
Number of Simulations	**10000**
Random Seed	**7622145** (for Reproducibility)
After Boundary Crossing	**Hold out**

Boundary Reports Tab	
All Reports	**Checked**

© NCSS, LLC. All Rights Reserved.

Boundary Plots Tab

t-Statistic vs Information.....**Checked**
t-Statistic vs Stage .....**Checked**
t-Statistic vs N.....**Checked**

Summary Reports Tab

All Reports .....**Checked**

Summary Plots Tab

All plots .....**Checked**

Output

Click the Calculate button to perform the calculations and generate the following output. The simulation results will differ slightly for each separate run.

Scenario Reports

All the scenario reports for each of the 15 scenarios are generated in the output, but they are not shown here.

Power and Sample Size Summary

Power and Sample Size Summary

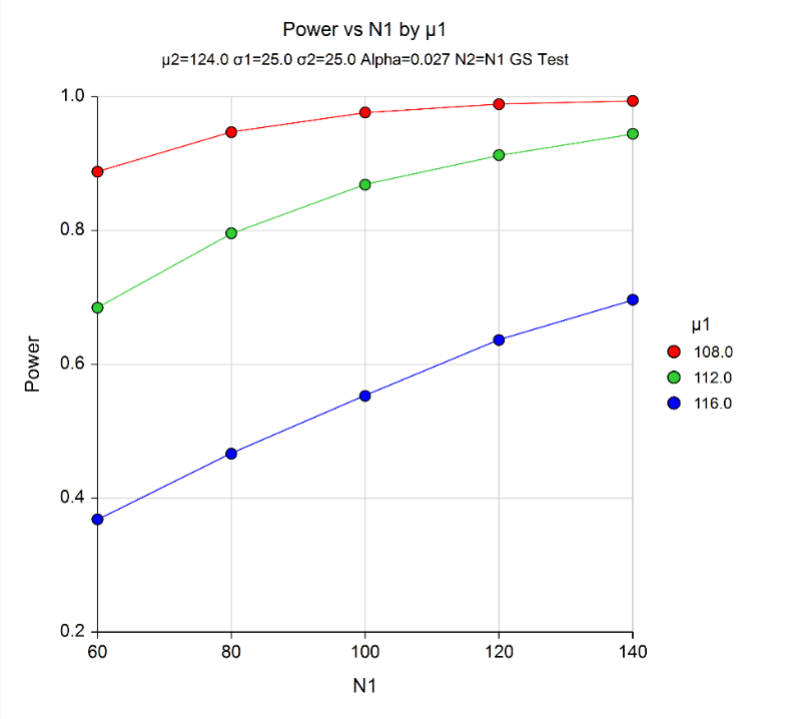
Solve For:	Power
Maximum Number of Stages:	5
Alternative Hypothesis:	μ1 - μ2 < 0 (one-sided)
Alpha Spending Function:	O'Brien-Fleming Analog
Beta Spending Function:	Hwang-Shih-DeCani (γ = 1.5)
Number of Simulations:	10000
Random Seed:	7622145 (User-Entered)

Sim Power	N1	N2	N	μ1	μ2	σ1	σ2	Target Alpha	Sim Alpha
0.8810	60	60	120	108	124	25	25	0.025	0.023
0.9475	80	80	160	108	124	25	25	0.025	0.029
0.9752	100	100	200	108	124	25	25	0.025	0.025
0.9875	120	120	240	108	124	25	25	0.025	0.024
0.9946	140	140	280	108	124	25	25	0.025	0.025
0.6747	60	60	120	112	124	25	25	0.025	0.027
0.7911	80	80	160	112	124	25	25	0.025	0.026
0.8731	100	100	200	112	124	25	25	0.025	0.026
0.9165	120	120	240	112	124	25	25	0.025	0.026
0.9459	140	140	280	112	124	25	25	0.025	0.025
0.3779	60	60	120	116	124	25	25	0.025	0.026
0.4710	80	80	160	116	124	25	25	0.025	0.027
0.5577	100	100	200	116	124	25	25	0.025	0.025
0.6207	120	120	240	116	124	25	25	0.025	0.024
0.6975	140	140	280	116	124	25	25	0.025	0.027

© NCSS, LLC. All Rights Reserved.

Plots Section for Power and Sample Size Summary

Plots



If the power results are compared to those with non-binding futility boundaries in Example 3, it is seen that the power for binding futility boundaries is several percent lower. Higher numbers of simulations might be used to fine-tune these differences.

© NCSS, LLC. All Rights Reserved.

Example 5 – Comparing Numbers of Stages

[Please note that each run will take a few to several minutes.]

Following the setup of Example 3, we wish to see the effect on power of changing the number of stages. This requires multiple runs with different numbers of stages. The numbers of stages examined here are 2, 3, 4, 5, 10 and 20. A mean for Group 1 of 112 and a sample size of 100 per group will be considered.

Setup

If the procedure window is not already open, use the PASS Home window to open it. The parameters for this example are listed below and are stored in the **Example 5** settings file. To load these settings to the procedure window, click **Open Example Settings File** in the Help Center or File menu.

Design Tab	
Solve For	**Power**
Alpha	**0.025**
Group Allocation	**Equal (N1 = N2)**
Sample Size Per Group	**100**
μ1	**112**
μ2	**124**
σ1	**25**
σ2	**σ1**
Maximum Number of Stages (K)	**2** (Also run with **3**, **4**, **5**, **10**, and **20**)
Info. Proportion at each Stage	**Equally incremented**
Boundaries Used	**One-sided Efficacy with Futility**
Hypothesis Direction	**Ha: μ1 - μ2 < 0**
Boundary Specification	**Spending Function Calculation**
Alpha Spending Function	**O'Brien-Fleming Analog**
Skipped Efficacy Stages	**<Empty>**
Design Beta	**0.10**
Beta Spending Function	**Hwang-Shih-DeCani (γ)**
γ	**1.5**
Skipped Futility Stages	**<Empty>**
Binding or Non-Binding Futility	**Non-Binding**

Options Tab	
Number of Simulations	**10000**
Random Seed	**Random**
After Boundary Crossing	**Hold out**

Boundary Reports Tab	
All Reports	**Checked**

© NCSS, LLC. All Rights Reserved.

Boundary Plots Tab	
t-Statistic vs Information	**Checked**
t-Statistic vs Stage	**Checked**
t-Statistic vs N	**Checked**

Summary Reports Tab	
All Reports	**Checked**

Summary Plots Tab	
All plots	**Checked**

Output

Click the Calculate button to perform the calculations and generate the following output. Each run will likely take a few minutes. The simulation results will differ slightly for each separate run.

Scenario Reports

All the scenario reports for each of the 6 scenarios are generated in the output, but they are not shown here.

Power and Sample Size Summary

Power and Sample Size Summary

Solve For:	Power
Maximum Number of Stages:	2 3 4 5 10 20 (separate runs)
Alternative Hypothesis:	μ1 - μ2 < 0 (one-sided)
Alpha Spending Function :	O'Brien-Fleming Analog
Beta Spending Function :	Hwang-Shih-DeCani (γ = 1.5)
Number of Simulations :	10000

Sim Power	N1	N2	N	μ1	μ2	σ1	σ2	Target Alpha	Sim Alpha	Number of Stages
0.92150	100	100	200	112	124	25	25	0.025	0.025	2
0.92020	100	100	200	112	124	25	25	0.025	0.026	3
0.91700	100	100	200	112	124	25	25	0.025	0.028	4
0.91550	100	100	200	112	124	25	25	0.025	0.028	5
0.91300	100	100	200	112	124	25	25	0.025	0.024	10
0.90980	100	100	200	112	124	25	25	0.025	0.028	20

It is seen that the impact of the number of stages on the overall power is very minor.

© NCSS, LLC. All Rights Reserved.

Example 6 – Two-Sided Boundaries

[Please note that this example has a calculation time of several minutes to an hour or more.]

Suppose that the scenario is similar to the setup of Example 3, but the boundary structure is changed to two-sided boundaries with an alpha of 0.05.

Setup

If the procedure window is not already open, use the PASS Home window to open it. The parameters for this example are listed below and are stored in the **Example 6** settings file. To load these settings to the procedure window, click **Open Example Settings File** in the Help Center or File menu.

Design Tab	
Solve For	**Power**
Alpha	**0.05**
Group Allocation	**Equal (N1 = N2)**
Sample Size Per Group	**60 80 100 120 140**
μ1	**108 112 116**
μ2	**124**
σ1	**25**
σ2	**σ1**
Maximum Number of Stages (K)	**5**
Info. Proportion at each Stage	**Equally incremented**
Boundaries Used	**Two-sided Efficacy with Futility (Symmetric)**
Boundary Specification	**Spending Function Calculation**
Alpha Spending Function	**O'Brien-Fleming Analog**
Skipped Efficacy Stages	**<Empty>**
Design Beta	**0.10**
Beta Spending Function	**Hwang-Shih-DeCani (γ)**
γ	**1.5**
Skipped Futility Stages	**<Empty>**
Binding or Non-Binding Futility	**Non-Binding**
Overlapped Futility Boundaries	**Remove (skip) overlapped futility boundaries**

Options Tab	
Number of Simulations	**10000**
Random Seed	**5712421** (for Reproducibility)
After Boundary Crossing	**Hold out**

Boundary Reports Tab	
All Reports	**Checked**

Boundary Plots Tab	
t-Statistic vs Information	**Checked**
t-Statistic vs Stage	**Checked**
t-Statistic vs N	**Checked**

© NCSS, LLC. All Rights Reserved.

Summary Reports Tab

All Reports ..... **Checked**

Summary Plots Tab

All plots ..... **Checked**

Output

Click the Calculate button to perform the calculations and generate the following output. The simulation results will differ slightly for each separate run.

Run Summary Report – Scenario 1

Run Summary Report - Scenario 1

Item	Value
Solve For	Power
Maximum Number of Stages (Design)	5
Current Stage	0
Alternative Hypothesis	μ1 - μ2 ≠ 0 (two-sided, symmetric)
t-Statistic	t-Statistic Allowing Unequal Variance (Welch)
Alpha Spending Function	O'Brien-Fleming Analog
Beta Spending Function	Hwang-Shih-DeCani (γ = 1.5)
Futility Boundaries	Non-Binding
Target Alpha	0.05
Alpha (from simulations)	0.1
μ1	108
μ2	124
σ1	25
σ2	25
N1 (if final stage reached)	60
N2 (if final stage reached)	60
Power (from simulations)	0.8586
Maximum Information	0.048

© NCSS, LLC. All Rights Reserved.

t-Value Boundaries – Scenario 1

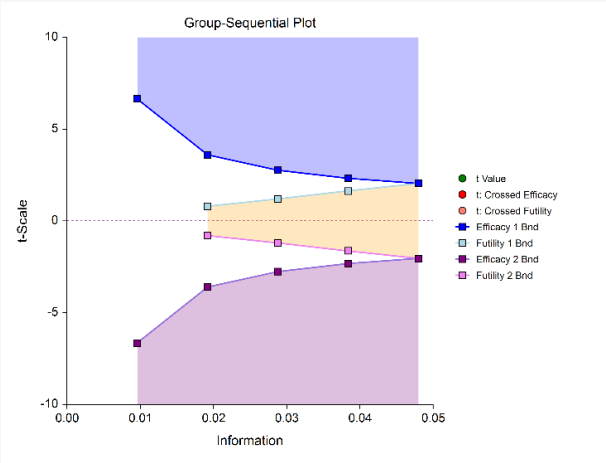
t-Value Boundaries

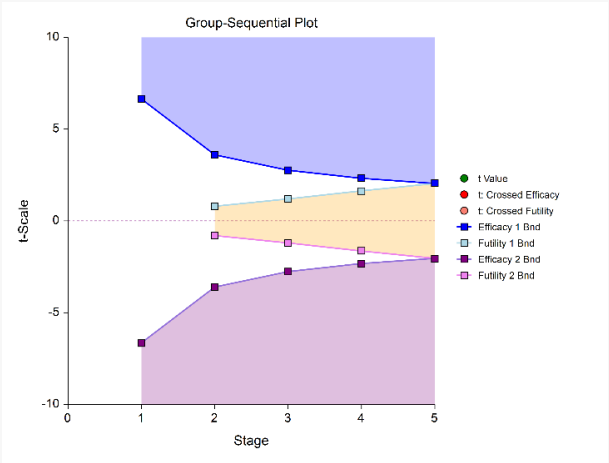
Maximum Information: 0.048
Alternative Hypothesis: μ1 - μ2 ≠ 0 (two-sided, symmetric)
Futility Boundaries: Non-Binding

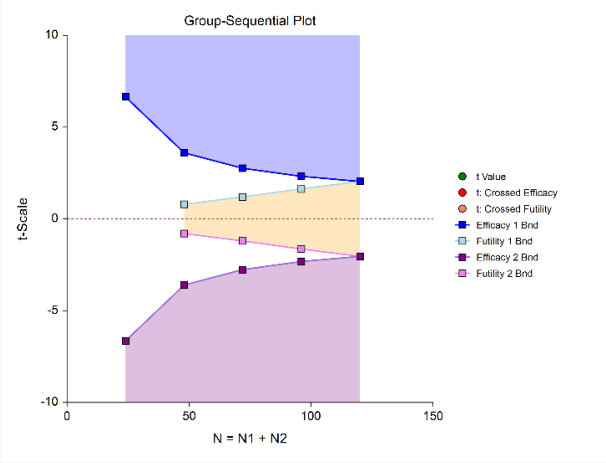
	Boundaries				
	Upper Side		Lower Side		
Stage	Efficacy 1	Futility 1	Efficacy 2	Futility 2	Information Proportion
1	6.6581		-6.6581		0.2
2	3.5949	0.7977	-3.5949	-0.7977	0.4
3	2.7608	1.1999	-2.7608	-1.1999	0.6
4	2.3284	1.6260	-2.3284	-1.6260	0.8
5	2.0533	2.0533	-2.0533	-2.0533	1.0

Boundary Plot(s) – Scenario 1

Boundary Plot(s)







© NCSS, LLC. All Rights Reserved.

Other Scenario Reports

All of the scenario reports for each of the 15 scenarios are generated in the output, but they are not shown here.

Power and Sample Size Summary

Power and Sample Size Summary

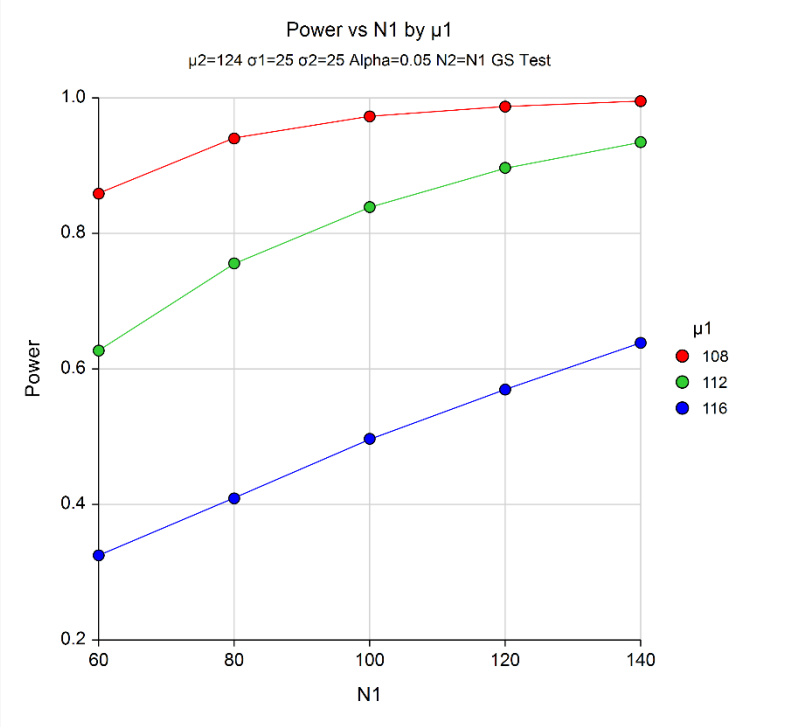
Solve For:	Power
Maximum Number of Stages:	5
Alternative Hypothesis:	μ1 - μ2 ≠ 0 (two-sided, symmetric)
Alpha Spending Function:	O'Brien-Fleming Analog
Beta Spending Function:	Hwang-Shih-DeCani (γ = 1.5)
Number of Simulations:	10000
Random Seed:	5712421 (User-Entered)

Sim Power	N1	N2	N	μ1	μ2	σ1	σ2	Target Alpha	Sim Alpha
0.8586	60	60	120	108	124	25	25	0.05	0.039
0.9400	80	80	160	108	124	25	25	0.05	0.032
0.9726	100	100	200	108	124	25	25	0.05	0.034
0.9870	120	120	240	108	124	25	25	0.05	0.033
0.9951	140	140	280	108	124	25	25	0.05	0.036
0.6270	60	60	120	112	124	25	25	0.05	0.035
0.7557	80	80	160	112	124	25	25	0.05	0.041
0.8388	100	100	200	112	124	25	25	0.05	0.036
0.8967	120	120	240	112	124	25	25	0.05	0.034
0.9346	140	140	280	112	124	25	25	0.05	0.036
0.3249	60	60	120	116	124	25	25	0.05	0.037
0.4089	80	80	160	116	124	25	25	0.05	0.038
0.4968	100	100	200	116	124	25	25	0.05	0.036
0.5695	120	120	240	116	124	25	25	0.05	0.038
0.6384	140	140	280	116	124	25	25	0.05	0.038

© NCSS, LLC. All Rights Reserved.

Plots Section for Power and Sample Size Summary

Plots



The power curve plot shows the effect of sample size and mean difference on the power for each scenario.

© NCSS, LLC. All Rights Reserved.

Example 7 – Validation Using O'Brien-Fleming Analog Boundaries

Reboussin (1992) presents an example for normally distributed data for a design with two-sided O'Brien-Fleming boundaries, 5 stages, alpha = 0.05, beta = 0.10, μ1 = 220, μ2 = 200, σ1 and σ2 = 30. They compute a sample size of 48.41 per group. The upper boundaries are: 4.8769, 3.3569, 2.6803, 2.2898, 2.0310. To show that **PASS** provides the same result for the boundaries, use the setup below. It is anticipated that the sample size will be slightly different, due to this procedure using the t-test rather than the z-test of Reboussin (1992).

[Please note that this example will take a half hour or so to compute, unless the number of simulations is decreased substantially.]

Setup

If the procedure window is not already open, use the PASS Home window to open it. The parameters for this example are listed below and are stored in the **Example 7** settings file. To load these settings to the procedure window, click **Open Example Settings File** in the Help Center or File menu.

Design Tab	
Solve For	**Sample Size**
Power	**0.90**
Alpha	**0.05**
Group Allocation	**Equal (N1 = N2)**
μ1	**220**
μ2	**200**
σ1	**30**
σ2	**σ1**
Maximum Number of Stages (K)	**5**
Info. Proportion at each Stage	**Equally incremented**
Boundaries Used	**Two-sided Efficacy Only (Symmetric)**
Boundary Specification	**Spending Function Calculation**
Alpha Spending Function	**O'Brien-Fleming Analog**
Skipped Efficacy Stages	**<Empty>**

Options Tab	
Number of Simulations	**100000**
Random Seed	**3159710** (for Reproducibility)
After Boundary Crossing	**Hold out**

Boundary Reports Tab	
Boundaries using Z Scale	**Checked**
Underlying Z Scale Boundaries	**Checked**
Information Report	**Checked**
Alpha Spending	**Checked**

© NCSS, LLC. All Rights Reserved.

Boundary Plots Tab

t-Statistic vs N ..... **Checked**

Summary Reports Tab

All Reports ..... **Checked**

Output

Click the Calculate button to perform the calculations and generate the following output. The simulation results will differ slightly for each separate run.

Run Summary Report

Run Summary Report

Item	Value
Solve For	Sample Size
Maximum Number of Stages (Design)	5
Current Stage	0
Alternative Hypothesis	μ1 - μ2 ≠ 0 (two-sided, symmetric)
t-Statistic	t-Statistic Allowing Unequal Variance (Welch)
Alpha Spending Function	O'Brien-Fleming Analog
Target Alpha	0.05
Alpha (from simulations)	0.05045
μ1	220
μ2	200
σ1	30
σ2	30
N1 (if final stage reached)	40
N2 (if final stage reached)	40
Target Power	0.9
Power (from simulations)	0.90515
Maximum Information	0.0278

t-Value Boundaries

t-Value Boundaries

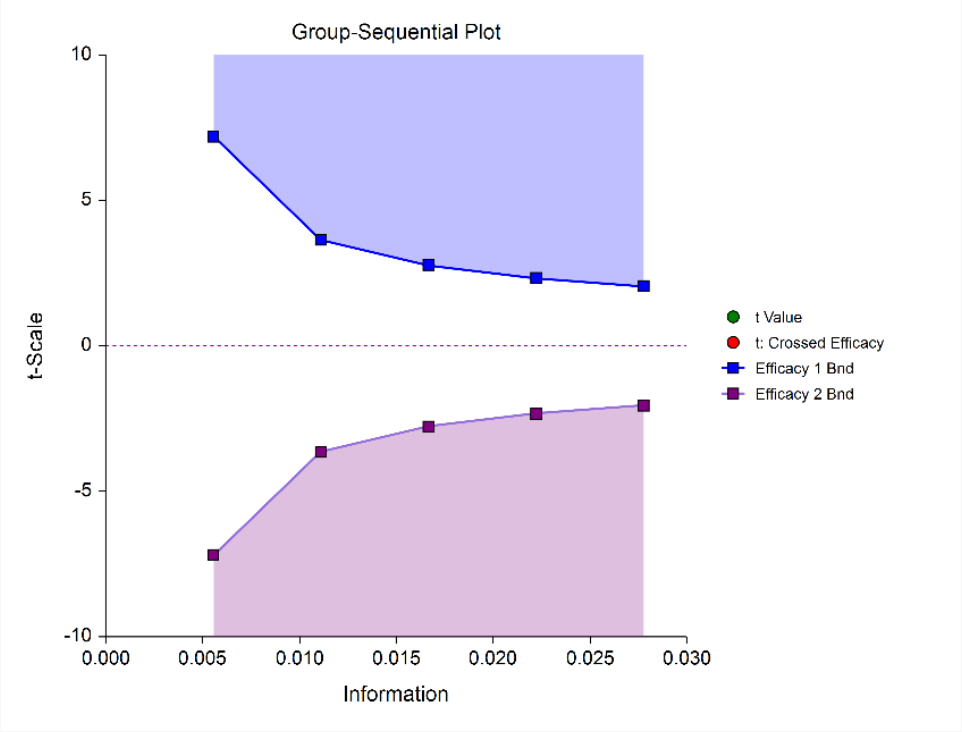
Maximum Information: 0.0222
Alternative Hypothesis: μ1 - μ2 ≠ 0 (two-sided, symmetric)

	Boundaries		
Stage	Efficacy 1	Efficacy 2	Information Proportion
1	7.1923	-7.1923	0.2
2	3.6489	-3.6489	0.4
3	2.7780	-2.7780	0.6
4	2.3365	-2.3365	0.8
5	2.0579	-2.0579	1.0

© NCSS, LLC. All Rights Reserved.

Boundary Plot(s)

Boundary Plot(s)



Underlying Z Scale Boundaries

Underlying Z Scale Boundaries

Maximum Information: 0.0278
Alternative Hypothesis: μ1 - μ2 ≠ 0 (two-sided, symmetric)

	Boundaries		
Stage	Efficacy 1	Efficacy 2	Information Proportion
1	4.8769	-4.8769	0.2
2	3.3569	-3.3569	0.4
3	2.6803	-2.6803	0.6
4	2.2898	-2.2898	0.8
5	2.0310	-2.0310	1.0

The boundary values match those of Reboussin (1992) exactly.

© NCSS, LLC. All Rights Reserved.

Information Report

Information Report

Maximum Information:	0.0278
Alternative Hypothesis:	μ1 - μ2 ≠ 0 (two-sided, symmetric)
Alpha:	0.05

Stage	Target Information Proportion	Target Information	Target Sample Size N1	N2	S1	S2	df
1	0.2	0.0056	10	10	30	30	18
2	0.4	0.0111	20	20	30	30	38
3	0.6	0.0167	30	30	30	30	58
4	0.8	0.0222	40	40	30	30	78
5	1.0	0.0278	50	50	30	30	98

Alpha Spending

Alpha Spending

Target Final Stage Alpha:	0.05
Spending Function:	O'Brien-Fleming Analog

Stage	Information Proportion	Alpha Spent this Stage	Cumulative Alpha Spent	(1-sided) Nominal (Boundary) Alpha	Percentage Alpha Spent this Stage	Cumulative Percentage Alpha Spent
1 *	0.2	0.0000	0.0000	0.000001	0.0	0.0
2 *	0.4	0.0008	0.0008	0.000394	1.6	1.6
3 *	0.6	0.0068	0.0076	0.003678	13.7	15.2
4 *	0.8	0.0168	0.0244	0.011017	33.6	48.8
5 *	1.0	0.0256	0.0500	0.021128	51.2	100.0

* projected

© NCSS, LLC. All Rights Reserved.

Boundary Probabilities for δ = 20

Boundary Probabilities for δ = 20

Number of Simulations:	100000
Random Seed:	3159710 (User-Entered)
After Efficacy Boundary Crossing:	Hold Out
Alternative Hypothesis:	μ1 - μ2 ≠ 0 (two-sided, symmetric)
μ1:	220
μ2:	200
δ:	20
σ1:	30
σ2:	30

			Efficacy 1		Efficacy 2	
Stage	**N1**	**N2**	**Boundary**	**Probability**	**Boundary**	**Probability**
1	*10	*10	7.1923	0.0001	-7.1923	0
2	*20	*20	3.6489	0.0817	-3.6489	0
3	*30	*30	2.7780	0.3506	-2.7780	0
4	*40	*40	2.3365	0.3149	-2.3365	0
5	*50	*50	2.0579	0.1579	-2.0579	0

* Simulation sample size (Non-integer sample sizes were rounded to the next highest integer.)

Average N1: 37.3852
Average N2: 37.3852

Boundary Probabilities for δ = 0 (Alpha)

Boundary Probabilities for δ = 0 (Alpha)

Number of Simulations:	100000
Random Seed:	3159710 (User-Entered)
After Efficacy Boundary Crossing:	Hold Out
Alternative Hypothesis:	μ1 - μ2 ≠ 0 (two-sided, symmetric)
μ1:	200
μ2:	200
δ:	0
σ1:	30
σ2:	30

			Efficacy 1		Efficacy 2	
Stage	**N1**	**N2**	**Boundary**	**Probability**	**Boundary**	**Probability**
1	*10	*10	7.1923	0.0000	-7.1923	0.0000
2	*20	*20	3.6489	0.0004	-3.6489	0.0004
3	*30	*30	2.7780	0.0035	-2.7780	0.0035
4	*40	*40	2.3365	0.0087	-2.3365	0.0088
5	*50	*50	2.0579	0.0126	-2.0579	0.0125

* Simulation sample size (Non-integer sample sizes were rounded to the next highest integer.)

Average N1: 49.6603
Average N2: 49.6603

© NCSS, LLC. All Rights Reserved.

Power and Sample Size Summary

Power and Sample Size Summary

Solve For:	Sample Size
Maximum Number of Stages:	5
Alternative Hypothesis:	μ1 - μ2 ≠ 0 (two-sided, symmetric)
Alpha Spending Function:	O'Brien-Fleming Analog
Number of Simulations:	100000
Random Seed:	3159710 (User-Entered)

Target Power	Sim Power	N1	N2	N	μ1	μ2	σ1	σ2	Target Alpha	Sim Alpha
0.9	0.90515	50	50	100	220	200	30	30	0.05	0.05

The sample size of 50 per group is very close to the value of 48.41 per group given by Reboussin (1992).

© NCSS, LLC. All Rights Reserved.