

## Chapter 700

# Logrank Tests (Freedman)

---

### Introduction

This module allows the sample size and power of the logrank test to be analyzed under the assumption of proportional hazards. Time periods are not stated. Rather, it is assumed that enough time elapses to allow for a reasonable proportion of responses to occur. If you want to study the impact of accrual and follow-up time, you should use the one of the other logrank modules also contained in **PASS**. The formulas used in this module come from Machin *et al.* (1997). They are also given in Fayers and Machin (2016) where they are applied to sizing quality of life studies.

A clinical trial is often employed to test the equality of survival distributions for two treatment groups. For example, a researcher might wish to determine if Beta-Blocker A enhances the survival of newly diagnosed myocardial infarction patients over that of the standard Beta-Blocker B. The question being considered is whether the pattern of survival is different.

The two-sample t-test is not appropriate for two reasons. First, the data consist of the length of survival (time to failure), which is often highly skewed, so the usual normality assumption cannot be validated. Second, since the purpose of the treatment is to increase survival time, it is likely (and desirable) that some of the individuals in the study will survive longer than the planned duration of the study. The survival times of these individuals are then said to be *censored*. These times provide valuable information, but they are not the actual survival times. Hence, special methods have to be employed which use both regular and censored survival times.

The logrank test is one of the most popular tests for comparing two survival distributions. It is easy to apply and is usually more powerful than an analysis based simply on proportions. It compares survival across the whole spectrum of time, not at just one or two points.

The power calculations used here assume an underlying exponential distribution. However, we are rarely in a position to assume exponential survival times in an actual clinical trial. How do we justify the exponential survival time assumption? First, the logrank test and the test derived using the exponential distribution have nearly the same power when the data are in fact exponentially distributed. Second, under the proportional hazards model (which is assumed by the logrank test), the survival distribution can be transformed to be exponential and the logrank test remains the same under monotonic transformations.

---

### Technical Details

In order to define the input parameters, we will present below some rather complicated looking formulas. You need not understand the formulas. However, you should understand the individual parameters used in these formulas.

We assume that a study is to be made comparing the survival of an existing (control) group with an experimental group. The control group consists of patients that will receive the existing treatment. In cases where no existing treatment exists, the control group consists of patients that will receive a placebo. This group is arbitrarily called group one. The experimental group will receive the new treatment. It is called group two.

## Logrank Tests (Freedman)

We assume that the critical event of interest is death and that two treatments have survival distributions with instantaneous death (hazard) rates,  $\lambda_1$  and  $\lambda_2$ . These hazard rates are an subject's probability of death in a short period of time. We want to test hypotheses about these hazard rates.

---

### Hazard Ratio

There are several ways to express the difference between two hazard rates. One way is to calculate the difference,  $\lambda_1 - \lambda_2$ . Another way is to form the hazard ratio (*HR*):

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

Note that since *HR* is formed by dividing the hazard rate of the experimental group by that of the control group, a treatment that does better than the control will have a hazard ratio that is less than one.

The hazard ratio may be formulated in other ways. If the proportions surviving during the study are called *S1* and *S2* for the control and experimental groups, the hazard ratio is given by

$$HR = \frac{\text{Log}(S2)}{\text{Log}(S1)}.$$

Furthermore, if the median survival times of the two groups are *M1* and *M2*, the hazard ratio is given by

$$HR = \frac{M1}{M2}.$$

Each of these expressions for the difference between hazards rates is useful in some situations. There is no one best way, so you will have to be a little flexible.

---

### Logrank Test

We assume that the logrank test will be used to analyze the data once they are collected. Basing the power calculations on the logrank test, we arrive at the following formula that gives the power based on several other parameters:

$$z_{1-\beta} = \frac{|HR - 1| \sqrt{N(1-w)\varphi[(1-S1) + \varphi(1-S2)] / (1+\varphi)}}{(1+\varphi HR)} - z_{1-\alpha/k}$$

where *k* is 1 for a one-sided hypothesis test or 2 for a two-sided test,  $\alpha$  and  $\beta$  are the error rates defined as usual, the *z*'s are the usual points from the standard normal distribution, *w* is the proportion that are lost to follow up, and  $\varphi$  represents the sample size ratio between the two groups. That is, *p1* is the proportion of the total sample size that is in the control group,  $\varphi$  is given by

$$\varphi = \frac{1 - p1}{p1}$$

Note that the hypothesis being tested is that the hazard rates are equal:

$$H_0: \lambda_1 = \lambda_2$$

This formulation assumes that the hazard rates are proportional. It does not assume that the failure times are exponentially distributed.

---

## Procedure Options

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

---

### Design Tab

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

---

#### Solve For

##### Solve For

This option specifies the parameter to be solved for from the other parameters. The parameters that may be selected are *Proportion Surviving 1*, *Proportion Surviving 2*, *Sample Size*, *Alpha*, or *Power*. 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* when you want to calculate the power of an experiment that has already been run.

If you select Survive 1, the search is made for values that are less than  $S_2$ . Likewise, if you select Survive 2, the search is made for values that are greater than  $S_1$ .

---

#### Test

##### One-Sided Test

Specify whether the test is one-sided (checked) or two-sided (unchecked). When a two-sided test is selected, the value of alpha is divided by two.

---

#### 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*.

## Logrank Tests (Freedman)

### Sample Size – Total Sample Size

#### N (Total Sample Size)

The combined sample size of both groups.

### Sample Size – Sample Proportions

#### Proportion in Group 1

The proportion of the sample size ( $N$ ) that is assigned to the first group. The rest of the patients are assigned to the second group. When exact values are not available, the number in group one is found by rounding up to the next integer. The number in group two is found by subtraction.

#### Proportion Lost During Follow Up

This is the proportion of patients in the trial that are lost to follow-up. Since they are lost, you can never obtain their outcome. The sample size is inflated to account for this loss using the formula:

$$N(\text{adjusted}) = N/(1-p)$$

where  $p$  is the proportion lost.

### Effect Size

#### S1 (Proportion Surviving in Group 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. A value for  $S1$  must be determined either from a pilot study or from previous studies.

#### S2 (Proportion Surviving in Group 2)

$S2$  is the proportion of patients belonging to group 2 (experimental) that are expected to survive during the study. Since  $S2$  is a proportion, it must be between zero and one.

This value is not necessarily the expected survival proportion under the treatment. Rather, you may set it to that value that, if achieved, would be of special interest. Values below this amount would not be of interest.

For example, if the standard 1-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 using it.

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

$$HR = \text{Log}(S2)/\text{Log}(S1)$$

we can solve for  $S2$  to obtain

$$S2 = \text{Exp}(\text{Log}(S1)HR).$$

Using this equation, you can determine an appropriate value for  $S2$  from a value for  $HR$ .

For example, suppose  $S1$  is 0.4 and you have determined that if the hazard rate is reduced by half, the improvement is sufficient to justify a change in treatment. The appropriate value for  $S2$  is  $\text{Exp}(\text{Log}(0.4)(0.5)) = 0.632$ .

When you do not have an anticipated value for  $HR$ , you can find an estimate base on the Median Survival Times ( $M1$  and  $M2$ ) since  $HR = M1/M2$ .

## Logrank Tests (Freedman)

**Example 1 – Finding the Power**

An experiment has been conducted to test the effectiveness of a new treatment on a total of 100 patients. The current treatment for this disease achieves 50% survival after two years. The proportion in the treatment group that survived two years was 0.70. Testing was done at the 0.05 significance level. Even though there was an increase of 0.20 for 0.50 to 0.70, the logrank test did not reject the hypothesis of equal hazard rates for the two treatments. Study the power of this test.

**Setup**

This section presents the values of each of the parameters needed to run this example. First, from the PASS Home window, load the procedure window. 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>
<b>Design Tab</b>	
Solve For .....	<b>Power</b>
One-Sided Test.....	<b>Not Checked</b>
Alpha.....	<b>0.05 0.10</b>
N (Total Sample Size).....	<b>50 to 300 by 50</b>
Proportion in Group 1 .....	<b>0.5</b>
Proportion Lost During Follow Up.....	<b>0</b>
S1 (Proportion Surviving in Group 1).....	<b>0.5</b>
S2 (Proportion Surviving in Group 2).....	<b>0.7</b>

**Annotated Output**

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

**Numeric Results**

Numeric Results									
<b>Power</b>	<b>N</b>	<b>N1</b>	<b>N2</b>	<b>S1</b>	<b>S2</b>	<b>Hazard Ratio</b>	<b>Two-Sided Alpha</b>	<b>Beta</b>	
0.2992	50	25	25	0.5000	0.7000	0.5000	0.0500	0.7008	
0.4162	50	25	25	0.5000	0.7000	0.5000	0.1000	0.5838	
0.5267	100	50	50	0.5000	0.7000	0.5000	0.0500	0.4733	
0.6488	100	50	50	0.5000	0.7000	0.5000	0.1000	0.3512	
0.6994	150	75	75	0.5000	0.7000	0.5000	0.0500	0.3006	
0.7989	150	75	75	0.5000	0.7000	0.5000	0.1000	0.2011	
0.8177	200	100	100	0.5000	0.7000	0.5000	0.0500	0.1823	
0.8891	200	100	100	0.5000	0.7000	0.5000	0.1000	0.1109	
0.8934	250	125	125	0.5000	0.7000	0.5000	0.0500	0.1066	
0.9406	250	125	125	0.5000	0.7000	0.5000	0.1000	0.0594	
0.9395	300	150	150	0.5000	0.7000	0.5000	0.0500	0.0605	
0.9690	300	150	150	0.5000	0.7000	0.5000	0.1000	0.0310	

## Logrank Tests (Freedman)

## Report Definitions

Power is the probability of rejecting a false null hypothesis.

N is the combined sample size.

N1 sample size in group 1.

N2 sample size in group 2.

S1 is the proportion surviving in group 1

S2 is the proportion surviving in group 2

The Hazard Ratio is the ratio of hazard2 and hazard1. It is  $\text{Log}(S2)/\text{Log}(S1)$ .

Alpha is the probability of rejecting a true null hypothesis.

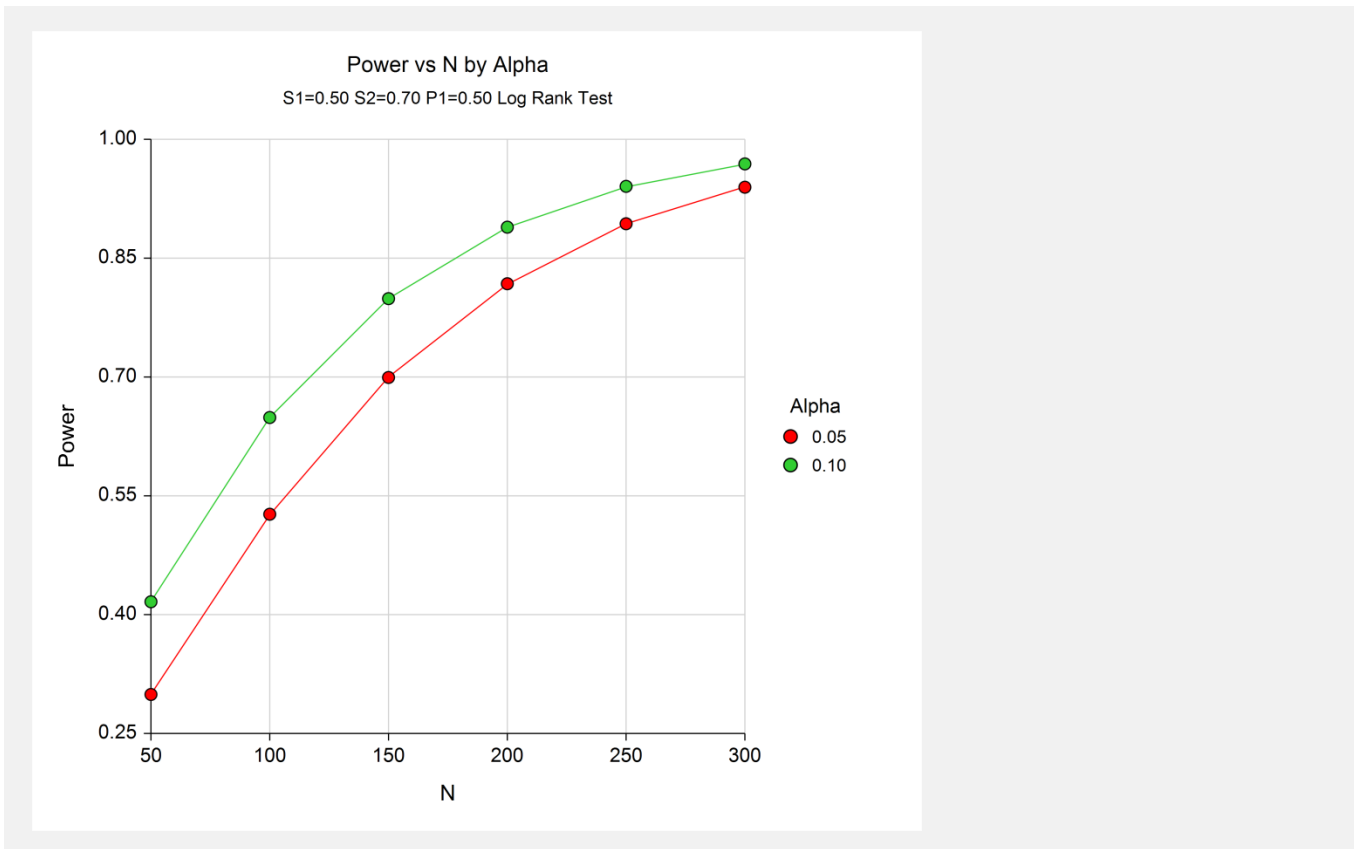
Beta is the probability of accepting a false null hypothesis.

## Summary Statements

A two-sided log rank test with an overall sample size of 50 subjects (of which 25 are in group 1 and 25 are in group 2) achieves 30% power at a 0.0500 significance level to detect a difference of 0.2000 between 0.5000 and 0.7000--the proportions surviving in groups 1 and 2, respectively. This corresponds to a hazard ratio of 0.5000. The proportion of patients lost during follow up was 0.0000. These results are based on the assumption that the hazard rates are proportional.

This report shows the values of each of the parameters, one scenario per row. The values from this table are in the chart below.

## Plots Section



This plot shows the relationship between alpha and power in this example. We notice that for alpha = 0.05, a power of 0.80 is reached when the sample size is about 200. A power of 90% is reached when the sample size is 250. Hence, we realize that this study should have had at least twice the number of patients that it had.

## Example 2 – Finding the Sample Size

Continuing with our example, the researcher decides that he wants to do it right this time. He believes that if the survival proportion of the treatment group is 0.6 or better, people will begin to use his treatment. He wants to know how many subjects he needs.

### Setup

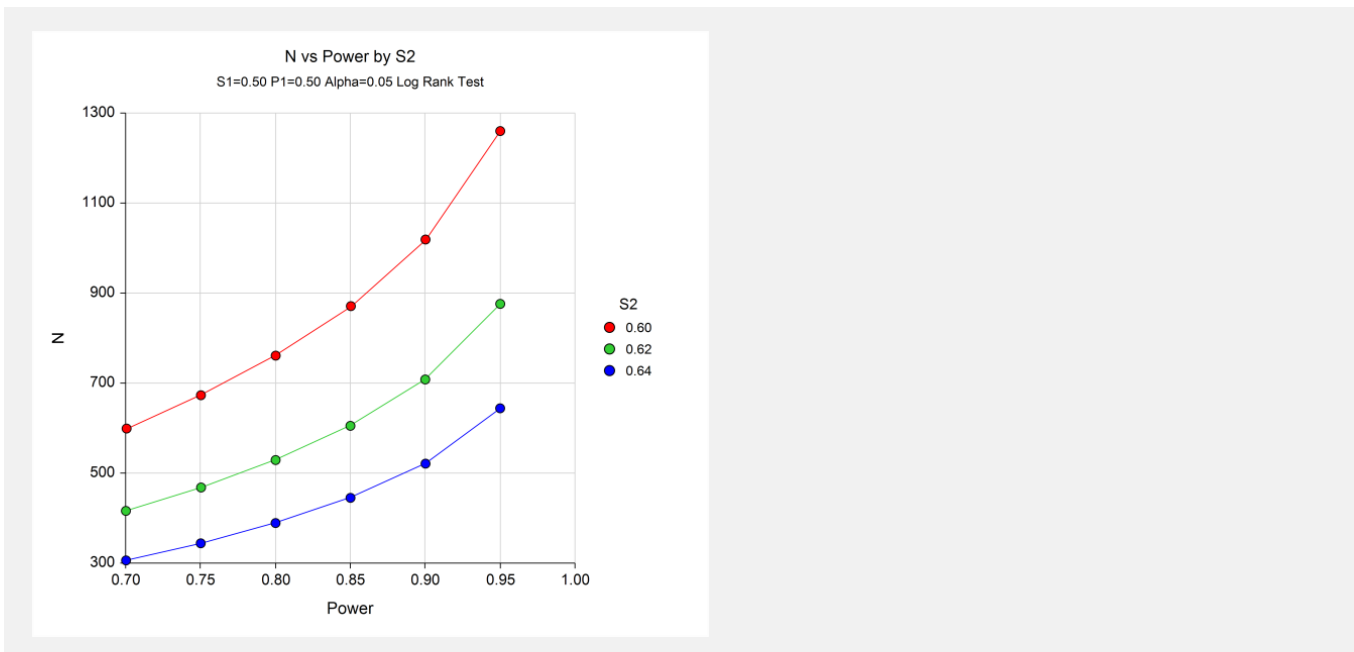
This section presents the values of each of the parameters needed to run this example. First, from the PASS Home window, load the procedure window. You may then make the appropriate entries as listed below, or open **Example 2** by going to the **File** menu and choosing **Open Example Template**.

<u>Option</u>	<u>Value</u>
<b>Design Tab</b>	
Solve For .....	<b>Sample Size</b>
One-Sided Test.....	<b>Not Checked</b>
Power.....	<b>0.70 to 0.95 by 0.05</b>
Alpha.....	<b>0.05</b>
Proportion in Group 1 .....	<b>0.5</b>
Proportion Lost During Follow Up.....	<b>0</b>
S1 (Proportion Surviving in Group 1).....	<b>0.5</b>
S2 (Proportion Surviving in Group 2).....	<b>0.60 0.62 0.64</b>

### Output

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

#### Plots Section



We will consider the chart since that allows us to understand the patterns more quickly. We note that changing  $S_2$  from 0.60 (the top line) to 0.62 (the middle line) decreases the sample size requirements by almost half. Our researcher decides to take a sample of 500 patients. This will achieve almost 80% power in detecting a shift from 0.50 to 0.62 in survivability.

## Example 3 – Validation using Machin (1997)

Machin *et al.* (1997) page 180 gives an example in which  $S1$  is 0.25,  $S2$  is 0.50, the one-sided significance level is 0.05, and the power is 90%. The total sample size is 124. We will now run this example through PASS.

### Setup

This section presents the values of each of the parameters needed to run this example. First, from the PASS Home window, load the procedure window. 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>
<b>Design Tab</b>	
Solve For .....	<b>Sample Size</b>
One-Sided Test.....	<b>Checked</b>
Power.....	<b>0.90</b>
Alpha.....	<b>0.05</b>
Proportion in Group 1 .....	<b>0.5</b>
Proportion Lost During Follow Up.....	<b>0</b>
S1 (Proportion Surviving in Group 1).....	<b>0.25</b>
S2 (Proportion Surviving in Group 2).....	<b>0.50</b>

### Output

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

#### Numeric Results

Numeric Results									
Power	N	N1	N2	S1	S2	Hazard Ratio	One-Sided Alpha	Beta	
0.9014	124	62	62	0.2500	0.5000	0.5000	0.0500	0.0986	

PASS also calculated the total sample size to be 124.



## Example 4 – Validation based on a HADS Study using Fayers and Machin (2016)

Fayers and Machin (2016) give an example of the use of this procedure in conjunction with a Hospital Anxiety and Depression Scale (HADS) study. The HADS survey provides an ordinal score which has 22 categories. The HADS survey is administered soon after a subject is admitted into the trial. The subject is followed until they improve two steps on the HADS survey. The response is the number of days that were needed to achieve the two-category improvement (score decrease). This outcome is elapsed time and it can be analyzed with Kaplan-Meier curves and logrank tests such as those which are shown in this procedure.

The example on page 297 of Fayers and Machin (2016) is of a two-group study that compares the response to a new treatment (group 2) to that of the standard treatment (group 1). The hazard ratio that is to be detected is represented by two proportions of those that improved two-steps by the end of 12 weeks. These proportions are 0.65 in group 1 and 0.75 in group 2. The two-sided significance level is 0.05 and the required power is 80%. The estimated sample size is 660.

### Setup

This section presents the values of each of the parameters needed to run this example. First, from the PASS Home window, load the procedure window. 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>
<b>Design Tab</b>	
Solve For .....	<b>Sample Size</b>
One-Sided Test.....	<b>Unchecked</b>
Power.....	<b>0.80</b>
Alpha.....	<b>0.05</b>
Proportion in Group 1 .....	<b>0.5</b>
Proportion Lost During Follow Up.....	<b>0</b>
S1 (Proportion Surviving in Group 1).....	<b>0.65</b>
S2 (Proportion Surviving in Group 2).....	<b>0.75</b>

### Output

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

#### Numeric Results

Numeric Results									
Power	N	N1	N2	S1	S2	Hazard Ratio	Two-Sided Alpha	Beta	
0.8003	660	330	330	0.65	0.75	0.66781	0.05	0.1997	
Event Report									
Power	E	E1	E2	S1	S2	Hazard Ratio	Two-Sided Alpha	Beta	
0.8003	198	99	99	0.65	0.75	0.66781	0.05	0.1997	

**PASS** calculated the total sample size to be 660. Note that 99 events per group are required. The sample size of 660 estimates how many subjects must be enrolled to achieve the required number of events.

## Example 5 – Calculating Sample Size for a COVID-19 Clinical Trial

This example will show how this procedure might be used in planning a clinical trial to compare the effectiveness of a new treatment with the standard treatment in combatting COVID-19. The study outcome in each group will be the number of days it takes patients to advance two categories on a multi-point ordinal scale of disease status. The hypothetical trial being planned here will use the following six-point ordinal scale.

- 0) Discharge (alive).
- 1) Hospital admission, not requiring supplemental oxygen.
- 2) Hospital admission, requiring supplemental oxygen.
- 3) Hospital admission, requiring high-flow nasal cannula or non-invasive mechanical ventilation.
- 4) Hospital admission, requiring extracorporeal membrane oxygenation or invasive mechanical ventilation.
- 5) Death.

Each patient in the study is rated on this scale every day, and the response is the number of days until the disease status improves two categories. Suppose the hazard ratio that is to be detected by the study is the proportion is represented by the two proportions improving two-steps by the end of 14 days. These proportions are 0.40 in group 1 (control) and 0.55 in group 2 (treatment). The two-sided significance level is 0.05 and the required power is 0.80. The goal is to estimate the necessary sample size.

### Setup

This section presents the values of each of the parameters needed to run this example. First, from the PASS Home window, load the procedure window. 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>
<b>Design Tab</b>	
Solve For .....	<b>Sample Size</b>
One-Sided Test.....	<b>Unchecked</b>
Power.....	<b>0.80</b>
Alpha.....	<b>0.05</b>
Proportion in Group 1 .....	<b>0.5</b>
Proportion Lost During Follow Up.....	<b>0</b>
S1 (Proportion Surviving in Group 1).....	<b>0.40</b>
S2 (Proportion Surviving in Group 2).....	<b>0.55</b>

### Output

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

Numeric Results									
Power	N	N1	N2	S1	S2	Hazard Ratio	Two-Sided Alpha	Beta	
0.80003	338	169	169	0.4	0.55	0.65245	0.05	0.19997	
Event Report									
Power	E	E1	E2	S1	S2	Hazard Ratio	Two-Sided Alpha	Beta	
0.80003	178	89	89	0.4	0.55	0.65245	0.05	0.19997	

PASS calculated the total sample size to be 338. Note that 89 events per group are required. The sample size of 338 estimates how many subjects must be enrolled to achieve the 178 events.