

Chapter 710

Group-Sequential Logrank Tests

Introduction

This procedure uses analytic methods for sample size and power calculations for group-sequential Logrank tests. PASS also has procedures for group-sequential Logrank tests based on simulations. The simulation procedures are more flexible in terms of input parameters, and also allow for both significance and futility boundaries.

Clinical trials are longitudinal. They accumulate data sequentially through time. The participants cannot be enrolled and randomized on the same day. Instead, they are enrolled as they enter the study. It may take several years to enroll enough patients to meet sample size requirements. Because clinical trials are long term studies, it is in the interest of both the participants and the researchers to monitor the accumulating information for early convincing evidence of either harm or benefit. This permits early termination of the trial.

Group sequential methods allow statistical tests to be performed on accumulating data while a phase III clinical trial is ongoing. Statistical theory and practical experience with these designs have shown that making four or five *interim analyses* is almost as effective in detecting large differences between treatment groups as performing a new analysis after each new data value. Besides saving time and resources, such a strategy can reduce the experimental subject's exposure to an inferior treatment and make superior treatments available sooner.

When repeated significance testing occurs on the same data, adjustments have to be made to the hypothesis testing procedure to maintain overall significance and power levels. The landmark paper of Lan & DeMets (1983) provided the theory behind the *alpha spending function* approach to group sequential testing. This paper built upon the earlier work of Armitage, McPherson, & Rowe (1969), Pocock (1977), and O'Brien & Fleming (1979). PASS implements the methods given in Reboussin, DeMets, Kim, & Lan (1992) to calculate the power and sample sizes of various group sequential designs.

This module calculates sample size and power for group sequential designs used to compare two survival curves. Other modules perform similar analyses for the comparison of means and proportions. The program allows you to vary number and times of interim tests, type of alpha spending function, and test boundaries. It also gives you complete flexibility in solving for power, significance level, sample size, or effect size. The results are displayed in both numeric reports and informative graphics.

Technical Details

In many clinical trials, patients are recruited and randomized to receive a particular treatment, either experimental or control, and then monitored until either a critical event occurs or the study is ended. The length of time the patient is monitored until the critical event occurs is called the *follow-up time*. After the study is ended, the follow-up times of the patients in the two groups are compared using the *logrank test* in what is often called a *survival analysis*.

When the critical event does not occur for a patient by the time the study is ended, the follow-up time is said to have been *censored*. Although the actual event time is not known for this patient, it is known that the event time

Group-Sequential Logrank Tests

will be greater than the follow-up time. Hence, some information is gleaned from these participants. Because of this censoring, the usual tests of means or proportions cannot be used. The logrank test was developed to provide a statistical test comparing the efficacy of the two treatments.

The Hazard Ratio (HR)

Suppose the critical event is death. The survival distribution of each treatment can be characterized by the *instantaneous death rates*, λ_1 and λ_2 . An instantaneous death rate, often called the *hazard*, is the probability of death in a short interval of time. The comparison of the efficacies of the two treatments is often formalized by considering the ratio of the hazard, or *hazard ratio* (HR).

$$HR = \frac{\lambda_2}{\lambda_1}$$

Although the logrank test concerns the ratio of the hazard rates of the two groups, when planning a study, it may be easier to obtain information about the expected proportion surviving during the trial. It turns out that the hazard ratio can be computed from the survival proportions, $S1$ and $S2$, using the equation

$$HR = \frac{\log(S2)}{\log(S1)}$$

when the hazard ratio is constant through time.

Assuming that group one is the control group, it may be easiest during the planning stages of a study to find $S1$ and state the minimum value of HR that would make the experimental treatment useful. The last equation can then be manipulated to calculate a value for $S2$ as follows

$$S2 = \exp\{HR(\log(S1))\}.$$

Sometimes it is more convenient to state hazard ratio in terms of the median survival times. In this case, the hazard ratio is estimated using

$$HR = \frac{M_1}{M_2}$$

when the hazard ratio is constant for different times.

The Logrank Statistic

The following results are excerpted from Reboussin (1992). The logrank statistic is given by the equation

$$L(d) = \sum_{i=1}^d \left(\frac{x_i r_{ic}}{r_{ic} + r_{it}} \right)$$

where d is the number of events, x_i is 1 if the event at time t_i is in the control group and 0 if it is in the treatment group, r_{ic} is the number of patients in the control group at risk just before t_i , and r_{it} is the corresponding number of patients at risk in the treatment group.

If $r_{ic} \approx r_{it}$ and HR is close to 1, then the sequential logrank statistic is (approximately)

$$z_k = 2 \frac{L(d_k)}{\sqrt{d_k}}.$$

Group-Sequential Logrank Tests

The subscript k indicates that the computations use all data that are available at the time of the k^{th} interim analysis or k^{th} look (k goes from 1 to K).

Spending Functions

Lan and DeMets (1983) introduced alpha spending functions, $\alpha(\tau)$, that determine a set of boundaries b_1, b_2, \dots, b_K for the sequence of test statistics z_1, z_2, \dots, z_K . These boundaries are the critical values of the sequential hypothesis tests. That is, after each interim test, the trial is continued as long as $|z_k| < b_k$. When $|z_k| \geq b_k$, the hypothesis of equal means is rejected and the trial is stopped early.

The time argument τ either represents the proportion of elapsed time to the maximum duration of the trial or the proportion of the sample that has been collected. When elapsed time is used, it is referred to as *calendar time*. When time is measured in terms of the sample, it is referred to as *information time*. Since it is a proportion, τ can only vary between zero and one.

Alpha spending functions have the characteristics:

$$\alpha(0) = 0$$

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

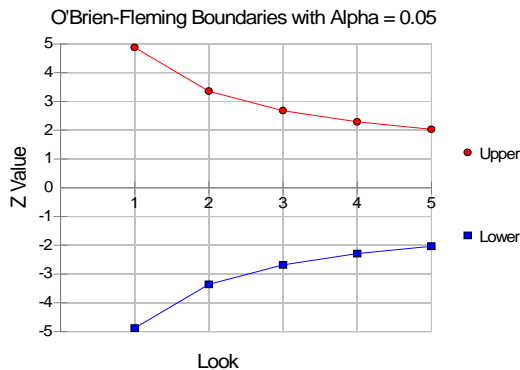
The last characteristic guarantees a fixed α level when the trial is complete. That is,

$$\Pr(|z_1| \geq b_1 \text{ or } |z_2| \geq b_2 \text{ or } \dots \text{ or } |z_k| \geq b_k) = \alpha(\tau)$$

This methodology is very flexible since neither the times nor the number of analyses must be specified in advance. Only the functional form of $\alpha(\tau)$ must be specified.

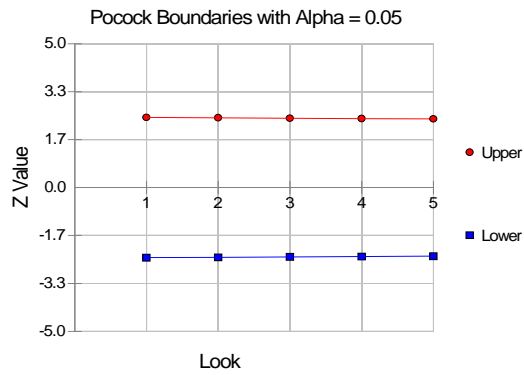
PASS provides five popular spending functions plus the ability to enter and analyze your own boundaries. These are calculated as follows:

1. **O'Brien-Fleming** $2 - 2\Phi\left(\frac{Z_{\alpha/2}}{\sqrt{\tau}}\right)$

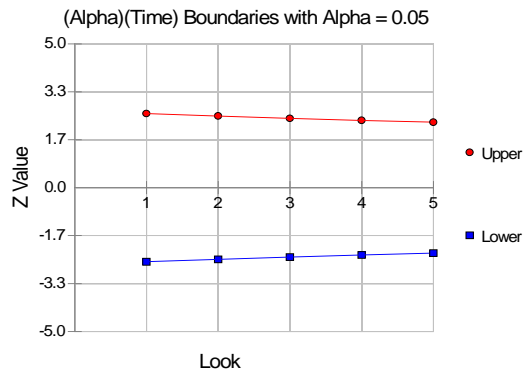


Group-Sequential Logrank Tests

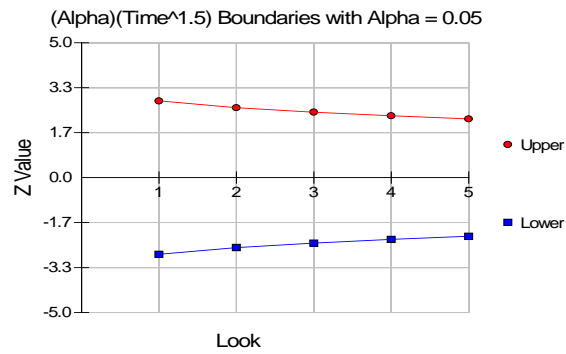
2. Pocock $\alpha \ln(1 + (e - 1)\tau)$



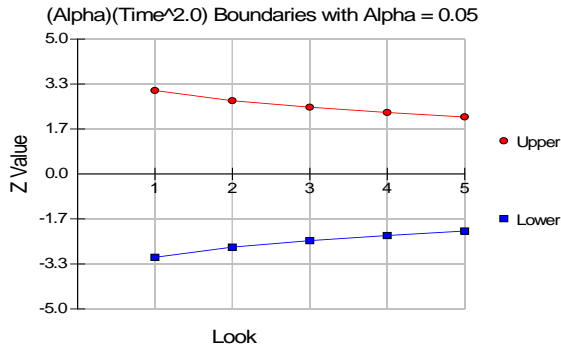
3. Alpha * time $\alpha\tau$



4. Alpha * time^1.5 $\alpha\tau^{3/2}$



Group-Sequential Logrank Tests

5. Alpha * time² $\alpha\tau^2$ 

6. User Supplied

A custom set of boundaries may be entered.

The O'Brien-Fleming boundaries are commonly used because they do not significantly increase the overall sample size and because they are conservative early in the trial. Conservative in the sense that the means must be extremely different before statistical significance is indicated. The Pocock boundaries are nearly equal for all times. The Alpha*t boundaries use equal amounts of alpha when the looks are equally spaced. You can enter your own set of boundaries using the User Supplied option.

Sequential Theory

A detailed account of the methodology is contained in Lan & DeMets (1983), DeMets & Lan (1984), Lan & Zucker (1993), and DeMets & Lan (1994). A brief summary of the theoretical basis of the method will be presented here.

Group sequential procedures for interim analysis are based on their equivalence to discrete boundary crossing of a Brownian motion process with drift parameter θ . The test statistics z_k follow the multivariate normal distribution with means $\theta\sqrt{\tau_k}$ and, for $j \leq k$, covariances $\sqrt{\tau_k / \tau_j}$. The drift is related to the parameters of the z-test through one of the equations

$$\theta = \frac{\log(HR)\sqrt{d_k}}{2} \quad (\text{exponential survival}) \quad \text{or} \quad \theta = \frac{[1 - HR]\sqrt{d_k}}{(1 + HR)} \quad (\text{proportional hazards}).$$

These equations may be solved for d_k , the required number of events, giving

$$d_k = \frac{4\theta^2}{[\log(HR)]^2} \quad (\text{exponential survival}) \quad \text{or} \quad d_k = \left(\frac{(1 + HR)\theta}{1 - HR} \right)^2 \quad (\text{proportional hazards}).$$

In survival analysis, the size of a sample is measured in terms of number of events rather than number of patients because it is probable that many of the patients will be censored—their event times are not known. In order to ensure that the sample size produces the required number of events, it must be inflated by the event rates.

The expected number of events can be computed from the proportion surviving using the equation

$$d_k = \frac{N(1 - S_1) + N(1 - S_2)}{2}$$

Group-Sequential Logrank Tests

where N is the total sample size (assumed to be split evenly between groups). This can be solved for N to give the sample size as

$$N = \frac{2d_k}{2 - S_1 - S_2}.$$

Hence, the algorithm is as follows:

1. Compute boundary values based on a specified spending function and alpha value.
2. Calculate the drift parameter based on those boundary values and a specified power value.
3. Use the drift parameter and the above equation to calculate the appropriate event size per group d_k .
4. Use the event size to compute the appropriate sample size, N .

Procedure Tabs

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

Design Tab

The Design tab contains the parameters associated with the z-test such as the survival rates, sample sizes, alpha, and power.

Solve For

Solve For

This option specifies the parameter to be solved for from the other parameters. The parameters that may be selected are *S1*, *S2*, *Alpha*, *Power*, or *Sample Size*. Under most situations, you will select either *Power* or *Sample Size*.

Select *Sample Size* when you want to calculate the sample size needed to achieve a given power and alpha level.

Select *Power and Beta* when you want to calculate the power of an experiment.

Test

Alternative Hypothesis

Specify whether the test is one-sided or two-sided. When a two-sided hypothesis is selected, the value of alpha is halved. Everything else remains the same.

Note that the accepted procedure is to use Two-Sided unless you can justify using a one-sided test.

Survival Time Assumption

Calculations can be based on the assumption of exponential survival times or proportional hazards. The proportional hazards assumption is the less restrictive and the one that is recommended.

- **Exponential Survival**

The survival curves follow exponential distributions. This is the most restrictive assumption.

- **Proportional Hazards**

The hazards of the two groups are proportional.

Group-Sequential Logrank Tests

Power and Alpha

Power

This option specifies one or more values for power. Power is the probability of rejecting a false null hypothesis, and is equal to one minus Beta. Beta is the probability of a type-II error, which occurs when a false null hypothesis is not rejected. In this procedure, a type-II error occurs when you fail to reject the null hypothesis of equal survival curves when in fact the curves are different.

Values must be between zero and one. Historically, the value of 0.80 (Beta = 0.20) was used for power. Now, 0.90 (Beta = 0.10) is also commonly used.

A single value may be entered here or a range of values such as *0.8 to 0.95 by 0.05* may be entered.

Alpha

This option specifies one or more values for the probability of a type-I error. A type-I error occurs when you reject the null hypothesis of equal survival curves when in fact the curves are equal.

Values of alpha must be between zero and one. Historically, the value of 0.05 has been used for alpha. This means that about one test in twenty will falsely reject the null hypothesis. You should pick a value for alpha that represents the risk of a type-I error you are willing to take in your experimental situation.

You may enter a range of values such as *0.01 0.05 0.10* or *0.01 to 0.10 by 0.01*.

Sample Size

N (Total Sample Size)

Enter a value (or range of values) for the total sample size. Each group is assumed to have a sample size of $N/2$. You may enter a range of values such as *100 to 1000 by 100*.

Note that the sample size is based implicitly on the length of the study since we used the equation

$$N = \frac{2d_k}{2 - S_1 - S_2}$$

to estimate the necessary sample size.

Effect Size

S1 (Proportion Surviving 1)

$S1$ is the proportion of patients belonging to group 1 (controls) that are expected to survive during the study. Since $S1$ is a proportion, it must be between zero and one. In many studies, this is the proportion surviving in the regular population with the standard treatment.

You may enter a range of values such as *0.1, 0.2, 0.3* or *0.1 to 0.9 by 0.2*.

S2 (Proportion Surviving 2)

$S2$ is the proportion of patients in the experimental group that survive during the study. This is not necessarily the expected proportion. Rather, you may set it to that proportion that, if achieved, would be of special interest. Values below (or above) this amount would not be of interest. For example, if the standard one-year survival proportion is 0.2 and the new treatment raises this proportion to 0.3 (a 50% increase in the proportion surviving), others may be interested in adopting this new treatment.

Group-Sequential Logrank Tests

Sometimes, researchers wish to state the alternative hypothesis in terms of the hazard ratio, HR , rather than the value of $S2$. Using the fact that

$$HR = \frac{\log(S2)}{\log(S1)},$$

An appropriate value for $S2$ may be calculated from $S1$ and HR using the equation

$$S2 = \exp\{HR(\log(S1))\}.$$

Sometimes it is more convenient to state hazard ratio in terms of the median survival times. In this case, the hazard ratio is estimated using

$$HR = \frac{M_1}{M_2}$$

and the above equation is used to find $S2$.

Since $S2$ is a proportion, it must remain between zero and one. You may enter a range of values such as 0.1 , 0.2 , 0.3 or 0.1 to 0.9 by 0.2 .

Look Details

The Sequential tab contains the parameters associated with Group Sequential Design such as the type of spending function, the times, and so on.

Number of Looks

This is the number of interim analyses (including the final analysis). For example, a five here means that four interim analyses will be run in addition to the final analysis.

Boundary Truncation

You can truncate the boundary values at a specified value. For example, you might decide that no boundaries should be larger than 4.0. If you want to implement a boundary limit, enter the value here.

If you do not want a boundary limit, enter *None* here.

Spending Function

Specify which alpha spending function to use. The most popular is the O'Brien-Fleming boundary that makes early tests very conservative. Select *User Specified* if you want to enter your own set of boundaries.

Max Time

This is the total running time of the trial. It is used to convert the values in the Times box to fractions. The units (months or years) do not matter, as long as they are consistent with those entered in the Times box.

For example, suppose Max Time = 3 and Times = 1, 2, 3. Interim analyses would be assumed to have occurred at 0.33, 0.67, and 1.00.

Times

Enter a list of time values here at which the interim analyses will occur. These values are scaled according to the value of the Max Time option.

For example, suppose a 48-month trial calls for interim analyses at 12, 24, 36, and 48 months. You could set Max Time to 48 and enter *12,24,36,48* here or you could set Max Time to *1.0* and enter *0.25,0.50,0.75,1.00* here.

Group-Sequential Logrank Tests

The number of times entered here must match the value of the Number of Looks.

- **Equally Spaced**

If you are planning to conduct the interim analyses at equally spaced points in time, you can enter *Equally Spaced* and the program will generate the appropriate time values for you.

Informations

You can weight the interim analyses on the amount of information obtained at each time point rather than on actual calendar time. If you would like to do this, enter the information amounts here. Usually, these values are the sample sizes obtained up to the time of the analysis.

For example, you might enter *50, 76, 103, 150* to indicate that 50 individuals were included in the first interim analysis, 76 in the second, and so on.

Upper and Lower Boundaries (Spending = User)

If the Spending Function is set to *User Supplied* you can enter a set of lower test boundaries, one for each interim analysis. The lower boundaries should be negative and the upper boundaries should be positive. Typical entries are *4,3,3,3,2* and *4,3,2,2,2*.

- **Symmetric**

If you only want to enter the upper boundaries and have them copied with a change in sign to the lower boundaries, enter *Symmetric* for the lower boundaries.

Options Tab

The Options tab controls the convergence of the various iterative algorithms used in the calculations.

Maximum Iterations

Maximum Iterations (Lan-Demets algorithm)

This is the maximum number of iterations used in the Lan-DeMets algorithm during its search routine. We recommend a value of at least 200.

Tolerance

Probability Tolerance

During the calculation of the probabilities associated with a set of boundary values, probabilities less than this are assumed to be zero.

We suggest a value of 0.0000000001.

Power Tolerance

This is the convergence level for the search for the spending function values that achieve a certain power. Once the iteration changes are less than this amount, convergence is assumed. We suggest a value of 0.000001.

If the search is too time consuming, you might try increasing this value.

Alpha Tolerance

This is the convergence level for the search for a given alpha value. Once the changes in the computed alpha value are less than this amount, convergence is assumed and iterations stop. We suggest a value of 0.0001.

This option is only used when you are searching for alpha.

If the search is too time consuming, you may try increasing this value.

Group-Sequential Logrank Tests

Example 1 – Finding the Sample Size

A clinical trial is to be conducted over a two-year period to compare the hazard rate of a new treatment to that of the current treatment. The proportion surviving for two years using the current treatment is 0.3. The health community will be interested in the new treatment if the proportion surviving is increased to 0.45, a 50% increase. So that the sample size requirements for several survival proportions can be compared, it is also of interest to compute the sample size at response rates of 0.30, 0.35, 0.40, and 0.50. Assume the survival times are exponential.

Testing will be done at the 0.05 significance level and the power should be set to 0.90. A total of four tests are going to be performed on the data as they are obtained. The O'Brien-Fleming boundaries will be used.

Find the necessary sample sizes and test boundaries assuming equal sample sizes for each arm and two-sided hypothesis tests.

Setup

This section presents the values of each of the parameters needed to run this example. First, from the PASS Home window, load the **Group-Sequential Logrank Tests** procedure window by expanding **Survival**, then **Two Survival Curves**, then clicking on **Group-Sequential**, and then clicking on **Group-Sequential Logrank Tests**. You may then make the appropriate entries as listed below, or open **Example 1** by going to the **File** menu and choosing **Open Example Template**.

<u>Option</u>	<u>Value</u>
Design Tab	
Solve For	Sample Size
Alternative Hypothesis	Two-Sided
Survival Time Assumption	Exponential Survival
Power	0.90
Alpha	0.05
S1 (Proportion Surviving 1)	0.3
S2 (Proportion Surviving 2)	0.35 0.40 0.45 0.50
Number of Looks	4
Spending Function	O'Brien-Fleming
Boundary Truncation	None
Max Time	2
Times	Equally Spaced
Informations	<i>Blank</i>

Annotated Output

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

Numeric Results

Numeric Results for Two-Sided LogRank Test (Assuming Exponential Survival)

Power	Total Sample Size (N)	Total Required Events	Alpha	Beta	Proportion Surv. (S1)	Proportion Surv. (S2)	Hazard Ratio
0.900003	3378	2280	0.050000	0.099997	0.3000	0.3500	0.8720
0.900267	884	575	0.050000	0.099733	0.3000	0.4000	0.7611
0.900626	407	254	0.050000	0.099374	0.3000	0.4500	0.6632
0.900018	234	140	0.050000	0.099982	0.3000	0.5000	0.5757

Group-Sequential Logrank Tests

Report Definitions

Power is the probability of rejecting a false null hypothesis. Power should be close to one.

N is the number of items sampled from each group.

Events is the number of events that must occur in each group.

Alpha is the probability of rejecting a true null hypothesis in at least one of the sequential tests.

Beta is the probability of accepting a false null hypothesis at the conclusion of all tests.

S1 is the proportion surviving in group 1.

S2 is the proportion surviving in group 2.

HR is the hazard ratio. It is calculated using $\text{Log}(S2)/\text{Log}(S1)$.

Summary Statements

A total sample size of 3378 (split equally between the two groups), or 2280 events, achieves 90% power to detect a hazard rate of 0.8720 when the proportions surviving in each group are 0.3000 and 0.3500 at a significance level (alpha) of 0.050000 using a two-sided log rank test.

These results assume that 4 sequential tests are made using the O'Brien-Fleming spending function to determine the test boundaries and that the survival times are exponential.

This report shows the values of each of the parameters, one scenario per row. Note that 254 events are required when $S2 = 0.45$. Based on the expected survival proportions, this many events will occur if the overall sample size is 407.

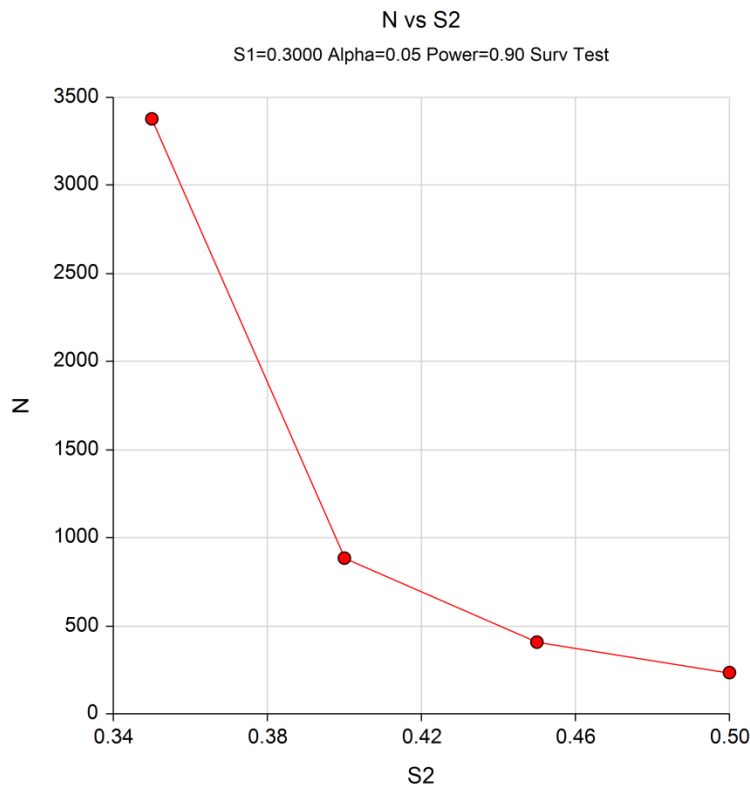
Total Sample Size (N)

This is the estimated sample size that is needed to obtain the necessary number of events.

Total Required Events

This is the number of events (deaths, etc.) that are required to achieve the desired power levels.

Plots Section



This plot shows that an increase in sample size from under 1000 to well over 3000 is necessary when the detectable proportion surviving is reduced from 0.4 to 0.35.

Group-Sequential Logrank Tests

Details Section

Details when Spending = O'Brien-Fleming, N = 407, d = 254, S1 = 0.3000, S2 = 0.4500								
Look	Time	Lower Bndry	Upper Bndry	Nominal Alpha	Inc Alpha	Total Alpha	Inc Power	Total Power
1	0.5000	-4.33263	4.33263	0.000015	0.000015	0.000015	0.003516	0.003516
2	1.0000	-2.96311	2.96311	0.003045	0.003036	0.003051	0.255183	0.258699
3	1.5000	-2.35902	2.35902	0.018323	0.016248	0.019299	0.427665	0.686364
4	2.0000	-2.01406	2.01406	0.044003	0.030701	0.050000	0.214262	0.900626
Drift	3.27466							

This report shows information about the individual interim tests. One report is generated for each scenario.

Look

These are the sequence numbers of the interim tests.

Time

These are the time points at which the interim tests are conducted. Since the Max Time was set to 2 (for two years), these time values are in years. Hence, the first interim test is at half a year, the second at one year, and so on.

We could have set Max Time to 24 so that the time scale was in months.

Lower and Upper Boundary

These are the test boundaries. If the computed value of the test statistic z is between these values, the trial should continue. Otherwise, the trial can be stopped.

Nominal Alpha

This is the value of alpha for these boundaries if they were used for a single, standalone, test. Hence, this is the significance level that must be found for this look in a standard statistical package that does not adjust for multiple looks.

Inc Alpha

This is the amount of alpha that is *spent* by this interim test. It is close to, but not equal to, the value of alpha that would be achieved if only a single test was conducted. For example, if we lookup the third value, 2.35902, in normal probability tables, we find that this corresponds to a (two-sided) alpha of 0.018323. However, the entry is 0.016248. The difference is due to the correction that must be made for multiple tests.

Total Alpha

This is the total amount of alpha that is used up to and including the current test.

Inc Power

These are the amounts that are added to the total power at each interim test. They are often called the exit probabilities because they give the probability that significance is found and the trial is stopped, given the alternative hypothesis.

Total Power

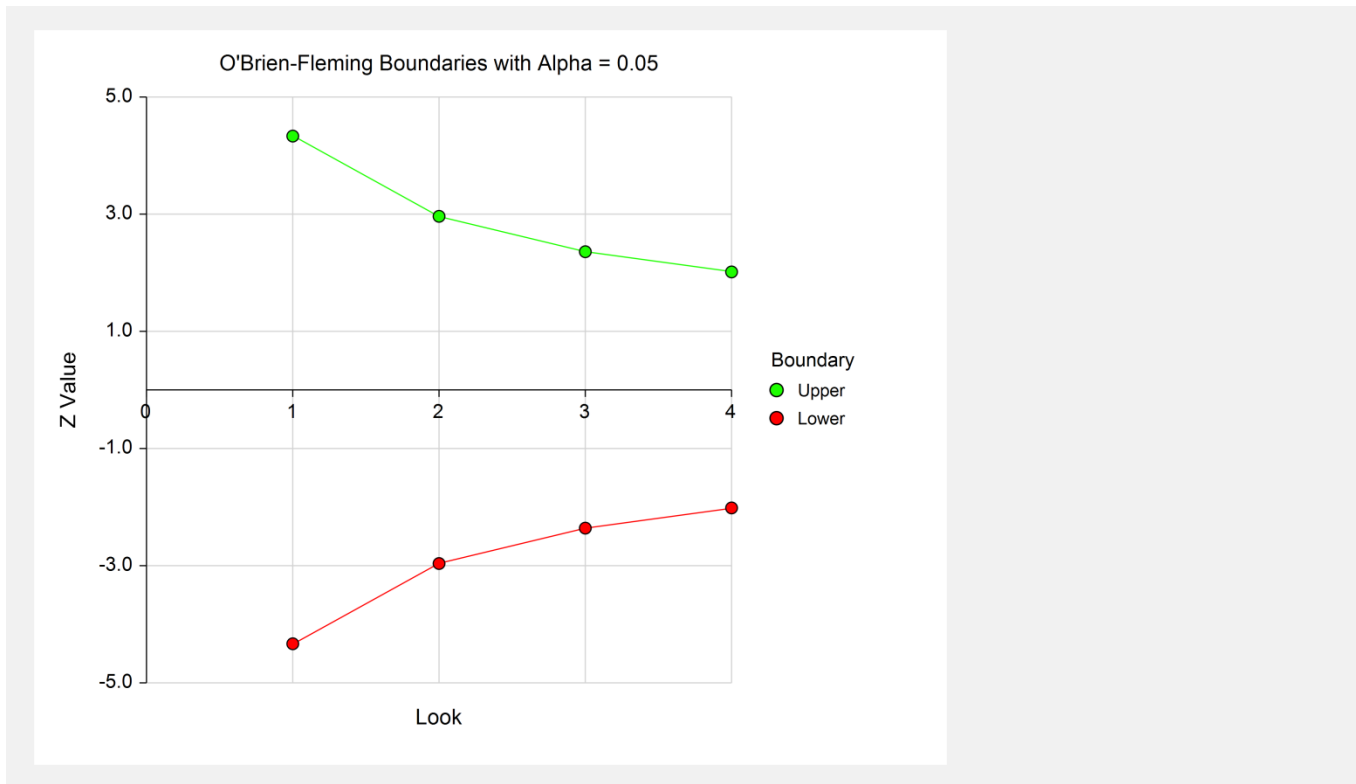
These are the cumulative power values. They are also the cumulative exit probabilities. That is, they are the probability that the trial is stopped at or before the corresponding time.

Drift

This is the value of the Brownian motion drift parameter.

Group-Sequential Logrank Tests

Boundary Plots



This plot shows the interim boundaries for each look. This plot shows very dramatically that the results must be extremely significant at early looks, but that they are near the single test boundary (1.96 and -1.96) at the last look.

Group-Sequential Logrank Tests

Example 2 – Finding the Power

Continuing the scenario began in Example 1, the researcher wishes to calculate the power of the design at sample sizes 50, 250, 450, 650, and 850. Testing will be done at the 0.01, 0.05, 0.10 significance levels and the overall power will be set to 0.10. Find the power of these sample sizes and test boundaries assuming equal sample sizes per arm and two-sided hypothesis tests.

Proceeding as in Example 1, we decide to translate the mean and standard deviation into a percent of mean scale.

Setup

This section presents the values of each of the parameters needed to run this example. First, from the PASS Home window, load the **Group-Sequential Logrank Tests** procedure window by expanding **Survival**, then **Two Survival Curves**, then clicking on **Group-Sequential**, and then clicking on **Group-Sequential Logrank Tests**. You may then make the appropriate entries as listed below, or open **Example 2** by going to the **File** menu and choosing **Open Example Template**.

<u>Option</u>	<u>Value</u>
Design Tab	
Solve For	Power
Alternative Hypothesis	Two-Sided
Survival Time Assumption	Exponential Survival
Alpha	0.01 0.05 0.10
N (Total Sample Size).....	50 to 850 by 200
S1 (Proportion Surviving 1)	0.3
S2 (Proportion Surviving 2)	0.45
Number of Looks	4
Spending Function.....	O'Brien-Fleming
Boundary Truncation	None
Max Time	2
Times	Equally Spaced
Informations	Blank

Output

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

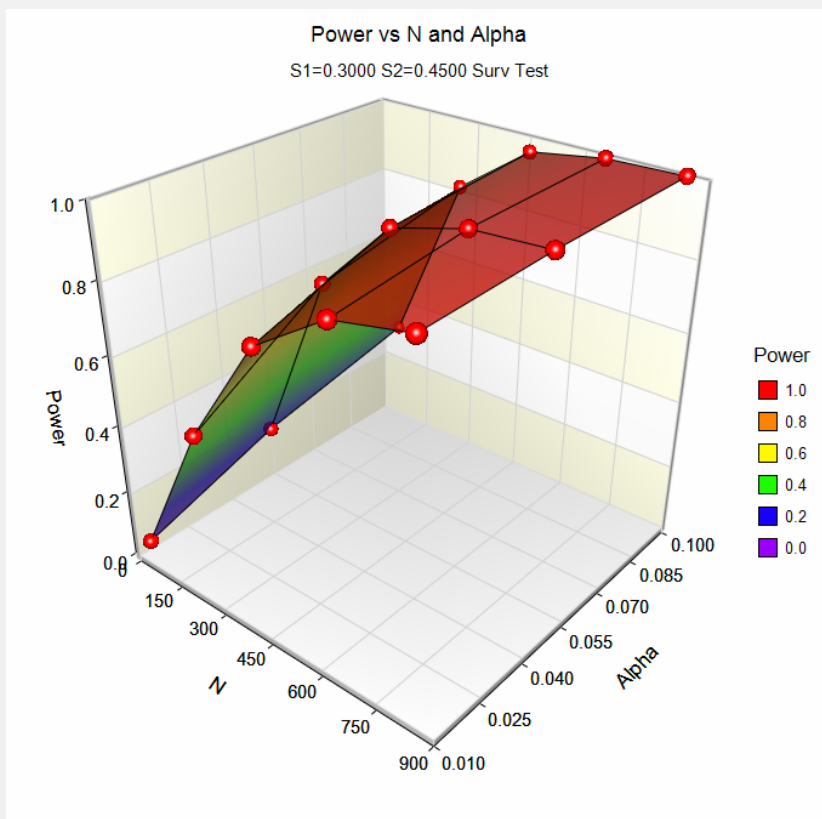
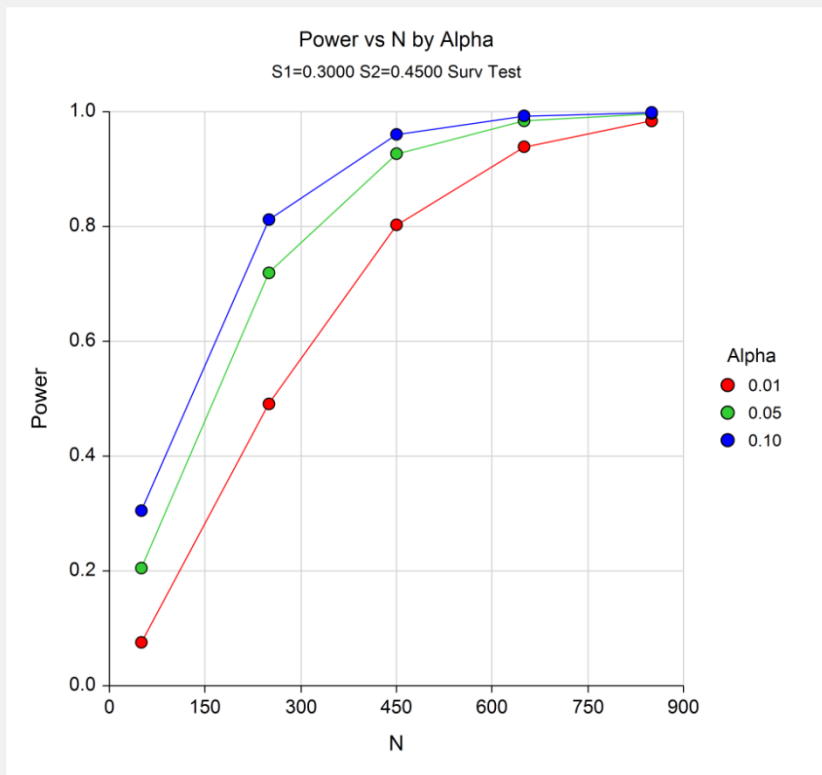
Numeric Results

Numeric Results for Two-Sided LogRank Test (Assuming Exponential Survival)

Power	Total Sample Size (N)	Total Required Events	Alpha	Beta	Proportion Surv. (S1)	Proportion Surv. (S2)	Hazard Ratio
0.075632	50	31	0.010000	0.924368	0.3000	0.4500	0.6632
0.491011	250	156	0.010000	0.508989	0.3000	0.4500	0.6632
0.802921	450	281	0.010000	0.197079	0.3000	0.4500	0.6632
0.938923	650	406	0.010000	0.061077	0.3000	0.4500	0.6632
0.983780	850	531	0.010000	0.016220	0.3000	0.4500	0.6632
0.205032	50	31	0.050000	0.794968	0.3000	0.4500	0.6632
0.719225	250	156	0.050000	0.280775	0.3000	0.4500	0.6632
0.926894	450	281	0.050000	0.073106	0.3000	0.4500	0.6632

(report continues)

Group-Sequential Logrank Tests



These data show the power for various sample sizes and alphas. It is interesting to note that once the sample size is greater than about 450, the value of alpha has comparatively little difference on the value of power.

Group-Sequential Logrank Tests

Example 3 – Effect of Number of Looks

Continuing with examples one and two, it is interesting to determine the impact of the number of looks on power. *PASS* allows only one value for the Number of Looks parameter per run, so it will be necessary to run several analyses. To conduct this study, set alpha to 0.05, N to 407, and leave the other parameters as before. Run the analysis with Number of Looks equal to 1, 2, 3, 4, 6, 8, 10, and 20. Record the power for each run.

Setup

This section presents the values of each of the parameters needed to run this example. First, from the *PASS* Home window, load the **Group-Sequential Logrank Tests** procedure window by expanding **Survival**, then **Two Survival Curves**, then clicking on **Group-Sequential**, and then clicking on **Group-Sequential Logrank Tests**. You may then make the appropriate entries as listed below, or open **Example 3** by going to the **File** menu and choosing **Open Example Template**.

<u>Option</u>	<u>Value</u>
Design Tab	
Solve For	Power
Alternative Hypothesis	Two-Sided
Survival Time Assumption	Exponential Survival
Alpha	0.05
N (Total Sample Size).....	407
S1 (Proportion Surviving 1)	0.3
S2 (Proportion Surviving 2)	0.45
Number of Looks	1 (Also run with 2, 3, 4, 6, 8, 10, and 20)
Spending Function.....	O'Brien-Fleming
Boundary Truncation	None
Max Time	2
Times	Equally Spaced
Informations	<i>Blank</i>

Output

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

Numeric Results

Numeric Results for Two-Sided Logrank Test (Assuming Exponential Survival)

Power	N	Events	Alpha	Beta	S1	S2	Looks
0.905693	407	254	0.050000	0.094307	0.3000	0.4500	1
0.904719	407	254	0.050000	0.095281	0.3000	0.4500	2
0.902412	407	254	0.050000	0.097588	0.3000	0.4500	3
0.900626	407	254	0.050000	0.099374	0.3000	0.4500	4
0.898243	407	254	0.050000	0.101757	0.3000	0.4500	6
0.896763	407	254	0.050000	0.103237	0.3000	0.4500	8
0.895758	407	254	0.050001	0.104242	0.3000	0.4500	10
0.893398	407	254	0.050001	0.106602	0.3000	0.4500	20

This analysis shows how little the number of looks impacts the power of the design. The power of a study with no interim looks is 0.905693. When twenty interim looks are made, the power falls to 0.893398—a very small change.

Example 4 – Studying a Boundary Set

Continuing with the previous examples, suppose that you are presented with a set of boundaries and want to find the quality of the design (as measured by alpha and power). This is easy to do with *PASS*. Suppose that the analysis is to be run with five interim looks at equally spaced time points. The upper boundaries to be studied are 3.5, 3.5, 3.0, 2.5, 2.0. The lower boundaries are symmetric. The analysis would be run as follows.

Setup

This section presents the values of each of the parameters needed to run this example. First, from the *PASS* Home window, load the **Group-Sequential Logrank Tests** procedure window by expanding **Survival**, then **Two Survival Curves**, then clicking on **Group-Sequential**, and then clicking on **Group-Sequential Logrank Tests**. You may then make the appropriate entries as listed below, or open **Example 4** by going to the **File** menu and choosing **Open Example Template**.

<u>Option</u>	<u>Value</u>
Design Tab	
Solve For	Power
Alternative Hypothesis	Two-Sided
Survival Time Assumption	Exponential Survival
Alpha	0.05 (will be calculated from boundaries)
N (Total Sample Size).....	407
S1 (Proportion Surviving 1)	0.30
S2 (Proportion Surviving 2)	0.45
Number of Looks	5
Spending Function.....	User Supplied
Boundary Truncation	None
Max Time	2
Times	Equally Spaced
Informations	<i>Blank</i>
Upper Boundaries.....	3.5 3.5 3.0 2.5 2.0
Lower Boundaries.....	Symmetric

Output

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

Numeric Results

Numeric Results for Two-Sided Logrank Test (Assuming Exponential Survival)								
Power	Total Sample Size (N)	Total Required Events	Alpha	Beta	Proportion Surv. (S1)	Proportion Surv. (S2)	Hazard Ratio	
0.900747	407	254	0.048157	0.099253	0.3000	0.4500	0.6632	
Details when Spending = User Supplied, N = 407, d=254, S1 = 0.3000, S2 = 0.4500								
Look	Time	Lower Bndry	Upper Bndry	Nominal Alpha	Inc Alpha	Total Alpha	Inc Power	Total Power
1	0.4000	-3.50000	3.50000	0.000465	0.000465	0.000465	0.020899	0.020899
2	0.8000	-3.50000	3.50000	0.000465	0.000408	0.000874	0.063132	0.084031
3	1.2000	-3.00000	3.00000	0.002700	0.002410	0.003284	0.243522	0.327553
4	1.6000	-2.50000	2.50000	0.012419	0.010331	0.013615	0.343072	0.670625
5	2.0000	-2.00000	2.00000	0.045500	0.034542	0.048157	0.230122	0.900747
Drift	3.27466							

Group-Sequential Logrank Tests

The power for this design is about 0.90. This value depends on both the boundaries and the sample size. The alpha level is about 0.048. This value only depends on the boundaries.

Example 5 – Validation using O’Brien-Fleming Boundaries

Reboussin (1992) presents an example for binomial distributed data for a design with two-sided O’Brien-Fleming boundaries, looks = 3, alpha = 0.05, beta = 0.10, $S1 = 0.30$, $S2 = 0.786$ (which gives a hazard ratio of 0.20). They compute a drift of 3.261 and the number of events at 16.42. The upper boundaries are: 4.8769, 3.3569, 2.6803, 2.2898, 2.0310.

To test that *PASS* provides the same result, enter the following.

Setup

This section presents the values of each of the parameters needed to run this example. First, from the PASS Home window, load the **Group-Sequential Logrank Tests** procedure window by expanding **Survival**, then **Two Survival Curves**, then clicking on **Group-Sequential**, and then clicking on **Group-Sequential Logrank Tests**. You may then make the appropriate entries as listed below, or open **Example 5** by going to the **File** menu and choosing **Open Example Template**.

<u>Option</u>	<u>Value</u>
Design Tab	
Solve For	Sample Size
Alternative Hypothesis	Two-Sided
Survival Time Assumption	Exponential Survival
Power	0.90
Alpha	0.05
S1 (Proportion Surviving 1)	0.30
S2 (Proportion Surviving 2)	0.786
Number of Looks	5
Spending Function	O’Brien-Fleming
Boundary Truncation	None
Max Time	1
Times	Equally Spaced
Informations	<i>Blank</i>

Output

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

Numeric Results

Numeric Results for Two-Sided Logrank Test (Assuming Exponential Survival)

Power	Total Sample Size (N)	Total Required Events	Alpha	Beta	Proportion Surv. (S1)	Proportion Surv. (S2)	Hazard Ratio
0.905173	37	17	0.050000	0.094827	0.3000	0.7860	0.2000

Group-Sequential Logrank Tests

Details when Spending = O'Brien-Fleming, N = 37, d = 17, S1 = 0.3000, S2 = 0.7860

Look	Time	Lower Bndry	Upper Bndry	Nominal Alpha	Inc Alpha	Total Alpha	Inc Power	Total Power
1	0.2000	-4.87688	4.87688	0.000001	0.000001	0.000001	0.000341	0.000341
2	0.4000	-3.35695	3.35695	0.000788	0.000787	0.000788	0.102760	0.103101
3	0.6000	-2.68026	2.68026	0.007357	0.006828	0.007616	0.352450	0.455551
4	0.8000	-2.28979	2.28979	0.022034	0.016807	0.024424	0.298953	0.754504
5	1.0000	-2.03100	2.03100	0.042255	0.025576	0.050000	0.150669	0.905173
Drift		3.30902						

The number of events, rounded to 17, matches the 16.42 reported in Reboussin (1992).