Chapter 720

Group-Sequential Analysis for Two Hazard Rates

Note: The corresponding sample size procedure, found in PASS Sample Size software, is *Group-Sequential Tests for Two Hazard Rates (Simulation)*.

Introduction

This procedure is used to test the difference of two hazard rates in stages (sometimes called looks or interim analyses) using group-sequential methods. This methodology assumes an underlying Exponential model. Unless the stage boundaries are entered directly, the stage boundaries are defined using a specified spending function. One- or two-sided tests may be performed with the option of binding or non-binding futility boundaries. Futility boundaries are specified through a beta-spending function.

Sample size re-estimation, based on current-stage sample sizes and parameter estimates, may also be obtained in this procedure.

The spending functions available in this procedure are the O'Brien-Fleming analog, the Pocock analog, the Hwang-Shih-DeCani gamma family, and the power family.

At each stage, the current and future boundaries are calculated based on the accumulated information proportion. Conditional and predictive power for future stages is also given.

A group-sequential boundary and analysis graph is produced in this procedure.
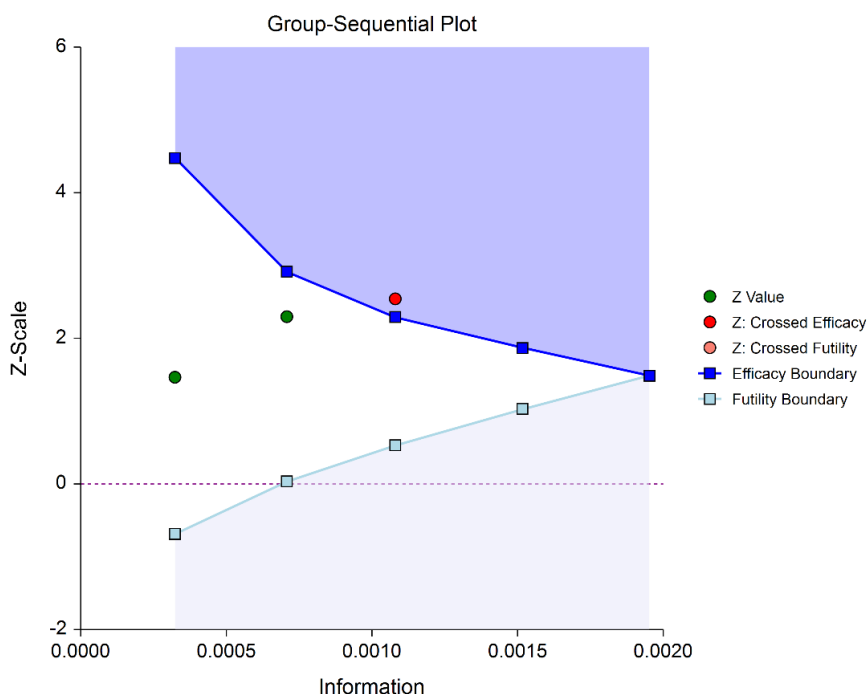
At each stage, stage-adjusted difference estimates, confidence intervals, and p-values are available.

The probabilities of crossing future boundaries may also be assessed, using simulation.

The format of the data for use in this procedure is three columns: one column for the response values, one column defining the two groups, and a third column defining the stage.

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

- Assumed survival rates

The design phase calculation may be done in the **PASS** sample size software program. **PASS** software permits the user to easily try a range of hazard rate differences, 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:
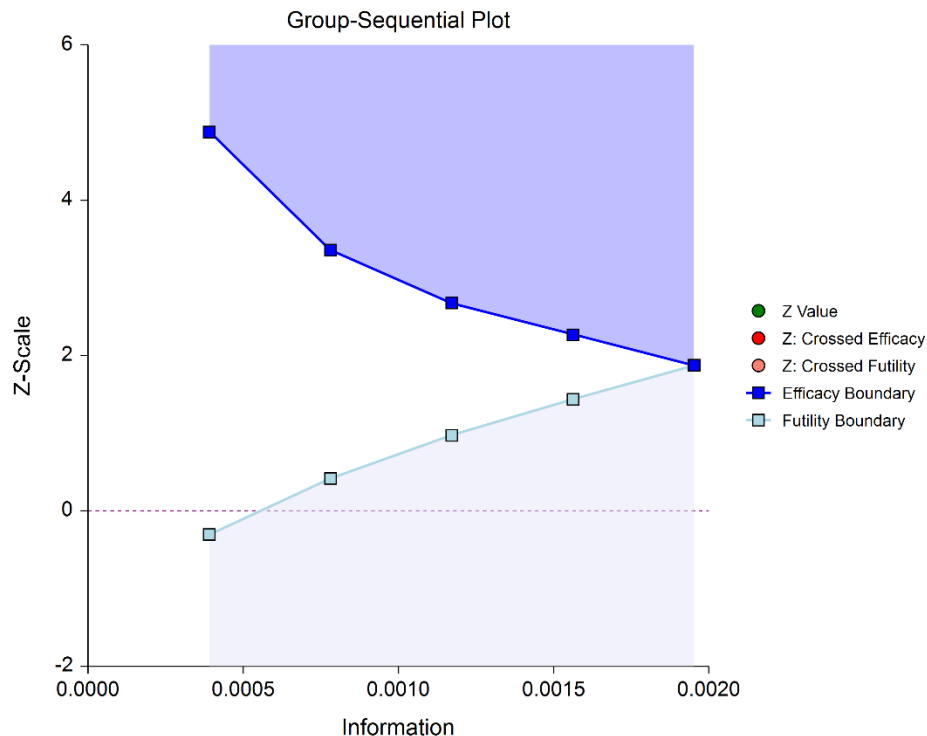
Group-Sequential Analysis Phase

A group sequential analysis consists of a series of stages where a decision to stop or continue is made at each stage.

First Interim Stage

The design phase gives the target number of subjects for the first stage, based on the time of the first stage and the accrual specification. 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. 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 *z*-statistic is calculated from the raw hazard rate difference, or other logrank-type method. The stage one *z*-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).

Group-Sequential Analysis for Two Hazard Rates
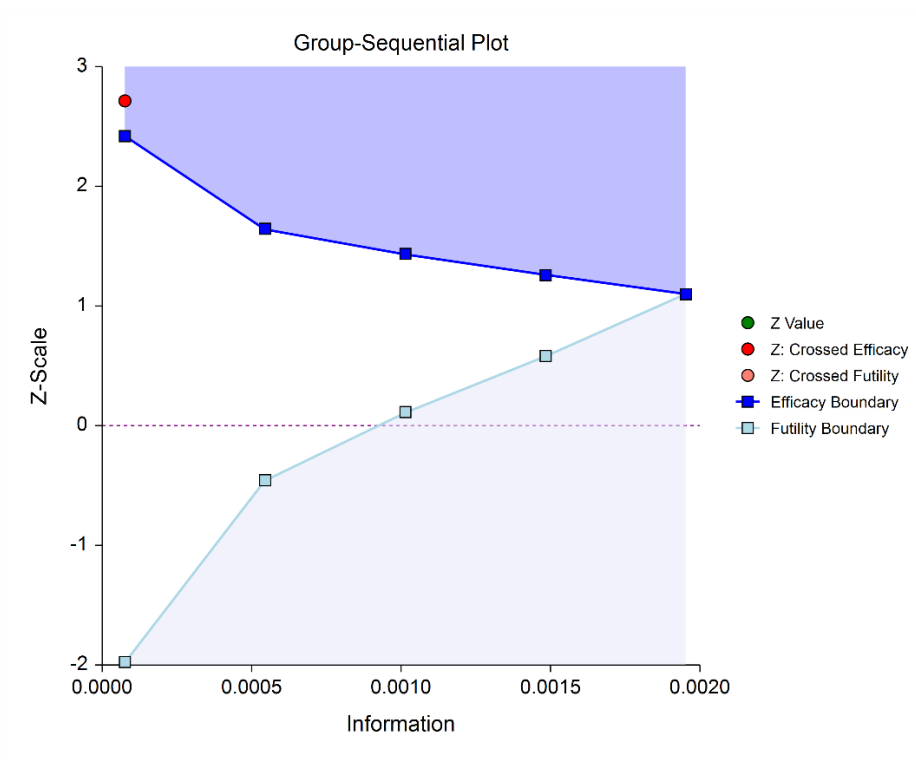
If none of the boundaries are crossed the study continues to the next stage.
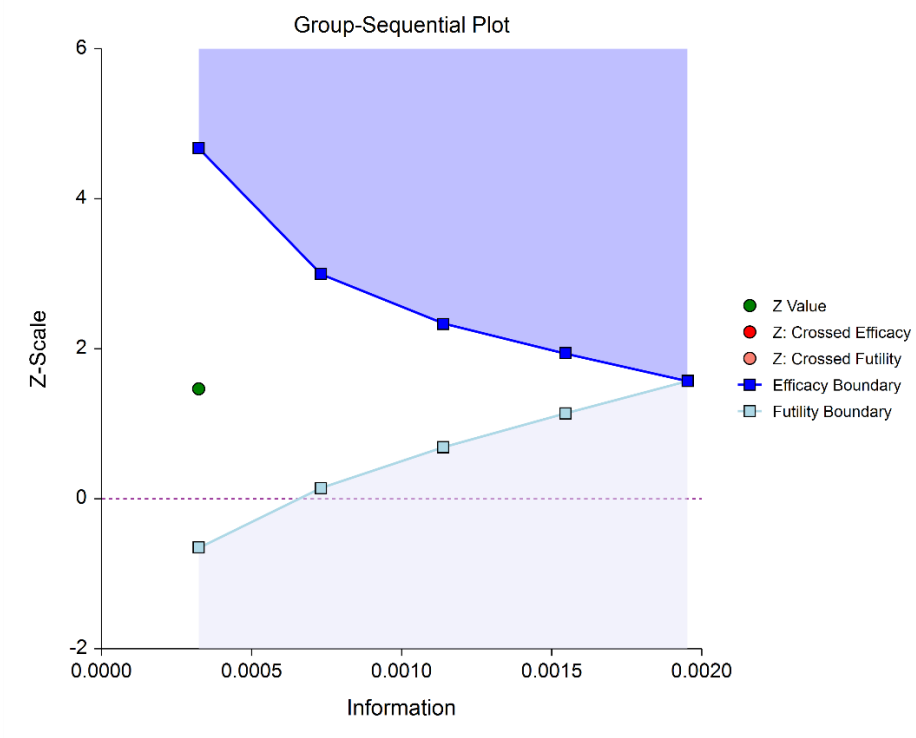
If none of the boundaries are crossed it may also be useful to examine the conditional power or stopping probabilities of future stages. Conditional power and stopping probabilities are based on the user-specified supposed true difference.

Second and other interim stages (if reached)

If the first stage time proportion is not equal to the design time proportion, a designation must be made at this point as to the target time of the second stage. Two options are available in this 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, 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 z-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.
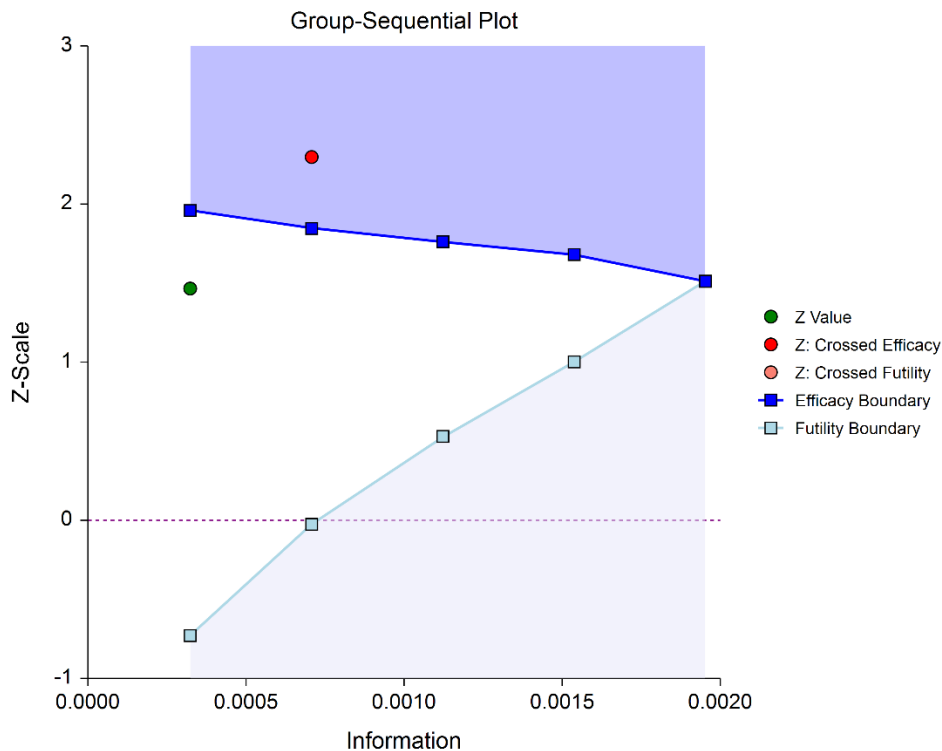
If none of the boundaries are crossed the study continues to the next stage.
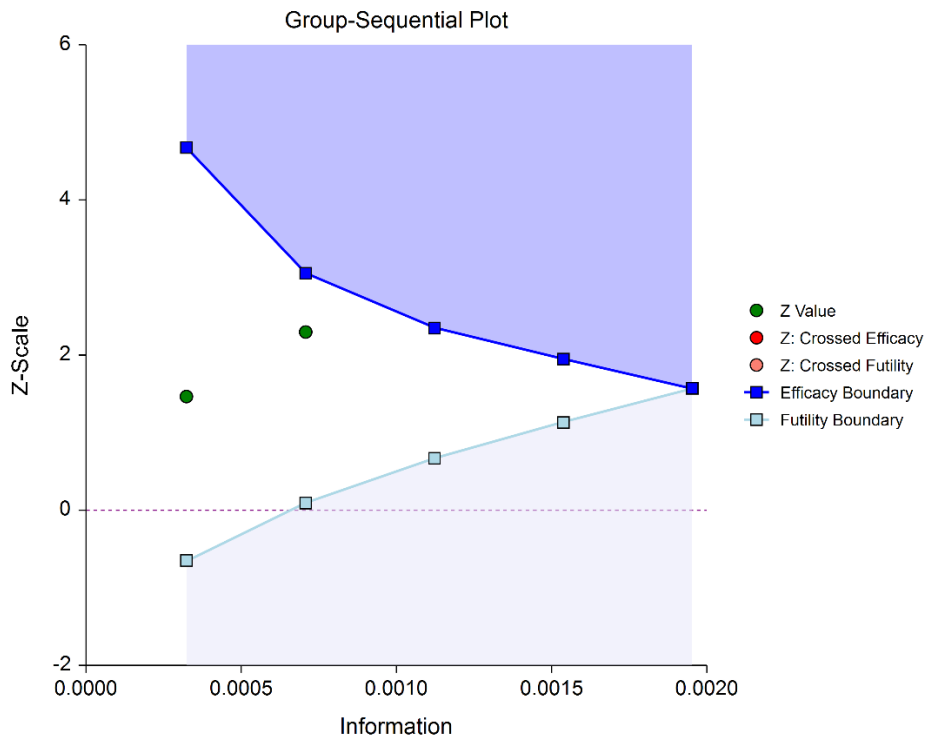
Once again, if no boundary is crossed, conditional power and stopping probabilities may be considered based on a choice of a supposed true difference.

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 *z*-statistic is computed from the sample hazard rates of the complete data from each group. The *z*-statistic is compared to the boundary and a decision of efficacy or futility is made.

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 hazard rate 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.
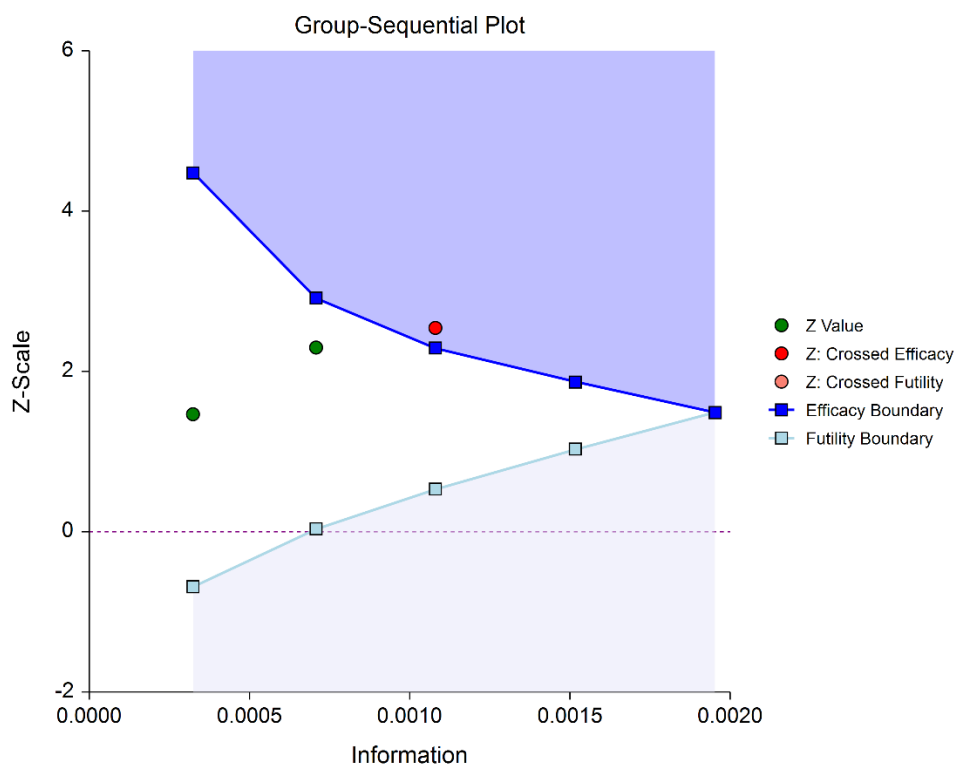
Adjusted confidence interval and estimate of the hazard rate difference

Due to the bias that is introduced in the group-sequential analysis process, the raw data confidence interval of the difference in hazard rates should not be used. An adjusted confidence interval should be used instead. See the Adjusted Confidence Interval topic of the Technical Details section for details. The mid-point of the adjusted confidence interval may be useful as a point estimate of the difference in hazard rates. Further, the confidence level at which the adjusted confidence interval limit is zero gives a rough adjusted p-value.

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.

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 hazard rates, the basic null hypothesis is that the hazard rates are equal,

$$H_0: h_1(T) = h_2(T)$$

with three common alternative hypotheses,

$$H_a: h_1(T) \neq h_2(T) \,,$$

$$H_a: h_1(T) < h_2(T) \,, \text{ or}$$

$$H_a: h_1(T) > h_2(T) \,,$$

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

In words, the null hypothesis is that the hazard rates of the two populations are equal at all times less than the maximum observed time and the alternative hypothesis is that the two hazard rates differ at some time less than the observed maximum time.

These hypotheses may be specified equivalently as

$$H_0: h_1(T) - h_2(T) = 0$$

versus

$$H_a: h_1(T) - h_2(T) \neq 0$$

$$H_a: h_1(T) - h_2(T) < 0$$

$$H_a: h_1(T) - h_2(T) > 0$$

A slightly different set of null and alternative hypotheses are used if the goal of the test is to determine whether $h_1(T)$ or $h_2(T)$ is greater than or less than the other by a given amount.

The null hypothesis then takes on the form

$$H_0: h_1(T) - h_2(T) = \text{Hypothesized Difference}$$

and the alternative hypotheses,

$$H_a: h_1(T) - h_2(T) \neq \text{Hypothesized Difference}$$

$$H_a: h_1(T) - h_2(T) < \text{Hypothesized Difference}$$

$$H_a: h_1(T) - h_2(T) > \text{Hypothesized Difference}$$

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

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 (MLE)

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

The general form of the test statistic is

$$z_k = \frac{\hat{h}_{1k} - \hat{h}_{2k}}{\sqrt{\hat{\sigma}_k^2(\hat{h}_{1k}) + \hat{\sigma}_k^2(\hat{h}_{2k})}}$$

with

$$\hat{h}_{ik} = \frac{\sum_{j=1}^{n_{ik}} c_{ijk}}{\sum_{j=1}^{n_{ik}} x_{ijk}}$$

$$\hat{\sigma}_k^2(\hat{h}_{ik}) = \frac{\sum_{j=1}^{n_{ik}} c_{ijk}}{\left(\sum_{j=1}^{n_{ik}} x_{ijk}\right)^2} = \frac{\hat{h}_{ik}^2}{\sum_{j=1}^{n_{ik}} c_{ijk}}$$

where

i = 1, 2 for the two groups

$\hat{h}_{ik}$ is the estimated group hazard rate at stage k

$\hat{\sigma}_k^2(\hat{h}_{ik})$ is the variance of the hazard rate estimator

c_{ijk} is an indicator of censoring

x_{ijk} is the elapsed time

Test Statistics (Logrank Type)

The general form of the Logrank-type test statistic is

$$z_k = \frac{\sum_{i=1}^{D} W(t_i)\left(d_{i1} - Y_{i1}\left(\frac{d_i}{Y_i}\right)\right)}{\sqrt{\sum_{i=1}^{D} W(t_i)^2 \frac{Y_{i1}}{Y_i}\left(1 - \frac{Y_{i1}}{Y_i}\right)\left(\frac{Y_i - d_i}{Y_i - 1}\right)d_i}}$$

where

D is the number of distinct event times

$W(t_i)$ is the weight function at time t_i

Y_{i1} is the number at risk in the Group 1 sample at time t_i

Y_i is the combined number at risk at time t_i

d_{i1} is the number of events in the Group 1 sample at time t_i

d_i is the combined number of events at time t_i

Details of the above formulas can be found in Klein and Moeschberger (1997), pages 191-202, and Andersen, Borgan, Gill, and Keiding (1992), pages 345-356.

Six different choices for the weight function, $W(T)$, with the flexible p and q for the Fleming-Harrington weight function, result in a variety of tests that are available in this procedure. The most commonly used test is the Logrank test, which has equal weighting. The other tests shift the heaviest weighting to the beginning or end of the trial. This may be appropriate in some studies, but the use of one of these other weighting schemes should be designated before the data have been seen. Because of the different weighting patterns, they will often give quite different results.

The following table describes each of these tests:

Test	**Weight**	**Comments**
Logrank	1	This is the most commonly used test. It places equal weights across all times. This test has optimum power when the hazard rates of the K populations are proportional to each other.
Gehan-Wilcoxon	Y_i	Places weight on hazards at the beginning of the study.
Tarone-Ware	$\sqrt{Y_i}$	Places weight on hazards at the beginning of the study.
Peto-Peto	$\tilde{S}(t_i)$	Places weight on hazards at the beginning of the study.
Modified Peto-Peto	$\tilde{S}(t_i)Y_i/(Y_i+1)$	Places weight on hazards at the beginning of the study.
Fleming-Harrington (1,0)	$\hat{S}(t_{i-1})$	Places weight on hazards at the beginning of the study.
Fleming-Harrington (0.5,0.5)	$\sqrt{\hat{S}(t_{i-1})\left(1-\hat{S}(t_{i-1})\right)}$	Places weight on hazards in the middle of the study.
Fleming-Harrington (1,1)	$\hat{S}(t_{i-1})\left(1-\hat{S}(t_{i-1})\right)$	Places weight on hazards in the middle of the study.
Fleming-Harrington (0,1)	$1-\hat{S}(t_{i-1})$	Places weight on hazards at the end of the study.
Fleming-Harrington (0.5,2)	$\sqrt{\hat{S}(t_{i-1})}\left(1-\hat{S}(t_{i-1})\right)^2$	Places weight on hazards at the end of the study.

This table uses the following definitions:

$$\hat{S}(t) = \prod_{t_i \le t}\left(1 - \frac{d_i}{Y_i}\right)$$

$$\tilde{S}(t) = \prod_{t_i \le t}\left(1 - \frac{d_i}{Y_i + 1}\right)$$

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, and the anticipated hazard rates with the corresponding estimate of the true difference in hazard rates.

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. The appropriate procedure in **PASS** (sample size software) can be used to make these planning phase sample size estimation calculations.

Information and Total Information

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

$$I_K^* = \frac{1}{\dfrac{\sigma_K^2(h_1, l_1, p_1)}{n_{1K}} + \dfrac{\sigma_K^2(h_2, l_2, p_2)}{n_{2K}}}$$

where

       $i = 1,2$ for the two groups

       $\sigma_K^2(h_i, l_i, p_i)$ is the variance of the hazard rate estimator

       h_i is the group hazard rate

       l_i is the group loss hazard rate

       p_i is the group patient entry parameter

and

$$\sigma_K^2(h_i, l_i, p_i) = h_i^2 \left(\frac{h_i}{h_i + l_i} + \frac{h_i p_i e^{-(h_i + l_i)T}\left(1 - e^{(h_i + l_i - p_i)T_0}\right)}{(1 - e^{-p_i T_0})(h_i + l_i)(h_i + l_i - p_i)} \right)^{-1}$$

       for $i = 1, 2,$

where

       T_0 is the accrual time

       T is the total time

If patient entry is uniform, the group variance is (Lachin and Foulkes, 1986):

$$\sigma_K^2(h_i, l_i, p_i) = h_i^2 \left(\frac{h_i}{h_i + l_i} \left[1 - \frac{e^{-(T - T_0)(h_i + l_i)} - e^{-T(h_i + l_i)}}{T_0(h_i + l_i)} \right] \right)^{-1}$$

The information at any data stage k may be calculated from the specified rates and the sample sizes, as

$$I_k = \frac{1}{\hat{\sigma}_k^2(\hat{h}_{1k}) + \hat{\sigma}_k^2(\hat{h}_{2k})}$$

with variance estimates as defined in the Test Statistic (MLE) section.

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

$$I_K = \frac{1}{\hat{\sigma}_K^2(\hat{h}_{1K}) + \hat{\sigma}_K^2(\hat{h}_{2K})} \neq I_K^* = \frac{1}{\frac{\sigma_K^2(h_1, l_1, p_1)}{n_{1K}} + \frac{\sigma_K^2(h_2, l_2, p_2)}{n_{2K}}}$$

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.

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.
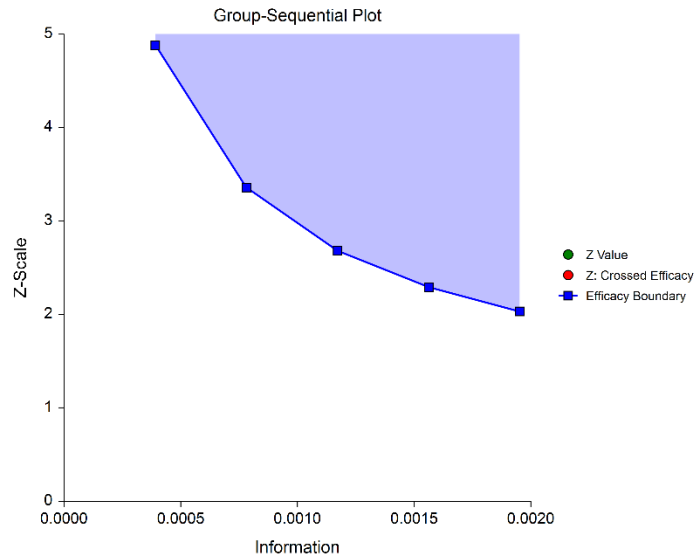
Efficacy Only (Two-Sided, Symmetric)

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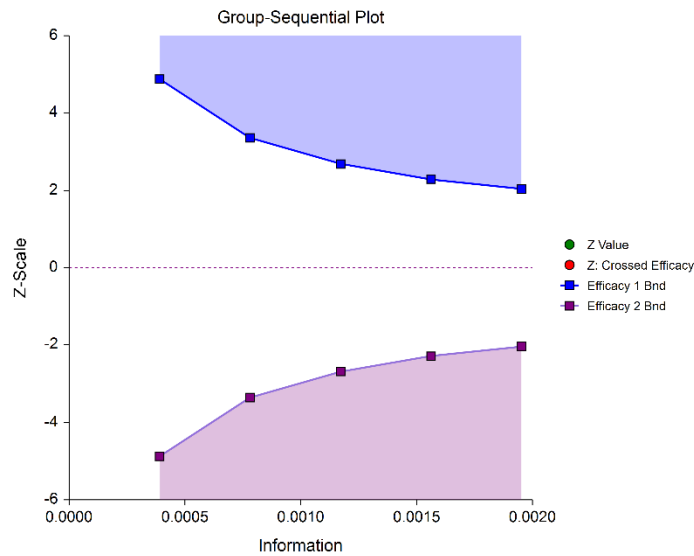
Efficacy 1 and Efficacy 2 / Harm (Two-Sided, Asymmetric)

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.
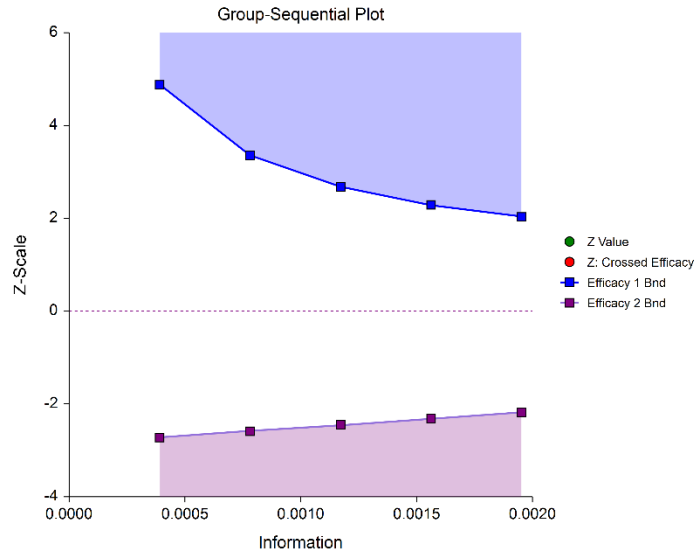
Efficacy and Binding Futility (One-Sided)

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.
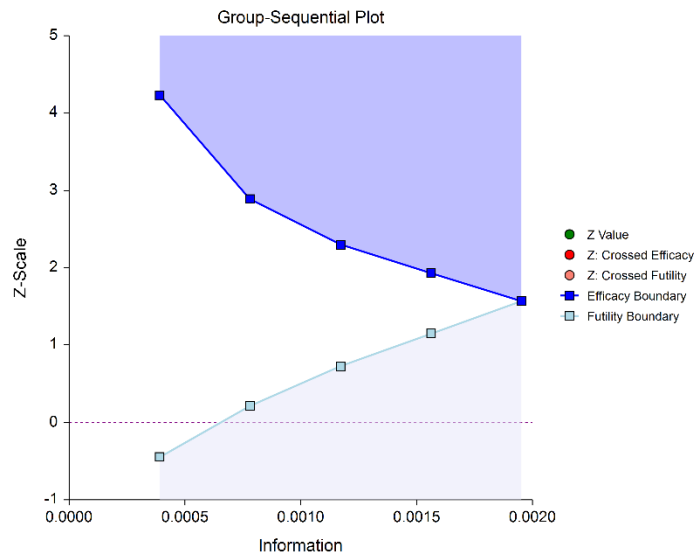
Efficacy and Non-Binding Futility (One-Sided)

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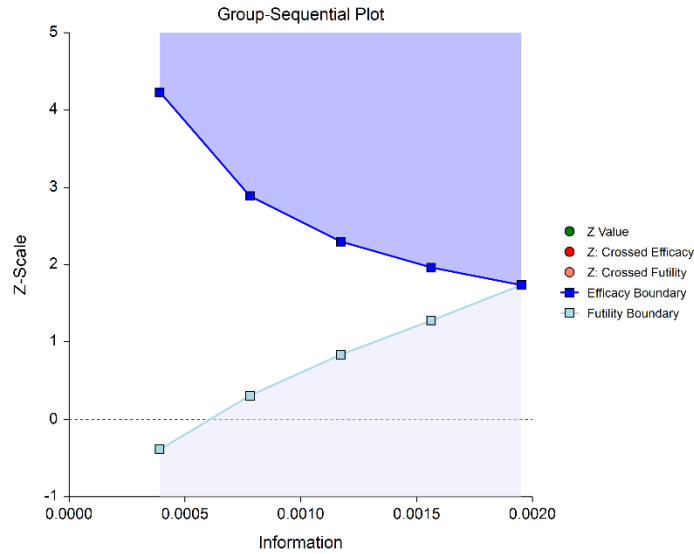
Efficacy and Binding Futility (Two-Sided, Symmetric)

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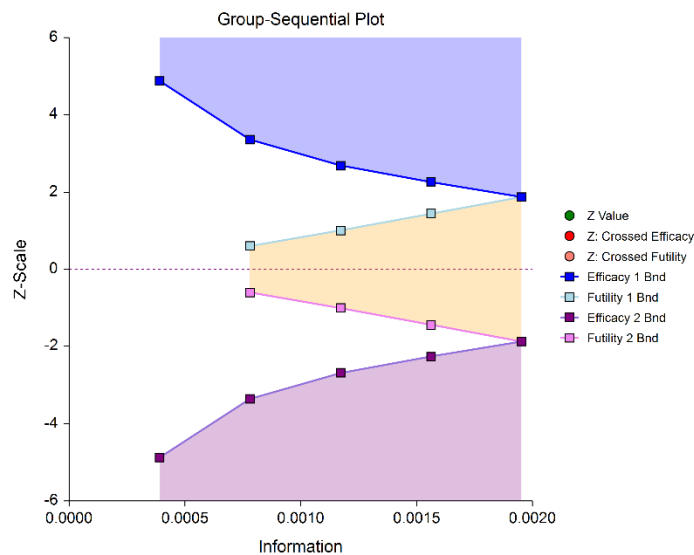
Efficacy and Non-Binding Futility (Two-Sided, Symmetric)

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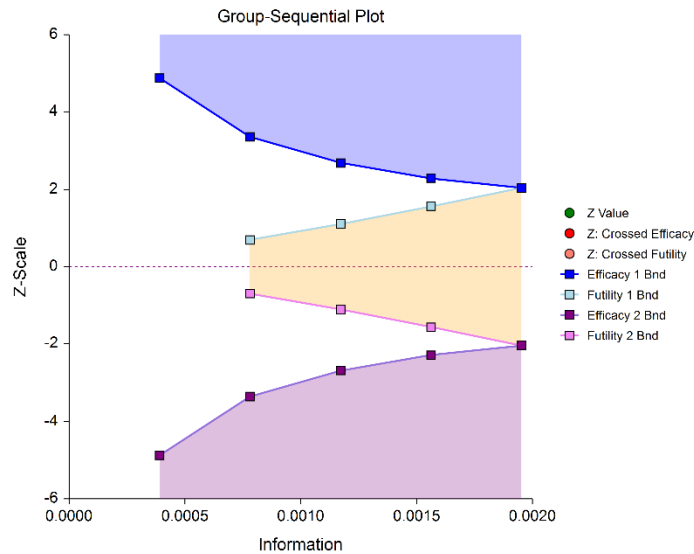
Efficacy 1, Efficacy 2 / Harm, and Binding Futility (Two-Sided, Asymmetric)

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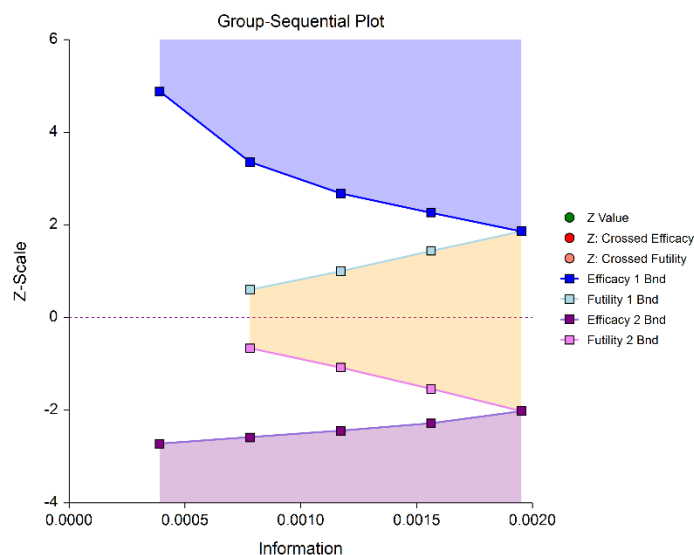
Efficacy 1, Efficacy 2 / Harm, and Non-Binding Futility (Two-Sided, Asymmetric)

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.
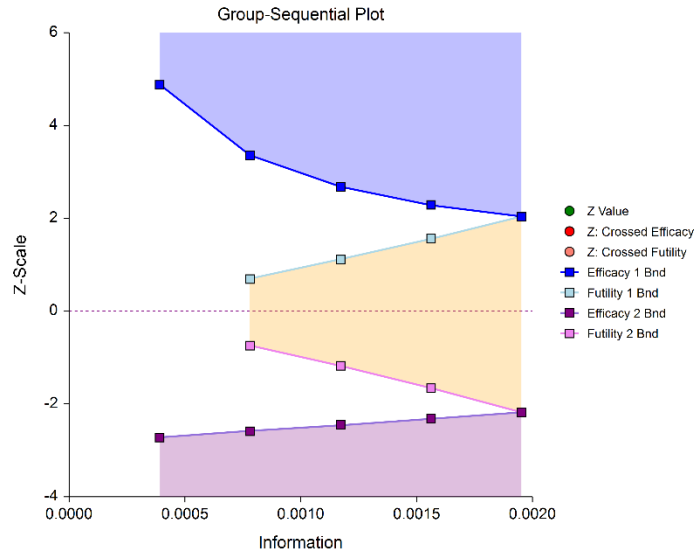
Futility Only (One-Sided)

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.
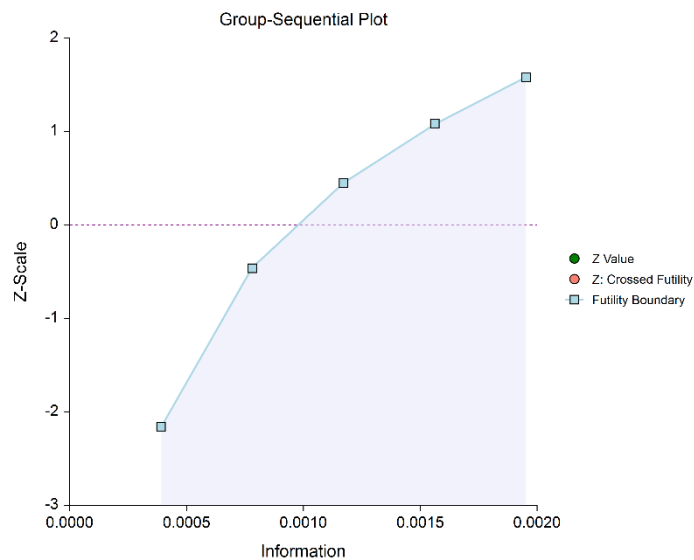
Futility Only (Two-Sided, Symmetric)

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.
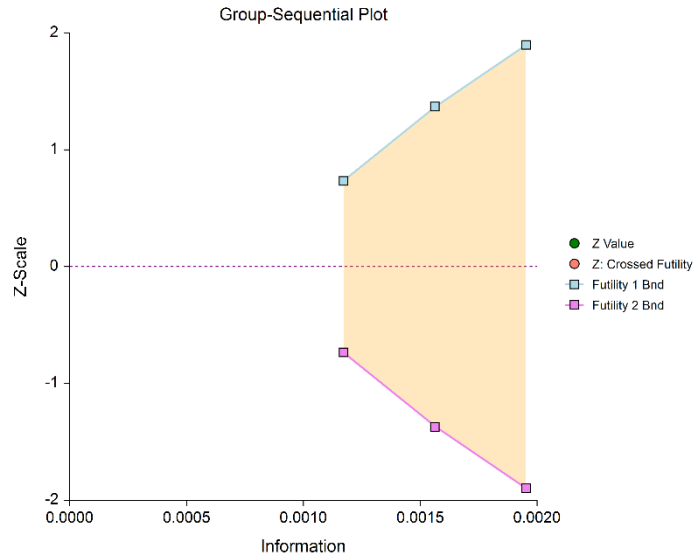
Futility Only (Two-Sided, Asymmetric)

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.
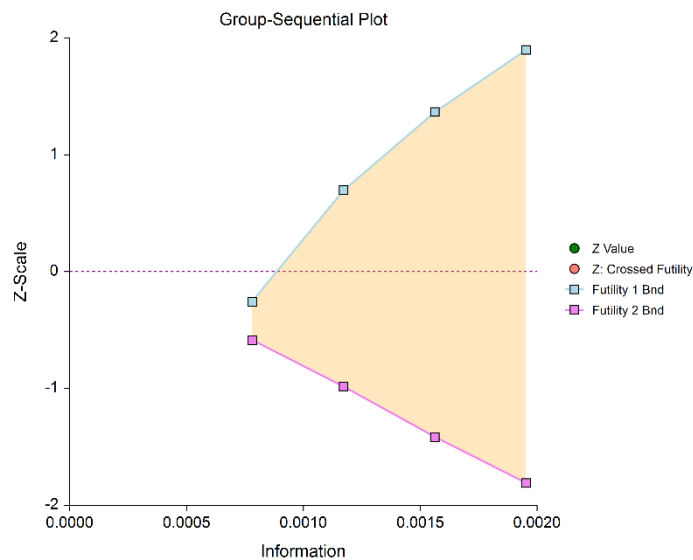
Boundary Calculations

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.

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.

- 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.

$$\alpha(0) = 0$$

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

$$\alpha(1) = \alpha$$
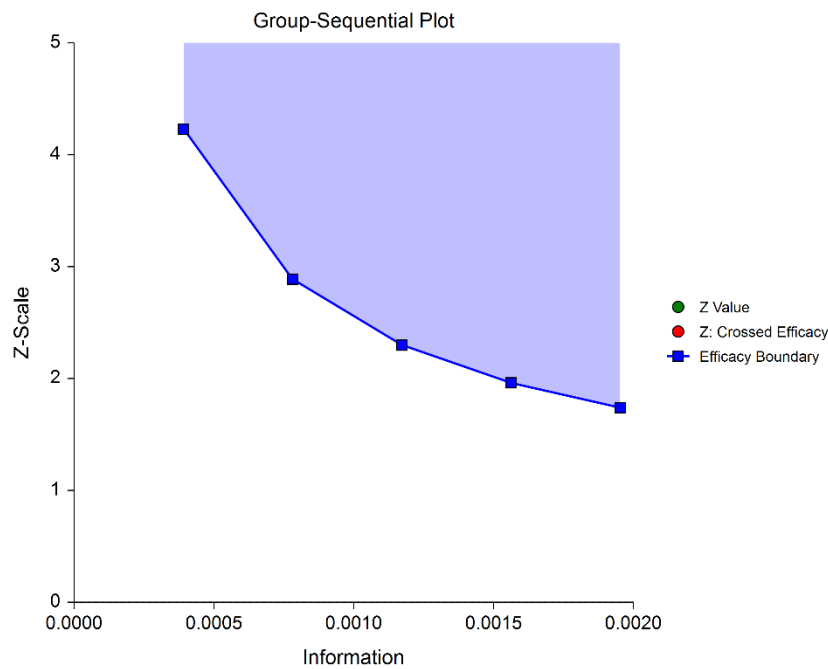
Pocock Analog

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.

$$\alpha(0) = 0$$

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

$$\alpha(1) = \alpha$$
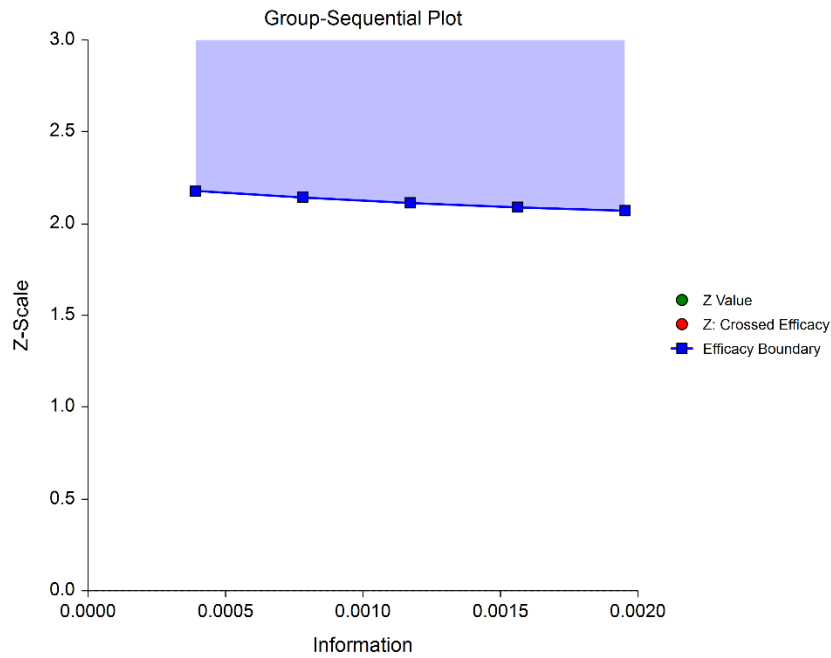
Power Family

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.

Group-Sequential Analysis for Two Hazard Rates

$\rho = 1$

$\rho = 2$

$\rho = 3$
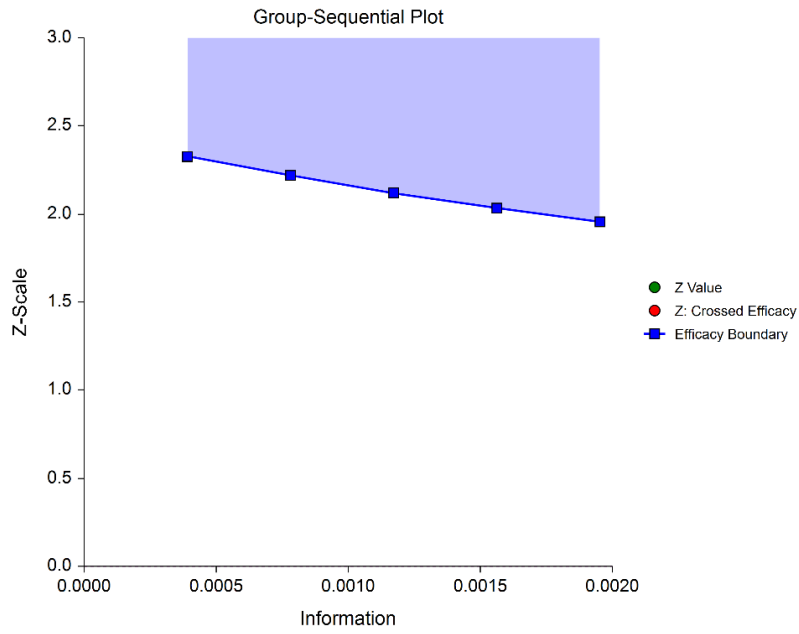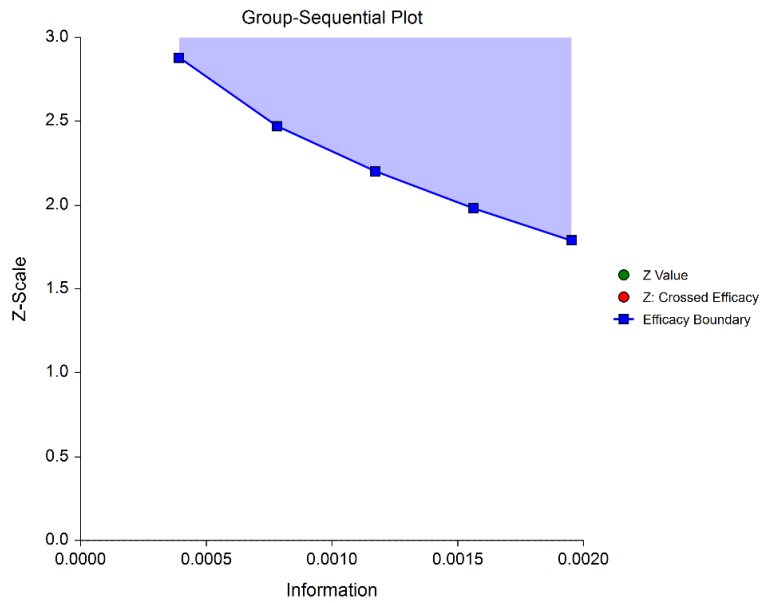
Group-Sequential Plot
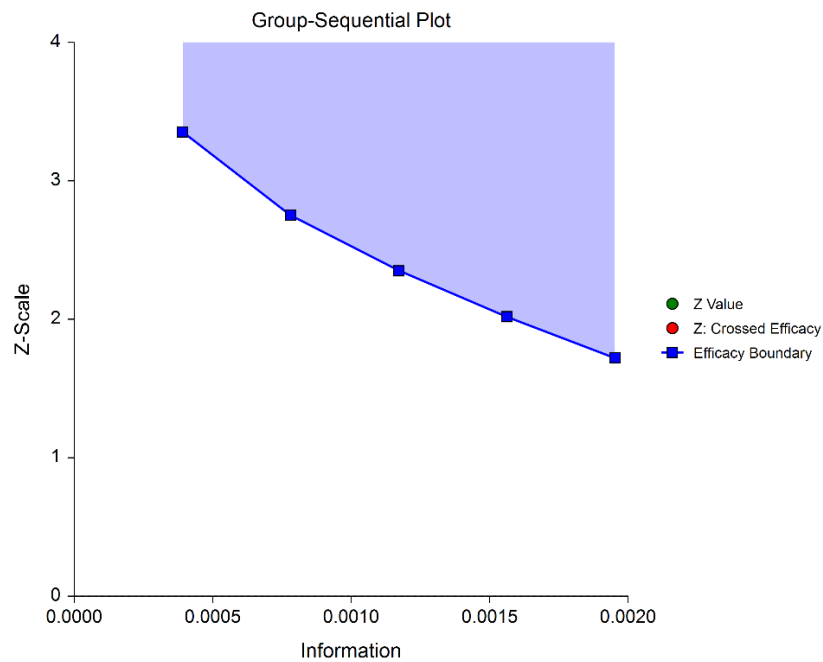
Hwang-Shih-DeCani (Gamma Family)

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

$$\alpha(0) = 0$$

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

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

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

Group-Sequential Analysis for Two Hazard Rates

$\gamma = -3$

$\gamma = -1$
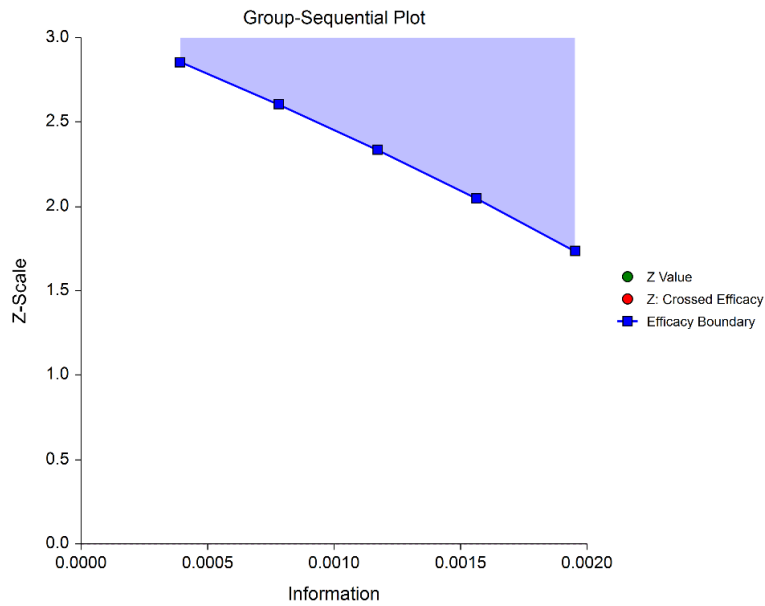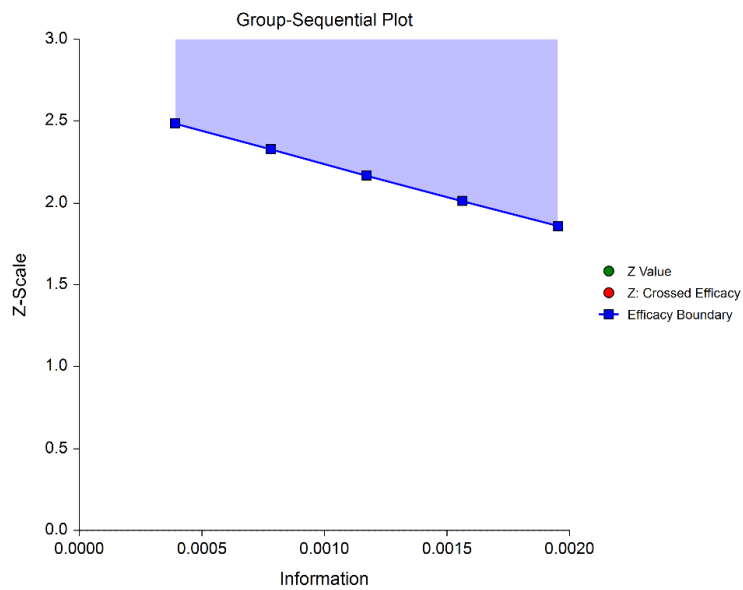
Group-Sequential Analysis for Two Hazard Rates

$\gamma = 1$
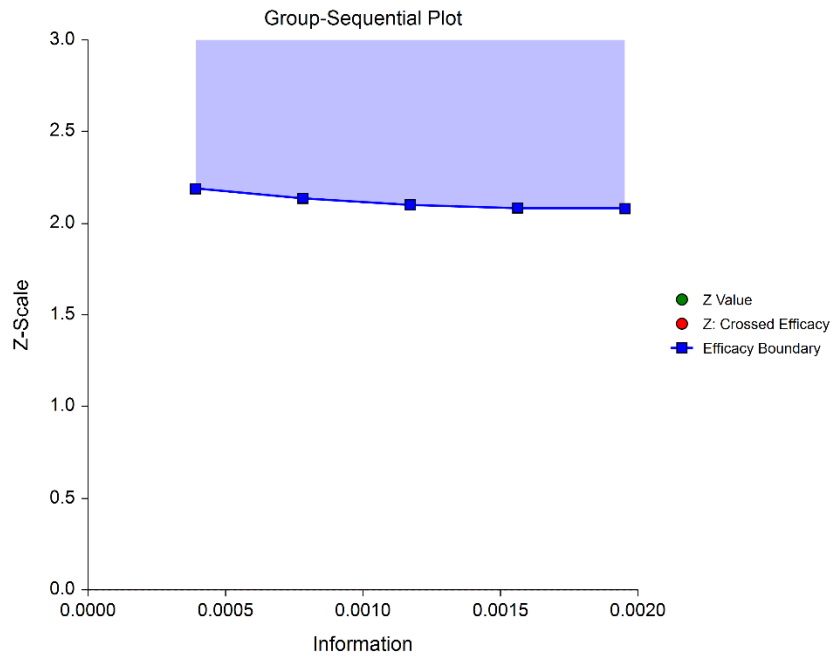
$\gamma = 3$
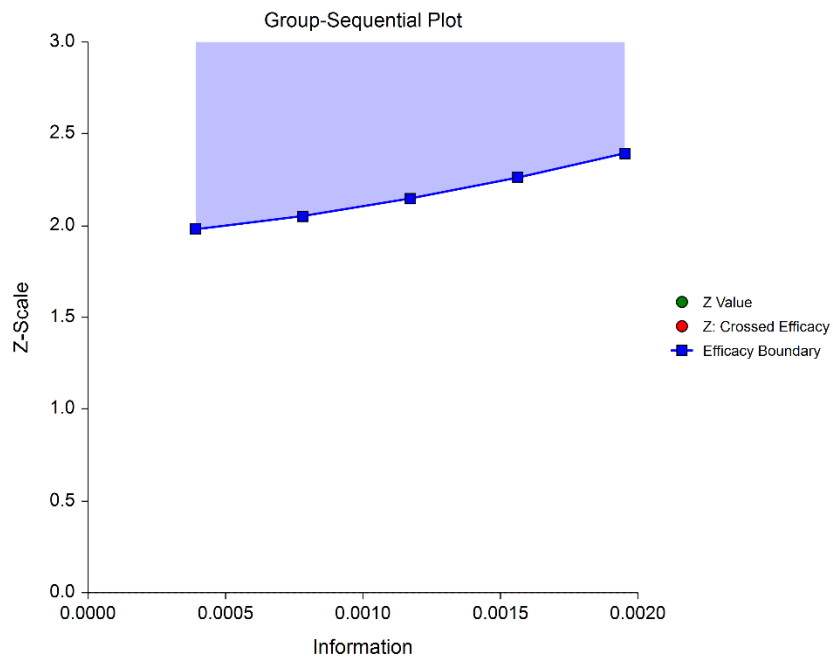
Adjusted Confidence Intervals

Except at the first stage, the raw (naïve) estimates of the confidence interval limits are inherently biased in the group-sequential analysis setting. The method given in Reboussin, DeMets, Kim, & Lan (1992) is used to calculate appropriately adjusted confidence limits and is based on Kim and DeMets (1987). This stage-wise ordering method is recommended by Jennison and Turnbull (2000) as "the only method available for use with unpredictable information sequences."

The methods used for the calculation of these confidence limits are based on the assumption that the current stage is the stopping stage of the study (typically from the crossing of a boundary).

As the methods in Reboussin *et al*. (1992) give only efficacy boundary results, the calculation adjustments are based only on the efficacy boundaries. The futility boundaries are not used in these calculations, except as they affect the efficacy boundaries.

Adjusted 'p-values'

As the raw (naïve) *p*-value is inherently biased in the group-sequential analysis setting, a sequential-test adjusted *p*-value should be used instead. Rather than giving a *p*-value in this procedure, a search is used to determine the adjusted confidence interval level at which an interval limit equals zero. As such, the methods used for this calculation are also based on the assumptions that the current stage is the stopping stage of the study. Again, only efficacy bounds are used.

Adjusted Hazard Rate Difference

The adjusted hazard rate difference is a rough estimate of the difference calculated simply as the midpoint of the adjusted confidence interval limits.

Conditional Power

From Jennison and Turnbull (2000) pages 205 to 208, the general upper one-sided conditional power at stage *k* for rejecting a null hypothesis about a parameter θ at the end of the study, given the observed test statistic, Z_k, is computed as

$$P_{uk}(\theta) = \Phi\left(\frac{Z_k\sqrt{I_k} - z_{1-\alpha}\sqrt{I_K} + \theta(I_K - I_k)}{\sqrt{I_K - I_k}}\right),$$

the general lower one-sided conditional power at stage *k* is computed as

$$P_{lk}(\theta) = \Phi\left(\frac{-Z_k\sqrt{I_k} - z_{1-\alpha}\sqrt{I_K} - \theta(I_K - I_k)}{\sqrt{I_K - I_k}}\right),$$

and the general two-sided conditional power at stage *k* is computed as

$$P_k(\theta) = \Phi\left(\frac{Z_k\sqrt{I_k} - z_{1-\alpha/2}\sqrt{I_K} + \theta(I_K - I_k)}{\sqrt{I_K - I_k}}\right) + \Phi\left(\frac{-Z_k\sqrt{I_k} - z_{1-\alpha/2}\sqrt{I_K} - \theta(I_K - I_k)}{\sqrt{I_K - I_k}}\right),$$

where

θ = the parameter being tested by the hypothesis

k = an interim stage at which the conditional power is computed (k = 1, ..., K – 1)

K = the stage at which the study is terminated, and the final test computed

Z_k = the test statistic calculated from the observed data that has been collected up to stage k

I_k = the information level at stage k

I_K = the information level at the end of the study

$z_{1-\alpha}$ = the standard normal value for the test with a type I error rate of α.

For the logrank test for two survival distributions with null hypothesis $H_0: HR = 1$, where HR is the hazard ratio, these components are computed in Chang (2008) page 71 as

$\theta = \log(HR) = \log(\lambda_2/\lambda_1)$ (the expected log(hazard ratio) under the alternative hypothesis)

$Z_k = S_k/\sqrt{\hat{I}_k}$ (the logrank test statistic computed from the observed data, see an alternate
     representation for Z_k in Jennison & Turnbull (2000) on page 216)

where

HR is the hazard ratio of treatment (group 2) to control (group 1)

λ_j is the hazard rate of group j

S_k is a logrank score statistic

$\hat{I}_k$ is the estimated information from the sample at stage k

E_k is the number of events at stage k

E is the total number of events

P_1 is the proportion of the subjects assigned to group 1, the control group

Computing conditional power requires you to set HR. This value can come from the values used during the planning of the study, from similar studies, or from estimates made from the data that has emerged.

Predictive Power

Predictive power (a Bayesian concept) is the result of averaging the conditional power over the posterior distribution of effect size. From Jennison and Turnbull (2000) pages 210 to 213, the general upper one-sided predictive power at stage *k* is given by

$$P_{uk} = \Phi\left(\frac{Z_k\sqrt{I_K} - z_{1-\alpha}\sqrt{I_k}}{\sqrt{I_K - I_k}}\right)$$

The general lower one-sided predictive power at stage *k* is given by

$$P_{lk} = \Phi\left(\frac{-Z_k\sqrt{I_K} - z_{1-\alpha}\sqrt{I_k}}{\sqrt{I_K - I_k}}\right)$$

The general two-sided predictive power at stage *k* is given by

$$P_k = \Phi\left(\frac{|Z_k|\sqrt{I_K} - z_{1-\alpha/2}\sqrt{I_k}}{\sqrt{I_K - I_k}}\right) + \Phi\left(\frac{-|Z_k|\sqrt{I_K} - z_{1-\alpha/2}\sqrt{I_k}}{\sqrt{I_K - I_k}}\right)$$

with all terms defined as in the equations for conditional power.

Using Simulation to obtain Future Boundary Crossing Probabilities

It can be useful to researchers to know the probability of crossing future group-sequential boundaries, given the data already obtained, and given specified assumed values for the hazard rates. The following steps are used to estimate these probabilities using simulation:

1. From the data obtained to the current stage, obtain a current estimate of the information.

2. Determine the target (cumulative) sample sizes and numbers of events for each future stage, including the final stage. Fractional sample sizes are rounded up to the next integer.

3. For each simulation, append simulated values to the current sample data to obtain a data set with the final stage sample sizes. Simulated values correspond to assumed hazard rates.

4. For each data set, determine which boundary or boundaries were crossed first (except in the case of non-binding futility boundaries). The proportion of simulations crossing each boundary (first) provides an estimate of the probability of crossing each boundary, given the specified assumed hazard rates.

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.

Data Structure

The data for this procedure is entered in four columns, with an additional column to identify the stage times. Two columns give the individual start and end times. A censor column identifies whether the end time is due to an event or to censoring. Another column identifies the group of the patient, individual, or experimental unit.

These five columns are assigned on the Variables tab of the procedure. Groups 1 and 2 are also assigned on the Variables tab.

The current stage is also identified directly on the Variables tab.

	StartDate	EndDate	Censor	Group	f_x Start	f_x End	StageTimes	C8	C9	C
1	1/2/2017	2/12/2017	0	Trt	0.003785	0.116455	1			
2	1/4/2017	5/26/2018	0	Trt	0.009687	1.398635	2			
3	1/10/2017	3/15/2017	0	Cntrl	0.025288	0.20252	3			
4	1/11/2017	2/17/2017	0	Cntrl	0.029976	0.1294				
5	1/18/2017	4/14/2017	0	Trt	0.046929	0.282283				
6	1/18/2017	7/23/2018	0	Trt	0.048199	1.557464				
7	1/20/2017	3/9/2017	0	Cntrl	0.054407	0.185398				
8	1/21/2017	6/1/2017	0	Trt	0.05543	0.416412				
9	1/21/2017	3/3/2017	0	Trt	0.057502	0.167194				
10	1/22/2017	1/26/2017	0	Cntrl	0.057688	0.069141				
11	1/22/2017	12/20/2017	0	Cntrl	0.059322	0.969148				
12	1/23/2017	2/2/2018	0	Trt	0.062278	1.088534				
13	1/24/2017	2/4/2018	0	Cntrl	0.064694	1.095082				
14	1/25/2017	5/17/2018	0	Cntrl	0.067065	1.373355				
15	1/27/2017	5/1/2018	0	Trt	0.073059	1.329166				
16	1/28/2017	2/28/2018	0	Trt	0.074532	1.159621				
17	1/28/2017	2/7/2017	0	Trt	0.075274	0.103676				
18	1/29/2017	12/23/2018	0	Trt	0.076946	1.977546				
19	2/3/2017	6/12/2017	0	Cntrl	0.091545	0.44497				
20	2/5/2017	1/12/2018	0	Trt	0.096764	1.030805				

For this data set, start dates and end dates are converted to start and end times through a transformation.

The transformation for converting to year times, is

$$Time = \frac{StartDate - Time0DateValue}{365}$$

For this dataset, the study began on January 1, 2017. The corresponding day value for January 1, 2017 is 42736. This value is found by entering the date in an empty column, changing the Data Type to Date & Time, and then changing the Data Type back to General. Thus, the formula is

$$Start\ Time = \frac{StartDate - 42736}{365}$$

and

$$End\ Time = \frac{EndDate - 42736}{365}$$

The units for the Start, End, and StageTimes columns are years. Any units may be used, but the units must be consistent throughout the use of the procedure. For example, if years are the units, then hazard rates must be yearly hazard rates. Conversions may be accomplished through the Survival Parameter Conversion Tool, which is available through the Tools menu.

If an individual has not had an event at the time of the stage analysis, but is still in the study, the end time should be left blank, as in

	StartDate	EndDate	Censor	Group	fx Start	fx End	StageTimes	C8	C9	C
496	6/3/2019		0	Trt	2.421438					
497	6/8/2019		0	Cntrl	2.435071					
498	6/11/2019	8/3/2019	0	Cntrl	2.442833	2.588857				
499	6/12/2019	12/22/2019	0	Trt	2.444262	2.973864				
500	6/17/2019		0	Trt	2.460232					
501	6/19/2019	10/31/2019	0	Cntrl	2.463922	2.831271				
502	6/20/2019	10/17/2019	0	Trt	2.46738	2.793332				
503	6/21/2019	7/26/2019	0	Cntrl	2.470428	2.567093				
504	6/21/2019	7/29/2019	0	Cntrl	2.470458	2.573955				
505	6/30/2019		0	Cntrl	2.49507					
506	7/2/2019	11/25/2019	0	Cntrl	2.498895	2.900524				
507	7/2/2019	8/9/2019	1	Cntrl	2.498976	2.603059				
508	7/2/2019	9/23/2019	0	Trt	2.499687	2.727671				
509	7/2/2019	7/16/2019	0	Trt	2.500937	2.53834				
510	7/3/2019	9/16/2019	1	Trt	2.502371	2.709578				
511	7/4/2019	7/30/2019	0	Trt	2.506072	2.576769				
512	7/10/2019		0	Trt	2.522188					
513	7/11/2019		0	Trt	2.523956					
514	7/11/2019	12/21/2019	0	Trt	2.525425	2.972529				
515	7/12/2019		0	Trt	2.527532					
516	7/12/2019		0	Cntrl	2.527868					
517	7/15/2019		0	Trt	2.536789					
518	7/16/2019	8/18/2019	0	Cntrl	2.537412	2.629215				
519	7/24/2019	8/22/2019	0	Trt	2.561504	2.640745				
520	7/25/2019	12/13/2019	0	Cntrl	2.561729	2.948912				
521	7/25/2019	8/13/2019	0	Cntrl	2.563042	2.61626				
522	7/26/2019	8/3/2019	0	Cntrl	2.56695	2.587287				
523	8/2/2019		0	Trt	2.586078					
524	8/4/2019		0	Trt	2.590614					

Example 1 – Group-Sequential Analysis

A colorectal cancer study is conducted to determine whether a new treatment following tumor excision will result in a longer time before tumor recurrence (lesser hazard rate). The new treatment is compared to the current standard treatment. The response for each patient is time, in years, before recurrence. A one-sided test with alpha equal to 0.025 is used. The MLE-based Z-Test for comparing two hazard rates will be used.

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

$$H_0: h_1 - h_2 = 0 \quad (H_0: h_{New} = h_{Std})$$

versus

$$H_a: h_1 - h_2 < 0 \quad (H_a: h_{New} < h_{Std})$$

The design calls for five stages of one year each, if the final stage is reached. The current stage is the 3rd stage. In the design phase, a needed power of 0.90 called for 505 patients per group if the final stage is reached, based on assumed hazard rates of 1.40 and 1.75 for the new and standard treatments, respectively. Both efficacy and non-binding futility boundaries are implemented. 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. Accrual is intended to be steady over the 5-year period of the study. Loss hazard rates of 0.03 for both groups are anticipated.

The boundary plot for stage 2 appeared as
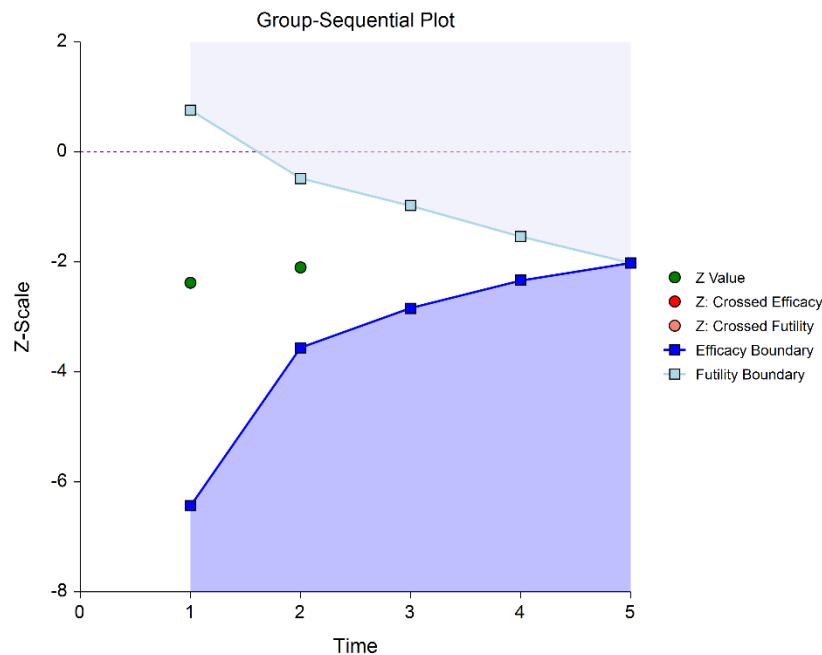
resulting in continuance of the study to stage 3.

Setup

To run this example, complete the following steps:

1   Open the GS Survival example dataset

- From the File menu of the NCSS Data window, select **Open Example Data**.
- Select **GS Survival** and click **OK**.

2   Specify the Group-Sequential Analysis for Two Hazard Rates procedure options

- Find and open the **Group-Sequential Analysis for Two Hazard Rates** procedure using the menus or the Procedure Navigator.
- The settings 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.

Variables Tab

Start Time Variable ....**Start**
End Time Variable ....**End**
Censor Indicator Variable ....**Censor**
  Censored **1**
  Not Censored...**0**
Group Variable...**Group**
  Group 1 Value...**Trt**
  Group 2 Value...**Cntrl**
Stage Times Variable ....**StageTimes**
Current Stage (k) **3**
Test Type..**MLE**
Maximum Number of Stages (K).....................**5**
Time Proportion at each Stage ....**Equally incremented**
N1 ....**505**
N2 ....**505**
h1..**1.4**
h2..**1.75**
T0 (Accrual or Recruitment Time)...................**5**
Accrual Parameter Entry.................................**Enter Accrual Parameter Directly**
Accrual Parameter **0**
T (Total Time) **5**
L1 (Loss Hazard Rate of Group 1)..................**0.03**
L2 (Loss Hazard Rate of Group 2)..................**0.03**
Future Stage Time Adjustment ....**Keep original design time proportions exactly**
Boundaries Used ....**One-sided Efficacy with Futility**
Hypothesis Direction.......................................**Ha: h1 - h2 < 0**
Boundary Specification ....**Spending Function Calculation**
Alpha **0.025**
Alpha Spending Function................................**O'Brien-Fleming Analog**
Skipped Efficacy Stages ....**<Empty>**

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

Reports Tab
_____

Reports for Current Stage
All Reports ....**Checked**
Confidence Level..**95**
Number of Simulations **2000**
All Three Boxes under Conditional ....**Checked**
and Predictive Power and Boundary
Crossing Probability Parameters
h1...**1.6**
h2...**1.6**

Planning Stage (Stage 0) Reports
Summary using Z Scale..................................**Checked**
Plots..**Checked**

Plots Tab
_____

Z-Statistic vs Information ....**Checked**
Z-Statistic vs Time ....**Checked**

Note: You can add the stage number to the plot by checking the **small box on the Plot Format button** (under Z-Statistic vs Information and Z-Statistic vs Time). When the procedure is run, select the **Stage Notes** tab and press the **Add Stage Number** button.

3   Run the procedure

- Click the **Run** button to perform the calculations and generate the output.

Run Summary Report

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

Run Summary Report
_____

Item	Value
Maximum Number of Stages (Design):	5
Current Stage:	3
Alternative Hypothesis:	h1 - h2 < 0 (one-sided)
Alpha Spending Function:	O'Brien-Fleming Analog
Beta Spending Function:	Hwang-Shih-DeCani (γ = 1.5)
Futility Boundaries:	Non-Binding
Alpha:	0.0250
Maximum Information:	86.5248
Z Statistic:	MLE

Z-Values and Boundaries at Stage 3

This section gives the Z-test values and boundaries, numerically. These values are reflected in the group-sequential boundary plot. The Decision column indicates whether a boundary was crossed at each stage.

Z-Values and Boundaries at Stage 3

Maximum Information: 86.5248
Alternative Hypothesis: h1 - h2 < 0 (one-sided)
Futility Boundaries: Non-Binding
Z Statistic: MLE

| | | Boundaries | | | Information | |
| | Z-Test | | | | | |
Stage	Value	Efficacy	Futility	Time	Proportion	Decision
1	-2.3797	-6.4316	0.7565	1.0	0.1173	Continue
2	-2.1001	-3.5628	-0.4866	2.0	0.3590	Continue
3	-3.3687	-2.7086	-1.1338	3.0	0.5871	Crossed Efficacy
4		-2.3412	-1.5201	4.0	0.7707	
5		-2.0218	-2.0218	5.0	1.0000	

Group-Sequential Boundary Plot at Stage 3

This plot shows the Z-test values and boundaries. The efficacy boundary is crossed at Stage 3.

Group-Sequential Boundary Plot at Stage 3
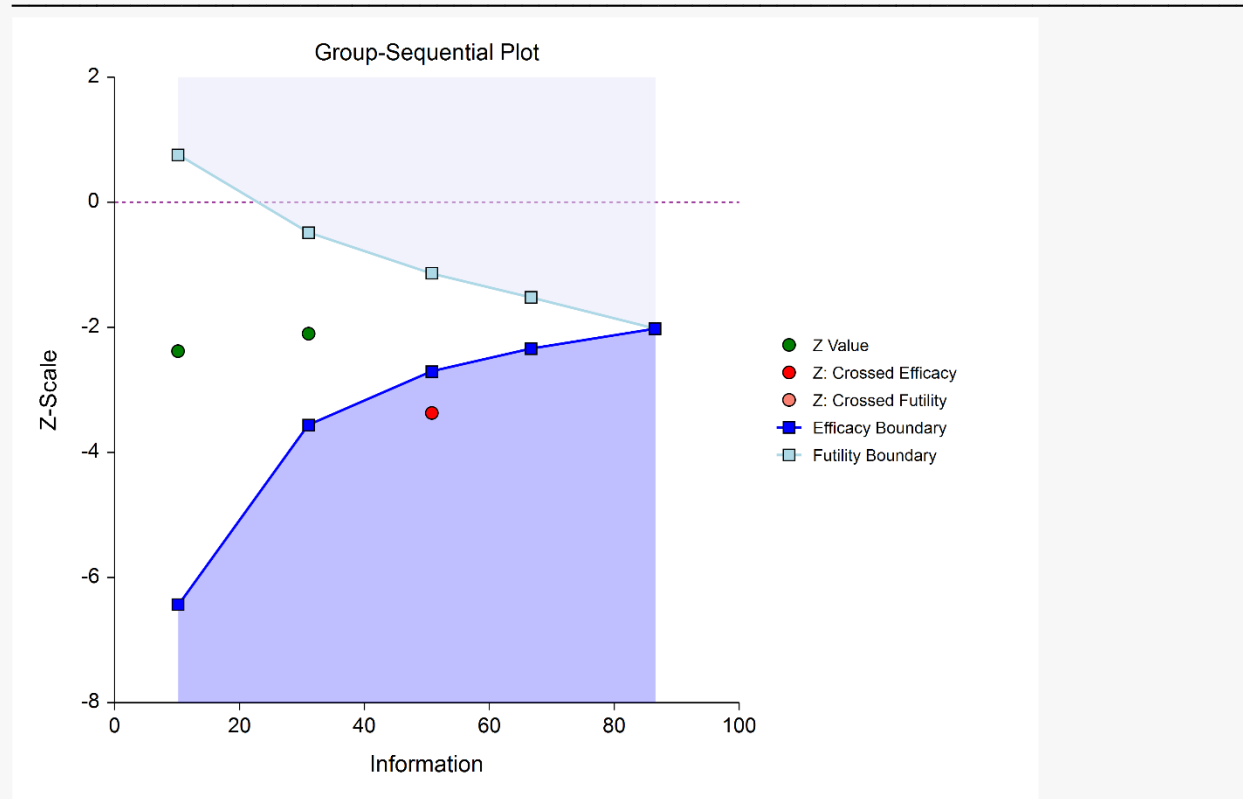
Group-Sequential Plot
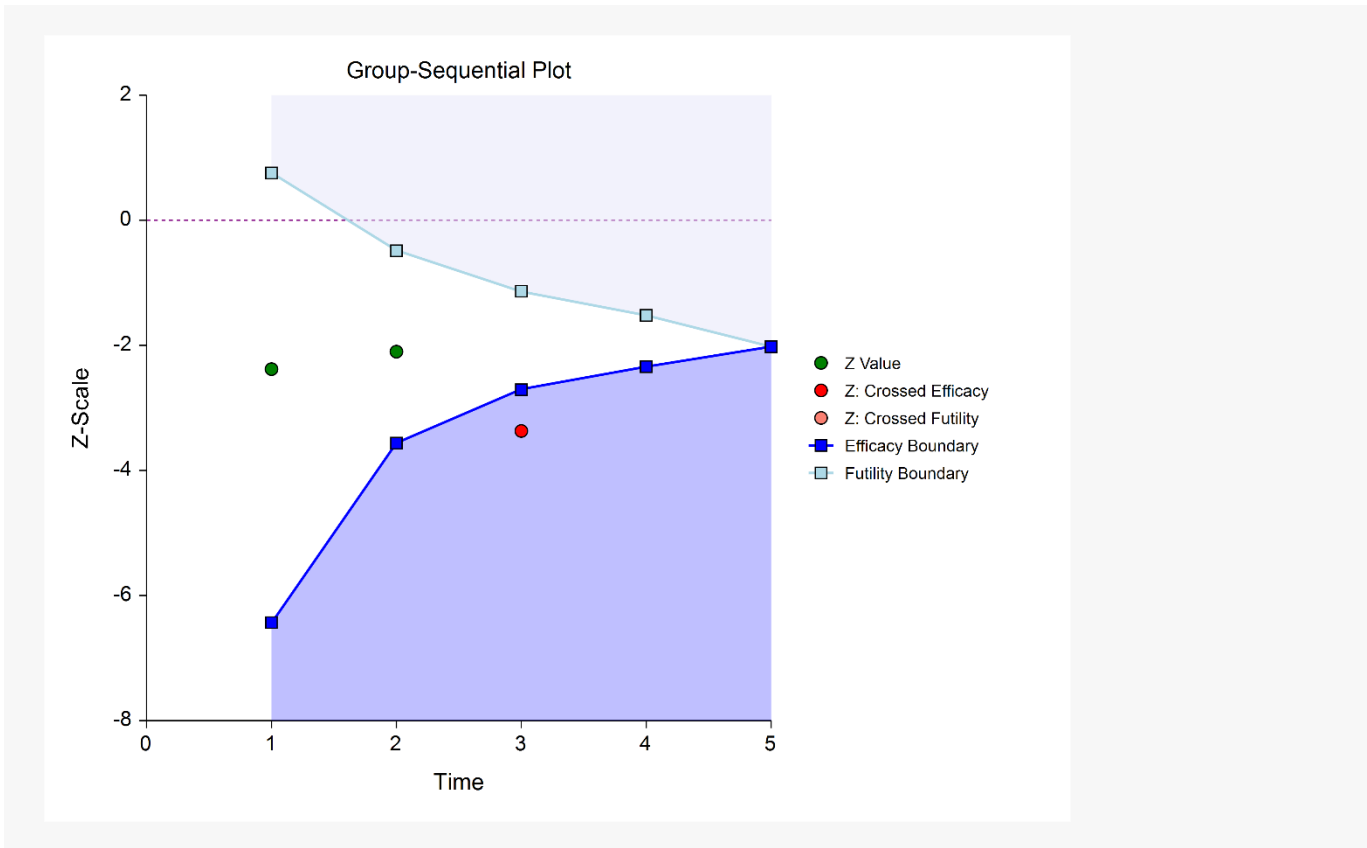
P-Values and Boundaries at Stage 3

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

P-Values and Boundaries at Stage 3

Maximum Information: 86.5248
Alternative Hypothesis: h1 - h2 < 0 (one-sided)
Futility Boundaries: Non-Binding
Z Statistic: MLE

| | | Boundaries | | | Information | |
Stage	Z-Test P-Value	Efficacy	Futility	Time	Proportion	Decision
1	0.00866	0.00000	0.77534	1.0	0.1173	Continue
2	0.01786	0.00018	0.31325	2.0	0.3590	Continue
3	0.00038	0.00338	0.12844	3.0	0.5871	Crossed Efficacy
4		0.00961	0.06424	4.0	0.7707	
5		0.02160	0.02160	5.0	1.0000	

P-values and P-value boundaries are one-sided values.

Current Stage Results Adjusted for Sequential Analysis (Stage 3)

This section gives appropriate adjustments to the raw results to reflect the group-sequential nature of the analysis. Additional explanation is given in early sections of this chapter: Adjusted Confidence Intervals, Adjusted p-values, and Adjusted Hazard Rate Difference.

Current Stage Results Adjusted for Sequential Analysis (Stage 3)

		Group-Sequential Adjusted Difference			Conf. Level
	Actual	95.0% C.I. of Diff.			Where Upper
	Difference			C.I.	Boundary
Stage	h1 - h2	Lower	Upper	Midpoint	Is Zero
3	-0.47265	-0.97316	-0.25003	-0.61279	99.898%

The adjustments are based only on the efficacy boundaries. The futility boundaries are not used in these calculations.

Descriptive Statistics up to Stage 3

This section gives the raw sample statistics for the cumulative data at each stage.

Descriptive Statistics up to Stage 3

| | Sample Size | | Event Count | | Sample Hazard Rates | | Difference | |
Stage	N1	N2	E1	E2	h1	h2	Difference	SE(Diff)
1	116	90	48	46	1.09335	1.84031	-0.74696	0.31389
2	219	184	145	122	1.24368	1.62048	-0.37680	0.17942
3	314	290	243	228	1.25946	1.73212	-0.47265	0.14031

Information Report at Stage 3

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

Information Report at Stage 3

Maximum Information: 86.5248
Alternative Hypothesis: h1 - h2 < 0 (one-sided)
Alpha: 0.0250

Stage	Target Information Proportion	Achieved Information Proportion	Target Information	Achieved Information	N1	N2	h1	h2
1	0.1153	0.1173	9.9780	10.1492	116.00	90.00	1.09335	1.84031
2	0.3211	0.3590	27.7831	31.0642	219.00	184.00	1.24368	1.62048
3	0.5448	0.5871	47.1361	50.7958	314.00	290.00	1.25946	1.73212
4	0.7720	*0.7707	66.7992	*66.6884	*371.33	*371.33	*1.25946	*1.73212
5	1.0000	*1.0000	86.5248	*86.5248	*464.16	*464.16	*1.25946	*1.73212

* Projected value

Alpha Spending at Stage 3

This section shows how alpha was spent (or is anticipated to be spent) across the stages.

Alpha Spending at Stage 3

Target Final Stage Alpha: 0.0250
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	0.1173	0.0000	0.0000	0.000000	0.0	0.0
2	0.3590	0.0002	0.0002	0.000183	0.7	0.7
3	0.5871	0.0033	0.0034	0.003379	13.0	13.8
4 *	0.7707	0.0072	0.0107	0.009611	28.9	42.7
5 *	1.0000	0.0143	0.0250	0.021597	57.3	100.0

* Projected

Beta Spending for Futility at Stage 3

This section shows how beta was spent (or is anticipated to be spent) across the stages.

Beta Spending for Futility at Stage 3

Target Cumulative Beta at Final Stage: 0.1000
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	0.1173	0.0208	0.0208	0.775340	20.8	20.8
2	0.3590	0.0328	0.0536	0.313254	32.8	53.6
3	0.5871	0.0218	0.0754	0.128444	21.8	75.4
4 *	0.7707	0.0128	0.0882	0.064237	12.8	88.2
5 *	1.0000	0.0118	0.1000	0.021597	11.8	100.0

* Projected

Conditional and Predictive Power Report at Stage 3

This section gives the conditional power for the various differences assumed. It also gives the predictive power.

Conditional and Predictive Power Report at Stage 3

Maximum Information: 86.5248
Information this Stage: 50.7958
Proportion of Maximum Information: 0.5871
Predictive Power: 0.9982

δ Name	δ Value	Conditional Power
Design	-0.3500	0.9989
Data	-0.4727	0.9999
δ1	0.0000	0.8331

The conditional power values are calculated based on Jennison and Turnbull (2000), pages 205 to 208. They do not account for future interim stages, nor futility boundaries, if applicable.

Boundary Probabilities for δ = -0.35

Using simulation based on the specified hazard rates, this section gives the estimated probabilities of crossing each of the future boundaries. Values given here will vary for each simulation.

Boundary Probabilities for δ = -0.35

Number of Simulations:	2000	
User-Entered Random Seed:	5689325	
Futility Boundaries:	Non-Binding	
After Efficacy Boundary Crossing:	Hold Out	
After Non-Binding Futility Boundary Crossing:	Leave In	
Alternative Hypothesis:	h1 - h2 < 0 (one-sided)	
Z Statistic:	MLE	
h1:	1.4	
h2:	1.75	
δ:	-0.35	

Stage	N1	N2	Z-Test Value	Efficacy Boundary	Efficacy Probability	Futility Boundary	Futility Probability
1	116.00	90.00	-2.3797	-6.4316		0.7565	
2	219.00	184.00	-2.1001	-3.5628		-0.4866	
3	314.00	290.00	-3.3687	-2.7086		-1.1338	
4	*371.33	*371.33		-2.3412	0.9955	-1.5201	0.0000
5	*464.16	*464.16		-2.0218	0.0035	-2.0218	0.0010

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

Event Summary for δ = -0.35

Using simulation based on the specified hazard rates, this section gives the estimated number of events at each future stage. Values given here will vary for each simulation.

Event Summary for δ = -0.35

Number of Simulations:	2000
User-Entered Random Seed:	5689325
h1:	1.4
h2:	1.75
δ:	-0.35

Stage	N1	N2	Z-Test Value	Average Cumulative Number of Events E1	Average Cumulative Number of Events E2	Time
1	116	90	-2.3797	48	46	1.00
2	219	184	-2.1001	145	122	2.00
3	314	290	-3.3687	243	228	3.00
4	*371.33	*371.33		**325.95	**319.41	*4.00
5	*464.16	*464.16		**401.89	**406.47	*5.00

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

Boundary Probabilities for δ = -0.4726526

Boundary Probabilities for δ = -0.4726526

───

Number of Simulations: 2000
User-Entered Random Seed: 5689325
Futility Boundaries: Non-Binding
After Efficacy Boundary Crossing: Hold Out
After Non-Binding Futility Boundary Crossing: Leave In
Alternative Hypothesis: h1 - h2 < 0 (one-sided)
Z Statistic: MLE
h1:   1.259464
h2:   1.732116
δ: -0.4726526

───

				Efficacy		Futility	
Stage	N1	N2	Z-Test Value	Boundary	Probability	Boundary	Probability
1	116.00	90.00	-2.3797	-6.4316		0.7565	
2	219.00	184.00	-2.1001	-3.5628		-0.4866	
3	314.00	290.00	-3.3687	-2.7086		-1.1338	
4	*371.33	*371.33		-2.3412	0.9990	-1.5201	0.0000
5	*464.16	*464.16		-2.0218	0.0010	-2.0218	0.0000

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

Event Summary for δ = -0.4726526

Event Summary for δ = -0.4726526

───

Number of Simulations: 2000
User-Entered Random Seed: 5689325
h1:   1.259464
h2:   1.732116
δ: -0.4726526

───

				Average Cumulative Number of Events		
Stage	N1	N2	Z-Test Value	E1	E2	Time
1	116	90	-2.3797	48	46	1.00
2	219	184	-2.1001	145	122	2.00
3	314	290	-3.3687	243	228	3.00
4	*371.33	*371.33		**321.42	**319.00	*4.00
5	*464.16	*464.16		**396.38	**406.02	*5.00

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

Boundary Probabilities for δ = 0

Boundary Probabilities for δ = 0

Number of Simulations:	2000	
User-Entered Random Seed:	5689325	
Futility Boundaries:	Non-Binding	
After Efficacy Boundary Crossing:	Hold Out	
After Non-Binding Futility Boundary Crossing:	Leave In	
Alternative Hypothesis:	h1 - h2 < 0 (one-sided)	
Z Statistic:	MLE	
h1:	1.6	
h2:	1.6	
δ:	0	

			Z-Test	Efficacy		Futility	
Stage	N1	N2	Value	Boundary	Probability	Boundary	Probability
1	116.00	90.00	-2.3797	-6.4316		0.7565	
2	219.00	184.00	-2.1001	-3.5628		-0.4866	
3	314.00	290.00	-3.3687	-2.7086		-1.1338	
4	*371.33	*371.33		-2.3412	0.8970	-1.5201	0.0010
5	*464.16	*464.16		-2.0218	0.0470	-2.0218	0.1370

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

Event Summary for δ = 0

Event Summary for δ = 0

Number of Simulations:	2000	
User-Entered Random Seed:	5689325	
h1:	1.6	
h2:	1.6	
δ:	0	

			Z-Test	Average Cumulative Number of Events		Time
Stage	N1	N2	Value	E1	E2	Time
1	116	90	-2.3797	48	46	1.00
2	219	184	-2.1001	145	122	2.00
3	314	290	-3.3687	243	228	3.00
4	*371.33	*371.33		**332.06	**316.00	*4.00
5	*464.16	*464.16		**409.04	**402.00	*5.00

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

Z-Values and Boundaries at Stage 0

This section gives the boundaries that were initially projected at the planning stage.

Z-Values and Boundaries at Stage 0

Maximum Information: 86.5248
Alternative Hypothesis: h1 - h2 < 0 (one-sided)
Futility Boundaries: Non-Binding
Z Statistic: MLE

| | | Boundaries | | | | |
| | Z-Test | | | | Information | |
Stage	Value	Efficacy	Futility	Time	Proportion	Decision
1		-6.4774	0.7745	1.0	0.1153	
2		-3.7865	-0.3140	2.0	0.3211	
3		-2.8249	-1.0151	3.0	0.5448	
4		-2.3268	-1.5457	4.0	0.7720	
5		-2.0211	-2.0211	5.0	1.0000	

Group-Sequential Boundary Plot at Planning Stage (Stage 0)

This plot shows the projected boundaries at the planning stage.

Group-Sequential Boundary Plot at Planning Stage (Stage 0)
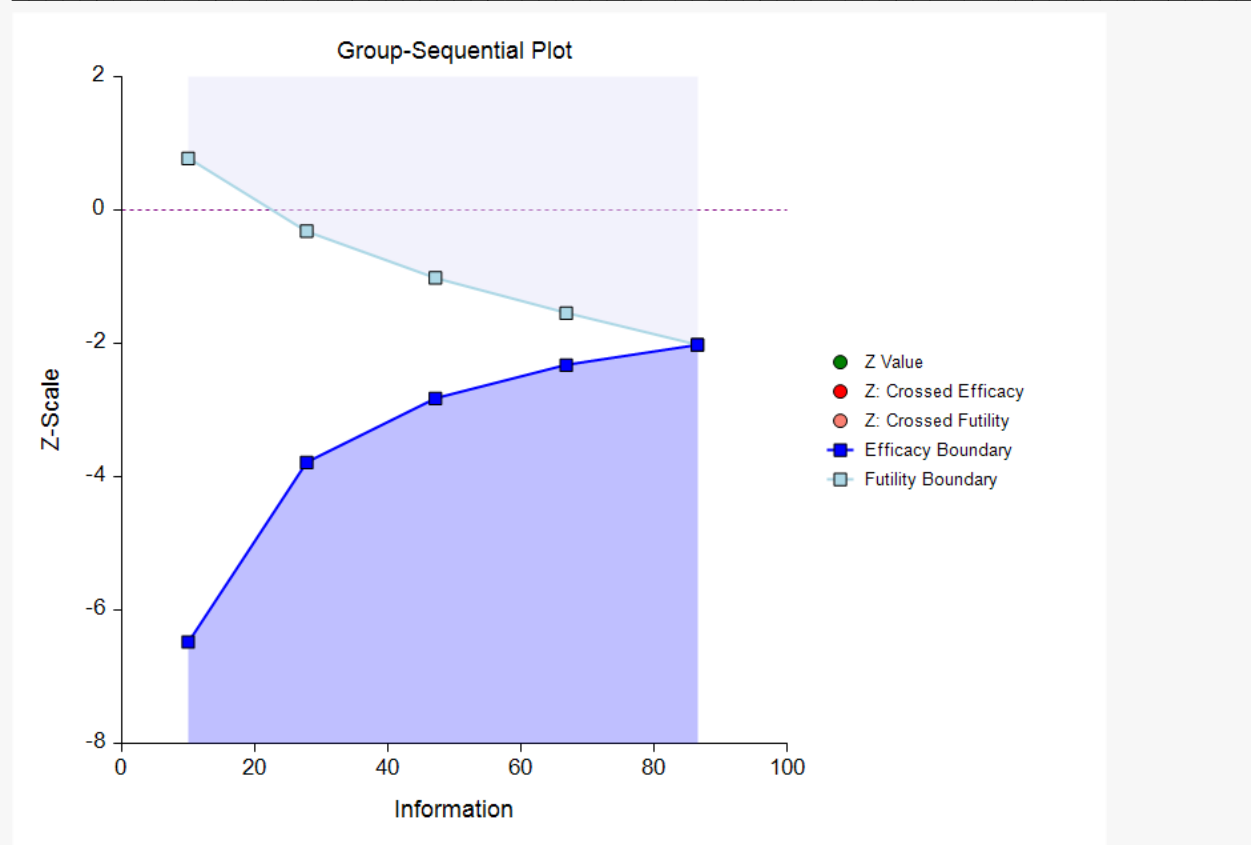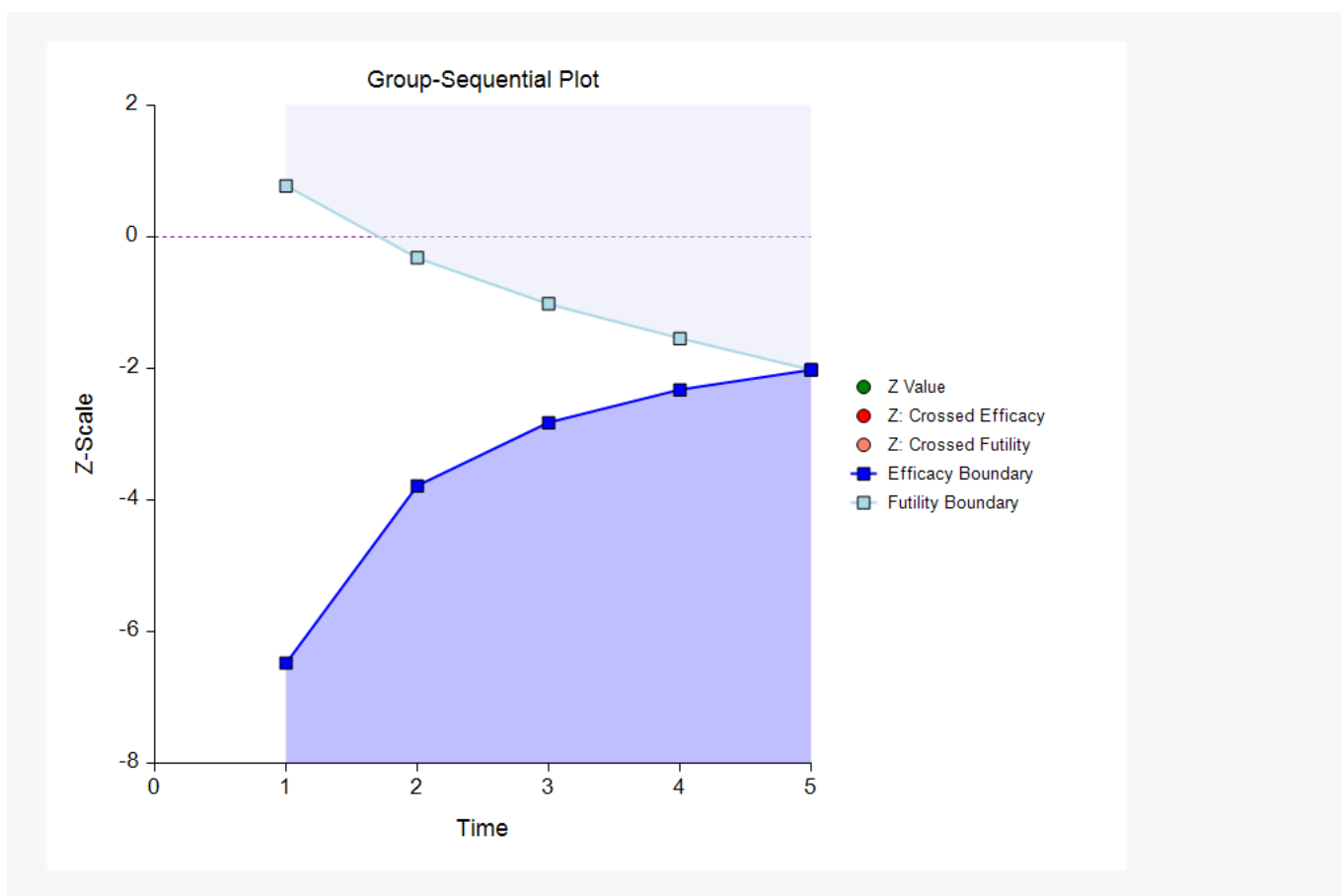
Kaplan-Meier Survival Plot

In order to obtain a Kaplan-Meier survival curve plot, the data must be converted into a form that can be used in the Kaplan-Meier Curves (Logrank Tests) procedure:

1. Rows with blank end times should be identified as censored.

2. Blank end times must be filled in with the current stage time.

3. An Elapsed Time column should be created by subtracting start times from end times.

When these steps are taken, the GS Survival dataset becomes the GS Survival B dataset.

Setup

To run this example, complete the following steps:

1   Open the GS Survival B example dataset

- From the File menu of the NCSS Data window, select **Open Example Data**.
- Select **GS Survival B** and click **OK**.

2   Specify the Kaplan-Meier Curves (Logrank Tests) procedure options

- Find and open the **Kaplan-Meier Curves (Logrank Tests)** procedure using the menus or the Procedure Navigator.

- The settings for this example are listed below and are stored in the **Example GS** settings file. To load these settings to the procedure window, click **Open Example Settings File** in the Help Center or File menu.

Variables Tab
_____

(Elapsed) Time Variable ....**Elapsed**
Censor Variable...**Censor**
  Failed..**0**
  Censored **1**
Group Variable..**Group**

Reports Tab
_____

Data Summary...**Checked**
All Other Reports ....**Unchecked**

Plots Tab
_____

Individual-Group Plots ....**Unchecked**
Combined Plot(s) ....**Checked**

Kaplan-Meier Survival/Reliability Plot Format (*Click the Button*)
_____

Kaplan-Meier Survival Line..............................**Checked**
Confidence Limits ....**Checked**
Confidence Limits - Fill ....**Checked**

3   Run the procedure

- Click the **Run** button to perform the calculations and generate the output.

Kaplan-Meier Survival Curve(s)

This plot shows the survival curves for the two groups.

Kaplan-Meier Survival Curve(s)
_____
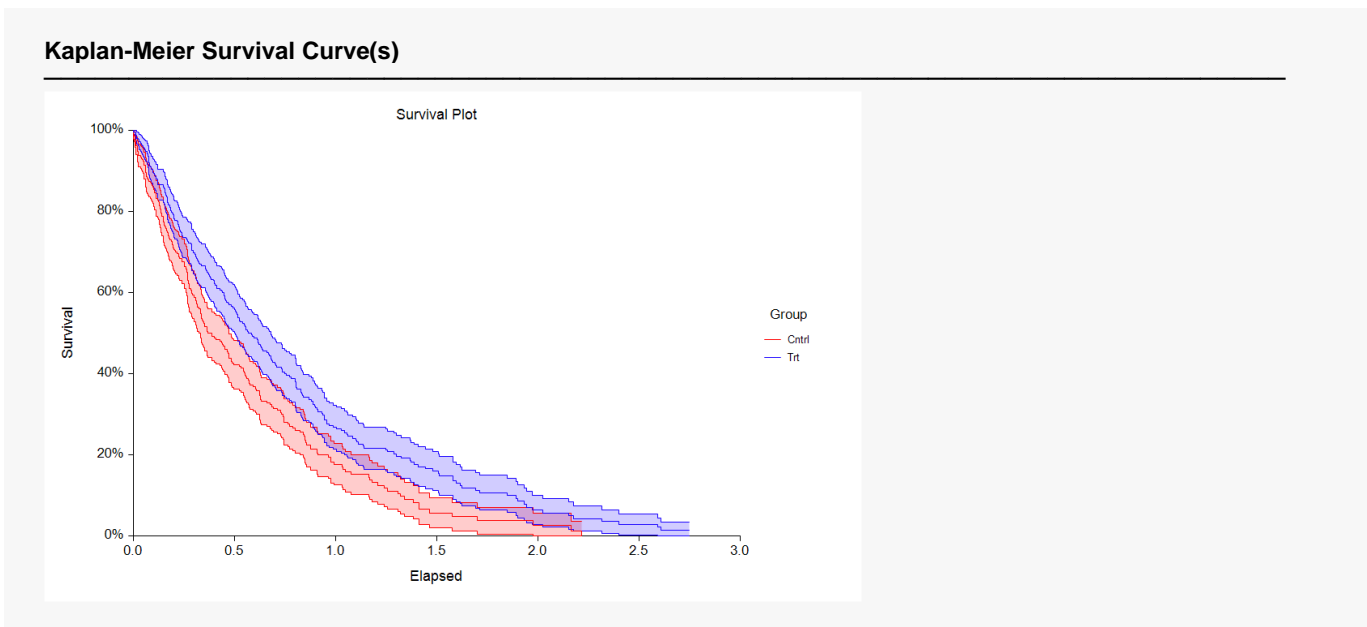
Example 2 – Skipping Stage Boundaries

Suppose that the setup is the same as in Example 1, except that the first two futility boundaries are skipped.

Setup

To run this example, complete the following steps:

1 Open the GS Survival example dataset

- From the File menu of the NCSS Data window, select **Open Example Data**.
- Select **GS Survival** and click **OK**.

2 Specify the Group-Sequential Analysis for Two Hazard Rates procedure options

- Find and open the **Group-Sequential Analysis for Two Hazard Rates** procedure using the menus or the Procedure Navigator.
- The settings 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.

  Variables Tab

  Skipped Futility Stages **1 2**

3 Run the procedure

- Click the **Run** button to perform the calculations and generate the output.

Run Summary Report

The skipped futility stages are now reported.

Run Summary Report

Item	Value
Maximum Number of Stages (Design):	5
Skipped Futility Stage(s):	1 2
Current Stage:	3
Alternative Hypothesis:	h1 - h2 < 0 (one-sided)
Alpha Spending Function:	O'Brien-Fleming Analog
Beta Spending Function:	Hwang-Shih-DeCani (γ = 1.5)
Futility Boundaries:	Non-Binding
Alpha:	0.0250
Maximum Information:	86.5248
Z Statistic:	MLE

Z-Values and Boundaries at Stage 3

The futility boundaries change slightly from those where no boundaries are skipped.

Z-Values and Boundaries at Stage 3

Maximum Information: 86.5248
Alternative Hypothesis: h1 - h2 < 0 (one-sided)
Futility Boundaries: Non-Binding
Z Statistic: MLE

		Boundaries			Information	
Stage	Z-Test Value	Efficacy	Futility	Time	Proportion	Decision
1	-2.3797	-6.4316		1.0	0.1173	Continue
2	-2.1001	-3.5628		2.0	0.3590	Continue
3	-3.3687	-2.7086	-1.3788	3.0	0.5871	Crossed Efficacy
4		-2.3412	-1.5678	4.0	0.7707	
5		-2.0218	-2.0218	5.0	1.0000	

Group-Sequential Boundary Plot at Stage 3

The plot now has two futility boundaries skipped.

Group-Sequential Boundary Plot at Stage 3
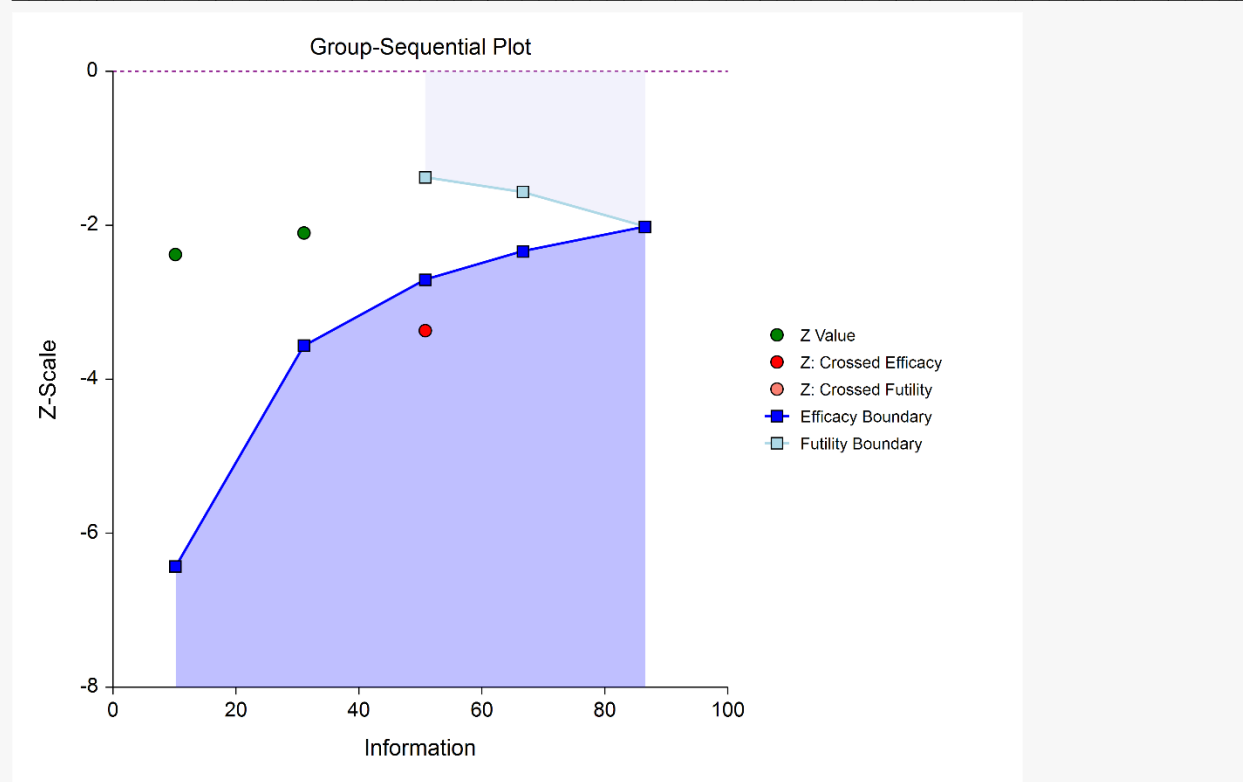
Group-Sequential Analysis for Two Hazard Rates
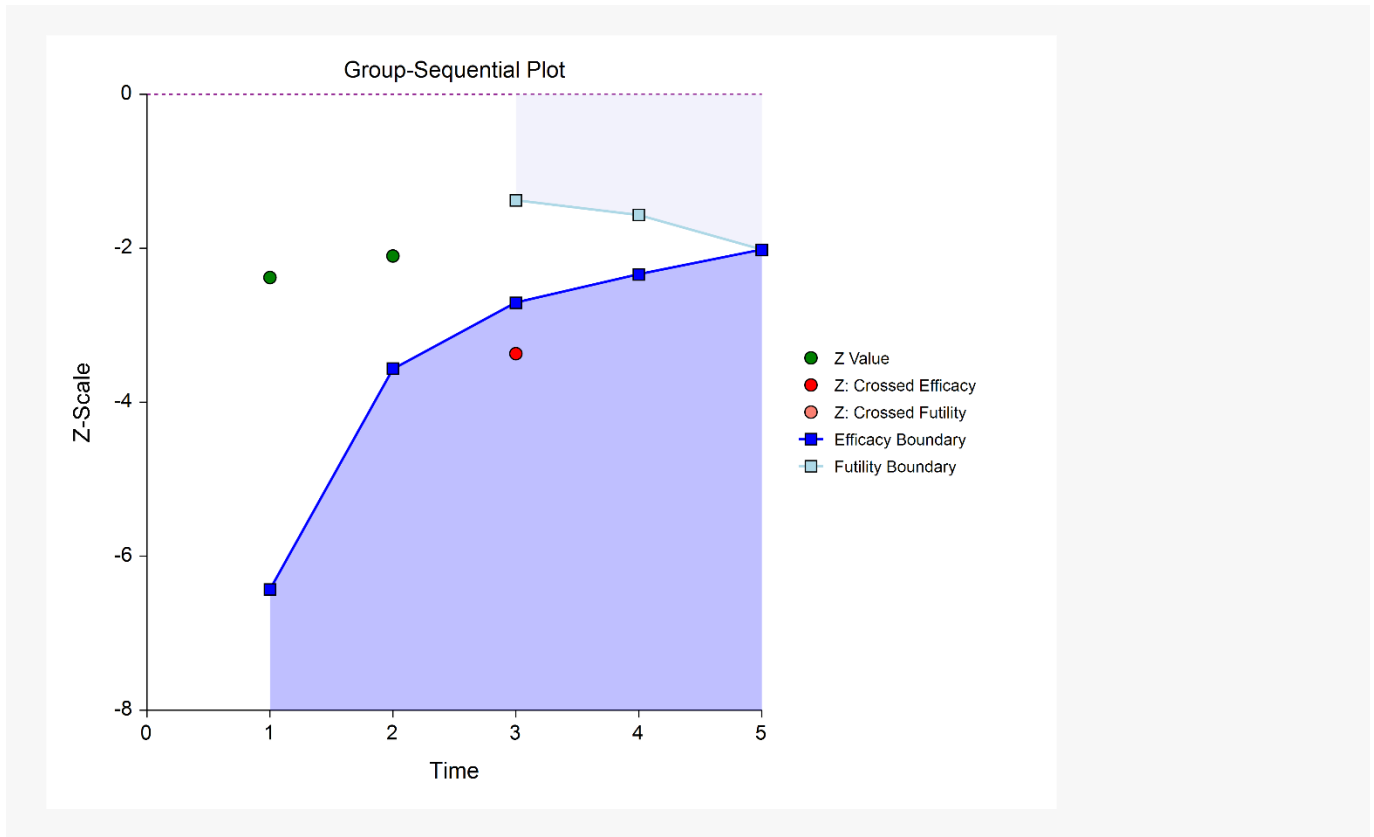
Example 3 – Two-Sided Boundaries

Suppose that the setup is the same as in Example 1, except that the boundaries are now two-sided.

Setup

To run this example, complete the following steps:

1 Open the GS Survival example dataset
- From the File menu of the NCSS Data window, select **Open Example Data**.
- Select **GS Survival** and click **OK**.

2 Specify the Group-Sequential Analysis for Two Hazard Rates procedure options
- Find and open the **Group-Sequential Analysis for Two Hazard Rates** procedure using the menus or the Procedure Navigator.
- The settings 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.

Variables Tab

Boundaries Used ....**Two-sided Efficacy with Futility (Symmetric)**
Alpha **0.05**
Overlapped Futility Boundaries........................**Remove (skip) overlapped futility boundaries**

3 Run the procedure
- Click the **Run** button to perform the calculations and generate the output.

Run Summary Report

Run Summary Report

Item	Value
Maximum Number of Stages (Design):	5
Current Stage:	3
Alternative Hypothesis:	h1 - h2 ≠ 0 (two-sided, symmetric)
Alpha Spending Function:	O'Brien-Fleming Analog
Beta Spending Function:	Hwang-Shih-DeCani (γ = 1.5)
Futility Boundaries:	Non-Binding
Alpha:	0.0500
Maximum Information:	86.5248
Z Statistic:	MLE

# Z-Values and Boundaries at Stage 3

Z-Values and Boundaries at Stage 3

Maximum Information: 86.5248
Alternative Hypothesis: h1 - h2 ≠ 0 (two-sided, symmetric)
Futility Boundaries: Non-Binding
Z Statistic: MLE

| | | Boundaries | | | | | | |
| | | Upper Side | | Lower Side | | | | |
Stage	Z-Test Value	Efficacy 1	Futility 1	Efficacy 2	Futility 2	Time	Information Proportion	Decision
1	-2.3797	6.4316		-6.4316		1.0	0.1173	Continue
2	-2.1001	3.5628	0.6212	-3.5628	-0.6212	2.0	0.3590	Crossed Fut 1
3	-3.3687	2.7086	1.1576	-2.7086	-1.1576	3.0	0.5871	Crossed F1 & E2
4		2.3412	1.5278	-2.3412	-1.5278	4.0	0.7707	
5		2.0218	2.0218	-2.0218	-2.0218	5.0	1.0000	

Group-Sequential Boundary Plot at Stage 3

Because they were overlapping, the first futility boundaries are skipped.

Group-Sequential Boundary Plot at Stage 3

Group-Sequential Analysis for Two Hazard Rates
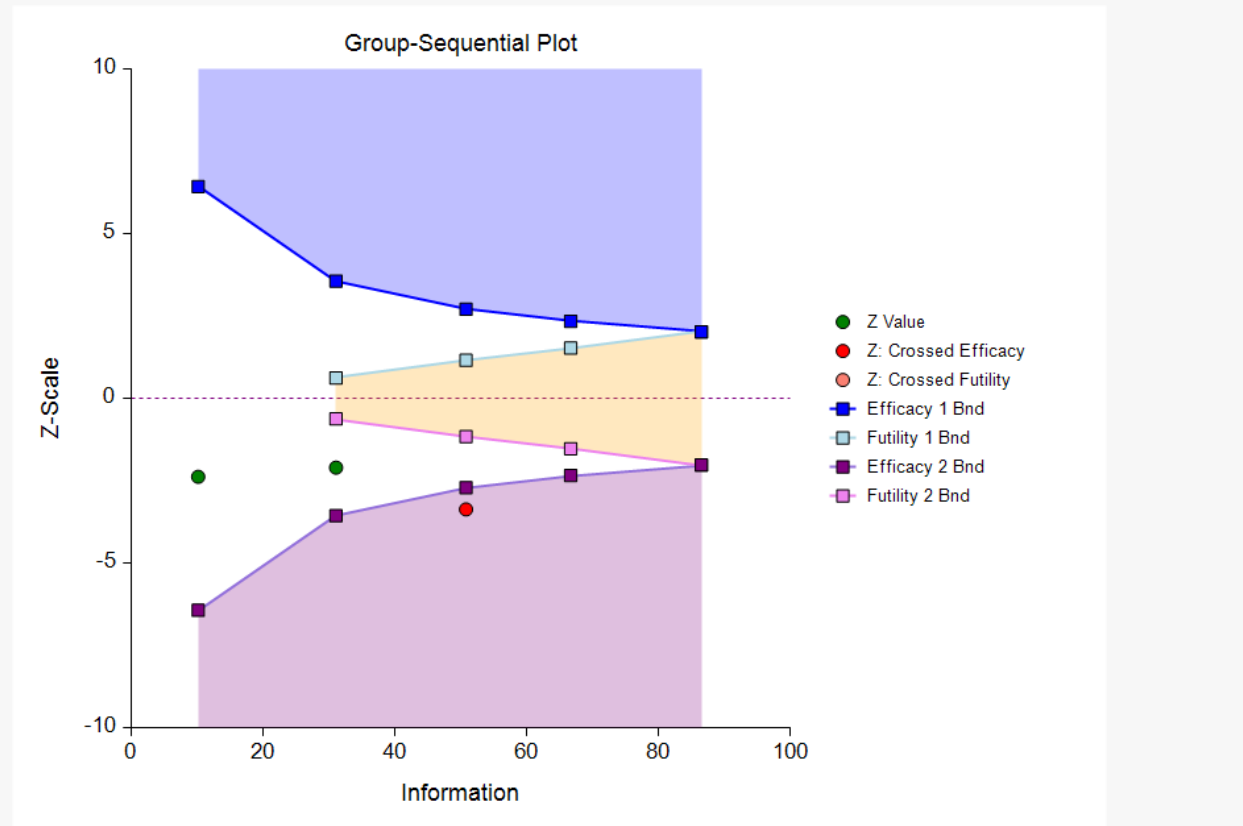
Group-Sequential Plot
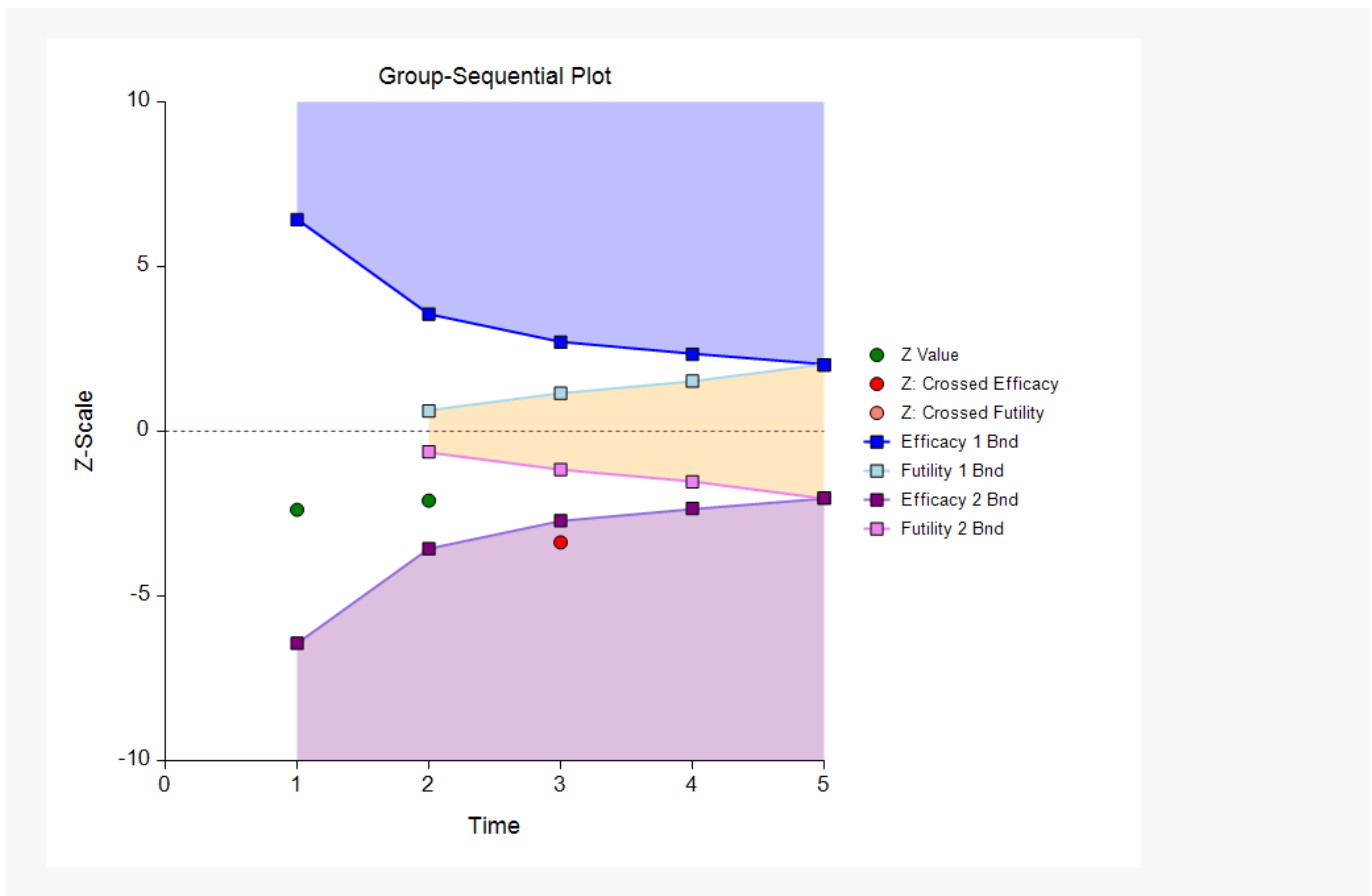
Example 4 – Sample Size Re-estimation and Boundary Probabilities

Suppose that the setup is the same as in Example 1, except that only the data up through Stage 2 has been collected.

Setup

To run this example, complete the following steps:

1 Open the GS Survival 2 Stages example dataset
- From the File menu of the NCSS Data window, select **Open Example Data**.
- Select GS Survival 2 Stages and click OK.

2 Specify the Group-Sequential Analysis for Two Hazard Rates procedure options
- Find and open the **Group-Sequential Analysis for Two Hazard Rates** procedure using the menus or the Procedure Navigator.
- The settings 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.

> Variables Tab
> _____
>
> Current Stage (k) **2**

3 Run the procedure
- Click the **Run** button to perform the calculations and generate the output.

Run Summary Report

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

Run Summary Report

Item	Value
Maximum Number of Stages (Design):	5
Current Stage:	2
Alternative Hypothesis:	h1 - h2 < 0 (one-sided)
Alpha Spending Function:	O'Brien-Fleming Analog
Beta Spending Function:	Hwang-Shih-DeCani (γ = 1.5)
Futility Boundaries:	Non-Binding
Alpha:	0.0250
Maximum Information:	86.5248
Z Statistic:	MLE

Z-Values and Boundaries at Stage 2

Examining the Z-Test values, the boundaries, and the decisions, no boundary has been crossed at this stage.

Z-Values and Boundaries at Stage 2

Maximum Information: 86.5248
Alternative Hypothesis: h1 - h2 < 0 (one-sided)
Futility Boundaries: Non-Binding
Z Statistic: MLE

| | | Boundaries | | | Information | |
Stage	Z-Test Value	Efficacy	Futility	Time	Proportion	Decision
1	-2.3797	-6.4316	0.7577	1.0	0.1173	Continue
2	-2.1001	-3.5628	-0.4846	2.0	0.3590	Continue
3		-2.8460	-0.9724	3.0	0.5394	
4		-2.3313	-1.5382	4.0	0.7691	
5		-2.0202	-2.0202	5.0	1.0000	

Group-Sequential Boundary Plot at Stage 2

This plot shows the Z-test values in the vicinity of the efficacy boundary side.

Group-Sequential Boundary Plot at Stage 2
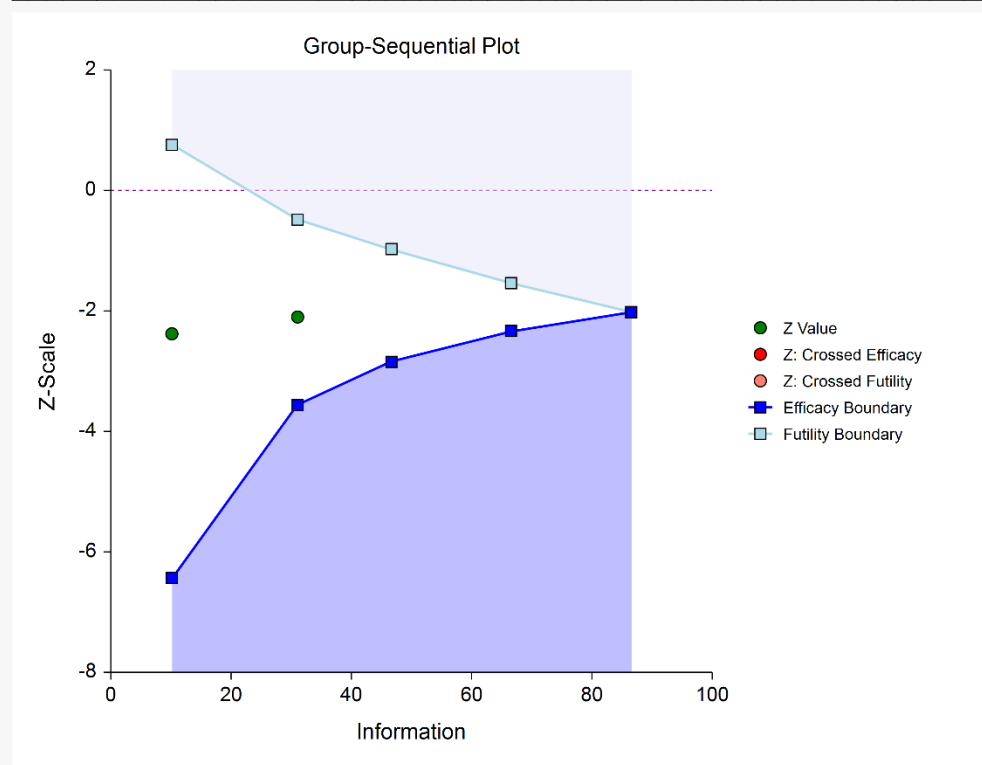
Group-Sequential Plot
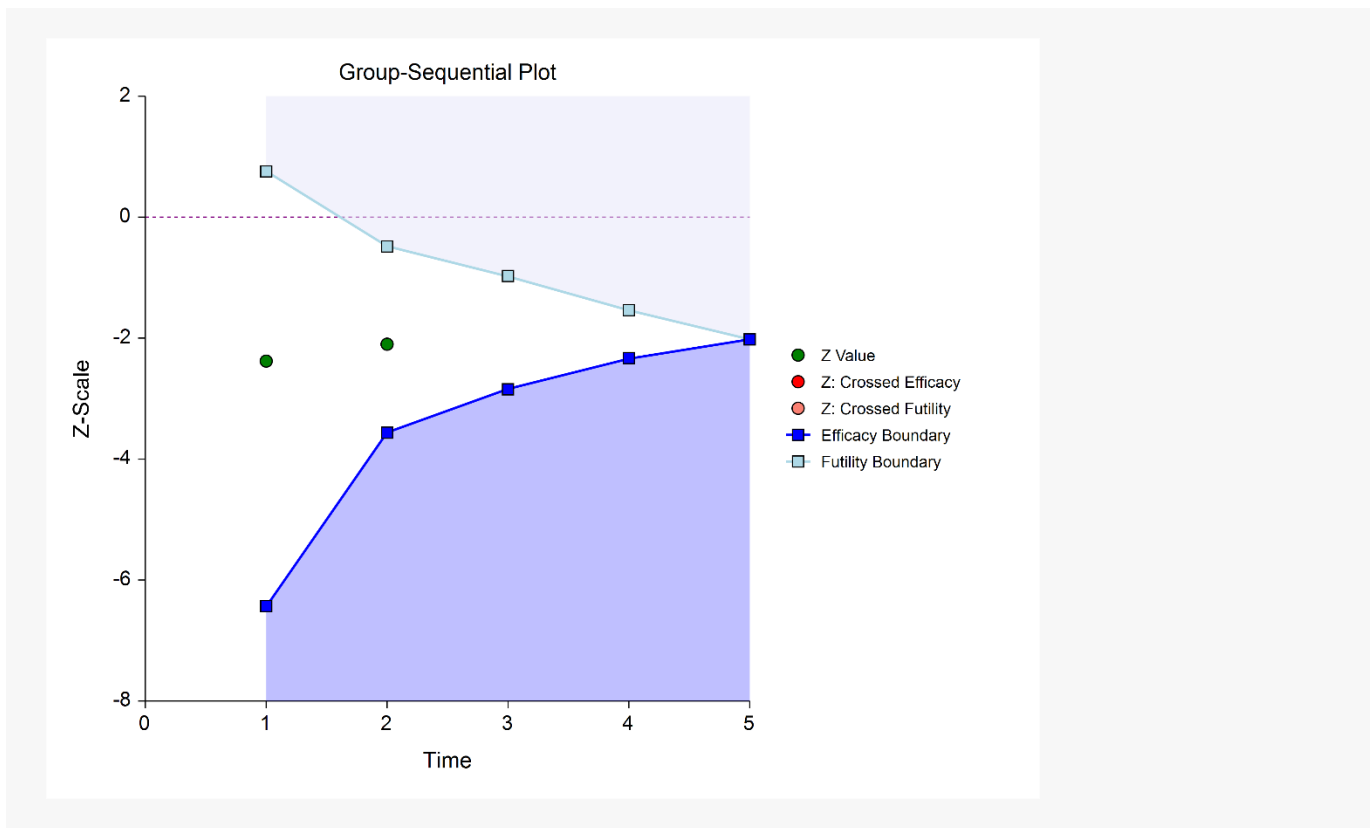
Descriptive Statistics up to Stage 2

Descriptive Statistics up to Stage 2

Stage	Sample Size		Event Count		Sample Hazard Rates		Difference	
	N1	N2	E1	E2	h1	h2	Difference	SE(Diff)
1	116	90	48	46	1.09335	1.84031	-0.74696	0.31389
2	219	184	145	122	1.24368	1.62048	-0.37680	0.17942

Information Report at Stage 2 (Gives Sample Size Re-estimation)

This section shows that the target sample size for the next stage should be 256 per group.

Information Report at Stage 2

Maximum Information: 86.5248
Alternative Hypothesis: h1 - h2 < 0 (one-sided)
Alpha: 0.0250

Stage	Target Information Proportion	Achieved Information Proportion	Target Information	Achieved Information	N1	N2	h1	h2
1	0.1153	0.1173	9.9780	10.1492	116.00	90.00	1.09335	1.84031
2	0.3211	0.3590	27.7831	31.0642	219.00	184.00	1.24368	1.62048
3	0.5448	*0.5394	47.1361	*46.6735	*255.51	*255.51	*1.24368	*1.62048
4	0.7720	*0.7691	66.7992	*66.5502	*340.68	*340.68	*1.24368	*1.62048
5	1.0000	*1.0000	86.5248	*86.5248	*425.86	*425.86	*1.24368	*1.62048

* Projected value

Alpha Spending at Stage 2

This section shows how alpha was spent (or is anticipated to be spent) across the stages.

Alpha Spending at Stage 2

Target Final Stage Alpha: 0.0250
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	0.1173	0.0000	0.0000	0.000000	0.0	0.0
2	0.3590	0.0002	0.0002	0.000183	0.7	0.7
3 *	0.5394	0.0021	0.0023	0.002214	8.4	9.1
4 *	0.7691	0.0083	0.0106	0.009870	33.3	42.4
5 *	1.0000	0.0144	0.0250	0.021681	57.6	100.0

* Projected

Beta Spending for Futility at Stage 2

This section shows how beta was spent (or is anticipated to be spent) across the stages.

Beta Spending for Futility at Stage 2

Target Cumulative Beta at Final Stage: 0.1000
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	0.1173	0.0208	0.0208	0.775687	20.8	20.8
2	0.3590	0.0328	0.0536	0.313973	32.8	53.6
3 *	0.5394	0.0178	0.0714	0.165424	17.8	71.4
4 *	0.7691	0.0167	0.0881	0.061997	16.7	88.1
5 *	1.0000	0.0119	0.1000	0.021681	11.9	100.0

* Projected

Conditional and Predictive Power Report at Stage 2

Conditional power and predictive power are described earlier in the Technical details section. The predictive power does not depend on an assumed difference, as does the conditional power.

Conditional and Predictive Power Report at Stage 2

Maximum Information: 86.5248
Information this Stage: 31.0642
Proportion of Maximum Information: 0.3590
Predictive Power: 0.8762

δ Name	δ Value	Conditional Power
Design	-0.3500	0.9582
Data	-0.3768	0.9732
δ1	0.0000	0.1904

The conditional power values are calculated based on Jennison and Turnbull (2000), pages 205 to 208. They do not account for future interim stages, nor futility boundaries, if applicable.

Boundary Probabilities for δ = -0.35

Given the data that have already accumulated in the first two stages, and assuming hazard rates of 1.4 and 1.75 going forward, this report gives the simulation probabilities of crossing each of the future boundaries. The sum of the efficacy boundary probabilities is another estimate of the conditional power.

Boundary Probabilities for δ = -0.35
_____

Number of Simulations:	2000
User-Entered Random Seed:	5863243
Futility Boundaries:	Non-Binding
After Efficacy Boundary Crossing:	Hold Out
After Non-Binding Futility Boundary Crossing:	Leave In
Alternative Hypothesis:	h1 - h2 < 0 (one-sided)
Z Statistic:	MLE
h1:	1.4
h2:	1.75
δ:	-0.35

_____

				Efficacy		Futility	
Stage	N1	N2	Z-Test Value	Boundary	Probability	Boundary	Probability
1	116.00	90.00	-2.3797	-6.4316		0.7577	
2	219.00	184.00	-2.1001	-3.5628		-0.4846	
3	*255.51	*255.51		-2.8460	0.3715	-0.9724	0.0055
4	*340.68	*340.68		-2.3313	0.4680	-1.5382	0.0160
5	*425.86	*425.86		-2.0202	0.1205	-2.0202	0.0470

_____

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

Event Summary for δ = -0.35

Using simulation based on the specified hazard rates, this section gives the estimated number of events at each future stage. Values given here will vary for each simulation.

Event Summary for δ = -0.35
_____

Number of Simulations:	2000
User-Entered Random Seed:	5863243
h1:	1.4
h2:	1.75
δ:	-0.35

_____

				Average Cumulative Number of Events		
Stage	N1	N2	Z-Test Value	E1	E2	Time
1	116	90	-2.3797	48	46	1.00
2	219	184	-2.1001	145	122	2.00
3	*255.51	*255.51		**228.77	**212.55	*3.00
4	*340.68	*340.68		**300.28	**293.53	*4.00
5	*425.86	*425.86		**368.59	**373.51	*5.00

_____

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

Boundary Probabilities for δ = - 0.376804

Given the data that have already accumulated in the first two stages, and assuming hazard rates of 1.243679 and 1.620483 going forward, this report gives the simulation probabilities of crossing each of the future boundaries.

Boundary Probabilities for δ = -0.376804
───

Number of Simulations:	2000
User-Entered Random Seed:	5863243
Futility Boundaries:	Non-Binding
After Efficacy Boundary Crossing:	Hold Out
After Non-Binding Futility Boundary Crossing:	Leave In
Alternative Hypothesis:	h1 - h2 < 0 (one-sided)
Z Statistic:	MLE
h1:	1.243679
h2:	1.620483
δ:	-0.376804

			Z-Test	Efficacy		Futility	
Stage	N1	N2	Value	Boundary	Probability	Boundary	Probability
1	116.00	90.00	-2.3797	-6.4316		0.7577	
2	219.00	184.00	-2.1001	-3.5628		-0.4846	
3	*255.51	*255.51		-2.8460	0.4115	-0.9724	0.0045
4	*340.68	*340.68		-2.3313	0.4815	-1.5382	0.0095
5	*425.86	*425.86		-2.0202	0.0865	-2.0202	0.0230

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

Event Summary for δ = -0.376804

Event Summary for δ = -0.376804
───

Number of Simulations:	2000
User-Entered Random Seed:	5863243
h1:	1.243679
h2:	1.620483
δ:	-0.376804

				Average Cumulative Number of Events		
Stage	N1	N2	Z-Test Value	E1	E2	Time
1	116	90	-2.3797	48	46	1.00
2	219	184	-2.1001	145	122	2.00
3	*255.51	*255.51		**223.38	**209.43	*3.00
4	*340.68	*340.68		**294.02	**290.13	*4.00
5	*425.86	*425.86		**362.44	**369.59	*5.00

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

Boundary Probabilities for δ = 0

Given the data that have already accumulated in the first two stages, and assuming hazard rates of 1.6 and 1.6 (no difference) going forward, this report gives the simulation probabilities of crossing each of the future boundaries.

Boundary Probabilities for δ = 0

Number of Simulations:	2000
User-Entered Random Seed:	5863243
Futility Boundaries:	Non-Binding
After Efficacy Boundary Crossing:	Hold Out
After Non-Binding Futility Boundary Crossing:	Leave In
Alternative Hypothesis:	h1 - h2 < 0 (one-sided)
Z Statistic:	MLE
h1:	1.6
h2:	1.6
δ:	0

				Efficacy		Futility	
Stage	N1	N2	Z-Test Value	Boundary	Probability	Boundary	Probability
1	116.00	90.00	-2.3797	-6.4316		0.7577	
2	219.00	184.00	-2.1001	-3.5628		-0.4846	
3	*255.51	*255.51		-2.8460	0.0365	-0.9724	0.0950
4	*340.68	*340.68		-2.3313	0.1355	-1.5382	0.4540
5	*425.86	*425.86		-2.0202	0.1085	-2.0202	0.7580

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

Event Summary for δ = 0

Event Summary for δ = 0

Number of Simulations:	2000
User-Entered Random Seed:	5863243
h1:	1.6
h2:	1.6
δ:	0

				Average Cumulative Number of Events		
Stage	N1	N2	Z-Test Value	E1	E2	Time
1	116	90	-2.3797	48	46	1.00
2	219	184	-2.1001	145	122	2.00
3	*255.51	*255.51		**234.44	**208.63	*3.00
4	*340.68	*340.68		**306.46	**289.08	*4.00
5	*425.86	*425.86		**375.10	**368.54	*5.00

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