Chapter 706

Non-Inferiority Logrank Tests

Introduction

This module computes the sample size and power for *non-inferiority* tests under the assumption of proportional hazards. Accrual time and follow-up time are included among the parameters to be set. The non-inferiority logrank test is used for data analysis.

Sometimes, the objective of a study is to show that an experimental therapy is not inferior to (no worse than) the standard therapy. The experimental therapy may be cheaper, less toxic, or have fewer side effects. Such studies are often called non-inferiority trials and have a one-sided hypothesis.

Power and sample size calculations for the non-inferiority logrank test have been developed by Jung et al. (2005), and we use their results. These calculations assume an underlying exponential survival distribution with a uniform patient accrual pattern during the accrual period.

Technical Details

Test Statistic

Suppose a clinical trial consists of two independent groups. Designate group one as the standard group with hazard rate h_1 and sample size n_1. Designate group two as the experimental group with hazard rate h_2 and sample size n_2. The total sample size is $N = n_1 + n_2$. Usually, you would plan to have $n_1 = n_2$.

Define the proportion of the total sample in each group as

$$Q_i = \frac{n_i}{N}, \quad i = 1, 2$$

Individuals are recruited during an accrual period of *R* years (or months or days). They are followed for an additional period of time until a total of *T* years is reached. Hence, the follow-up period is *T-R* years. At the end of the study, the non-inferiority logrank test is conducted at significance level α with power $1 - \beta$. Under the proportion hazards assumption, the hazard ratio $HR = h_2 / h_1$ is constant across time.

For a given non-inferiority margin HR_0 (>1) (the maximum ratio of clinical insignificance), the statistical hypotheses tested are

$$H_0\colon\ HR \geq HR_0 \quad \text{vs.} \quad H_1\colon HR < HR_0$$

Define the partial score function as

$$W(HR) = HR\sum_{i=1}^{n_1}\frac{\delta_{1i}\sum_{j=1}^{n_2} I(X_{2j} \geq X_{1i})}{\sum_{j=1}^{n_1} I(X_{1j} \geq X_{1i}) + HR\sum_{j=1}^{n_2} I(X_{2j} \geq X_{1i})}$$

$$-\sum_{i=1}^{n_2}\frac{\delta_{2i}\sum_{j=1}^{n_2} I(X_{1j} \geq X_{2i})}{\sum_{j=1}^{n_1} I(X_{1j} \geq X_{2i}) + HR\sum_{j=1}^{n_2} I(X_{2j} \geq X_{2i})}$$

and the information function as

$$\sigma_N^2(HR) = HR\sum_{k=1}^{2}\sum_{i=1}^{n_k}\frac{\delta_{ki}\left\{\sum_{j=1}^{n_1} I(X_{1j} \geq X_{ki})\right\}\left\{\sum_{j=1}^{n_2} I(X_{2j} \geq X_{ki})\right\}}{\left\{\sum_{j=1}^{n_1} I(X_{1j} \geq X_{ki}) + HR\sum_{j=1}^{n_2} I(X_{2j} \geq X_{ki})\right\}^2}$$

where X_{ki} is the minimum of the survival time, and the censoring time, δ_{ki}, is an event indicator taking 1 if there was an event or 0 otherwise, and $I(.)$ is an indicator function. Note that *W*(1) is the standard logrank test statistic.

Under H_0, $W(HR_0)/\sigma_N(HR_0)$ is asymptotically normal with mean 0 and variance 1. Reject H_0 in favor of H_1 if $W(HR_0)/\sigma_N(HR_0) > z_{1-\alpha}$ with one-sided type I error probability α.

The partial MLE, $\widehat{HR}$, is obtained by solving $W(HR) = 0$. Let HR^* denote the true value of *HR*. It can be shown that $\widehat{HR}$ is asymptotically normal with mean HR^* and variance $\sigma_N^{-2}(HR^*)$.

An asymptotic $100(1-\alpha)\%$ confidence interval for *HR* is

$$\widehat{HR} \pm z_{1-\alpha/2}\sigma_N^{-1}(\widehat{HR}).$$

Power Calculations

Jung (2005) shows that the power of the non-inferiority logrank test can be expressed as

$$1-\beta = \Phi\left(\frac{(HR_0 - 1)DQ_1Q_2 - z_{1-\alpha}\sqrt{HR_0}}{Q_1 + Q_2HR_0}\right)$$

where *D* is the observed number of deaths (events). The total sample size *N* is obtained by inflating *D* according to the relationship E(*d*)*N* = *D*, where E(*d*) is the expected death rate for the trial.

Following the proposal of Yateman and Skene (1992) and the results of Lakatos (1988), we compute E(*d*) using the Markov Model given in chapter 715 as

$$E(d) = Q_1S_{1,2} + Q_2S_{2,2}$$

where $S_{1,2}$ and $S_{2,2}$ are the occupancy probabilities for the event state for the standard and experimental groups, respectively. This formulation allows the inclusion of loss to follow-up, noncompliance, and drop-in along with various accrual patterns.

Example 1 – Finding the Power

A non-inferiority trial is planned in which the primary analysis will use the non-inferiority logrank test. After extensive discussion, the researchers have decided that the upper bound on non-inferiority is 1.3.

The trial will include a recruitment period of two-years after which participants will be followed for three more years. It is assumed that patients will enter the study uniformly over the accrual period. The researcher estimates a loss-to-follow rate of 5% per year in both the reference and experimental groups. Past experience leads to a base line hazard rate of 0.04. An equal sample allocation design will be used with a target power of 0.90 and significance level of 0.05.

Setup

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

Design Tab	
Solve For	**Power**
Alpha	**0.05**
Group Allocation	**Enter total sample size and percentage in Group 1**
Total Sample Size (N)	**1000 to 5000 by 1000**
Percent in Group 1	**50**
HR0 (Non-Inferiority Hazard Ratio)	**1.3**
h1 (Hazard Rate of Reference Group)	**0.04**
Accrual Time (Integers Only)	**2**
Accrual Pattern	**Uniform or Equal**
Total Time (Integers Only)	**5**
References Lost	**0.05**
References Switch to Treatment	**0.0**
Treatments Lost	**0.0**
Treatments Switch to Reference	**0.0**
Reports Tab	
Show Detail Numeric Reports	**Checked**

Output

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

Numeric Reports in Terms of Sample Size

Numeric Results (Sample Size)

Solve For: Power
Test Type: Two-Sample Logrank Test
Groups: 1 = Reference, 2 = Treatment
Hypotheses: H0: HR ≥ HR0 vs. H1: HR < HR0

	Sample Size			Hazard Ratio		Ref. Group Hazard		Time		Proportion Lost*		Proportion Switching Groups*		
Power	N1	N2	N	Non-Inferiority HR0	Actual HR1	Rate h1	Accrual Pattern	Accrual	Total	Ref.	Treat.	Ref. to Treat.	Treat. to Ref.	Alpha
0.4665	500	500	1000	1.3	1	0.04	Equal	2	5	0.05	0	0	0	0.05
0.7111	1000	1000	2000	1.3	1	0.04	Equal	2	5	0.05	0	0	0	0.05
0.8528	1500	1500	3000	1.3	1	0.04	Equal	2	5	0.05	0	0	0	0.05
0.9282	2000	2000	4000	1.3	1	0.04	Equal	2	5	0.05	0	0	0	0.05
0.9662	2500	2500	5000	1.3	1	0.04	Equal	2	5	0.05	0	0	0	0.05

* The reported proportions are during a single time period.

Power	The probability of rejecting a false null hypothesis when the alternative hypothesis is true.
N1, N2, N	The sample sizes of the reference group, treatment group, and both groups, respectively.
HR0	The Non-Inferiority Hazard Ratio is the upper bound for the hazard ratio that still leads to the conclusion of non-inferiority.
HR1	The Actual Hazard Ratio is assumed to be the actual value of the hazard ratio. This is always set to 1.
h1	The Reference Group Hazard Rate is the hazard (instantaneous failure) rate of the reference group. Its scale is events per time period.
Accrual Pattern	The pattern of accrual used for each time period.
Accrual Time	The number of time periods (years or months) during which accrual takes place.
Total Time	The total number of time periods in the study. Follow-up time = (Total Time) - (Accrual Time).
Reference Proportion Lost	The proportion of the reference group that is lost (drops out) during a single time period (e.g., year or month).
Treatment Proportion Lost	The proportion of the treatment group that is lost (drops out) during a single time period (e.g., year or month).
Reference to Treatment	Drop-In. The proportion of the reference group that switch to a group with a hazard rate equal to the treatment group.
Treatment to Reference	Noncompliance. The proportion of the treatment group that switch to a group with a hazard rate equal to the reference group.
Alpha	The probability of rejecting a true null hypothesis.

This report shows the values of each of the parameters, one scenario per row. We see that almost 4000 subjects will be required for this study.

Numeric Reports in Terms of Events

Next, a report displaying the number of required events rather than the sample size is displayed.

Numeric Results (Events)

Solve For: Power
Test Type: Two-Sample Logrank Test
Groups: 1 = Reference, 2 = Treatment
Hypotheses: H0: HR ≥ HR0 vs. H1: HR < HR0

	Number of Events			Hazard Ratio		Ref. Group Hazard		Time		Proportion Lost*		Proportion Switching Groups*		
Power	E1	E2	E	Non-Inferiority HR0	Actual HR1	Rate h1	Accrual Pattern	Accrual	Total	Ref.	Treat.	Ref. to Treat.	Treat. to Ref.	Alpha
0.4665	66.8	73.8	140.6	1.3	1	0.04	Equal	2	5	0.05	0	0	0	0.05
0.7111	133.6	147.6	281.3	1.3	1	0.04	Equal	2	5	0.05	0	0	0	0.05
0.8528	200.4	221.5	421.9	1.3	1	0.04	Equal	2	5	0.05	0	0	0	0.05
0.9282	267.3	295.3	562.5	1.3	1	0.04	Equal	2	5	0.05	0	0	0	0.05
0.9662	334.1	369.1	703.2	1.3	1	0.04	Equal	2	5	0.05	0	0	0	0.05

* The reported proportions are during a single time period.

Power	The probability of rejecting a false null hypothesis when the alternative hypothesis is true.
E1, E2, E	The number of events in the reference group, the treatment group, and both groups, respectively.
HR0	The Non-Inferiority Hazard Ratio is the upper bound for the hazard ratio that still leads to the conclusion of non-inferiority.
HR1	The Actual Hazard Ratio is assumed to be the actual value of the hazard ratio. This is always set to 1.
h1	The Reference Group Hazard Rate is the hazard (instantaneous failure) rate of the reference group. Its scale is events per time period.
Accrual Pattern	The pattern of accrual used for each time period.
Accrual Time	The number of time periods (years or months) during which accrual takes place.
Total Time	The total number of time periods in the study. Follow-up time = (Total Time) - (Accrual Time).
Reference Proportion Lost	The proportion of the reference group that is lost (drops out) during a single time period (e.g., year or month).
Treatment Proportion Lost	The proportion of the treatment group that is lost (drops out) during a single time period (e.g., year or month).
Reference to Treatment	Drop-In. The proportion of the reference group that switch to a group with a hazard rate equal to the treatment group.
Treatment to Reference	Noncompliance. The proportion of the treatment group that switch to a group with a hazard rate equal to the reference group.
Alpha	The probability of rejecting a true null hypothesis.

Most of this report is identical to the last report, except that the sample sizes are replaced by the number of required events.

Reports of Detailed Input

Next, reports displaying the individual settings year-by-year for each scenario are displayed.

Detailed Input when Power = 0.4665, N1 = 500, N2 = 500, N = 1000, HR0 = 1.3, HR1 = 1, Accrual Time = 2, Total Time = 5, Alpha = 0.05

Time Period	Reference Hazard Rate h1 (0.04)	Percent Accrual (Equal)	Percent Administratively Censored (Calculated)	Proportion Lost		Proportion Switching Groups	
				Reference (0.05)	Treatment (0)	Reference to Treatment (0)	Treatment to Reference (0)
1	0.04	50	0	0.05	0	0	0
2	0.04	50	0	0.05	0	0	0
3	0.04	0	0	0.05	0	0	0
4	0.04	0	50	0.05	0	0	0
5	0.04	0	100	0.05	0	0	0

(More Reports Follow)

This report shows the individual settings for each time period (year). It becomes very useful when you want to document a study in which these parameters vary from year to year.

Summary Statements and References

Next, summary statements and references are displayed.

Summary Statements

A parallel, two-group design (where higher hazard rates are considered worse) will be used to test whether the Group 2 (treatment) hazard rate is non-inferior to the Group 1 (control) hazard rate, with a non-inferiority hazard ratio of 1.3 (H0: HR ≥ 1.3 versus Ha: HR < 1.3, HR = h2 / h1). The comparison will be made using a one-sided, two-sample logrank test with a Type I error rate (α) of 0.05. The total duration of the study will be 5 time periods with subject accrual (entry) occurring in the first 2 time periods. The accrual pattern across time periods will be uniform (all periods equal). The proportion of subjects dropping out of the control group during each time period will be 0.05. It is anticipated that no subjects will drop out of the treatment group. The proportion of subjects switching from the control group to another group with a hazard rate equal to the treatment group is 0 (per time period). The proportion of subjects switching from the treatment group to another group with a hazard rate equal to the control group is 0 (per time period). To detect a hazard ratio (h2 / h1) of 1 (h1 = 0.04, h2 = 0.04), with a sample size of 500 subjects in Group 1 and 500 subjects in Group 2 (totaling 1000 subjects), the power is 0.4665. The corresponding number of events is 66.8 in Group 1 and 73.8 in Group 2 (totaling 140.6 events).

References

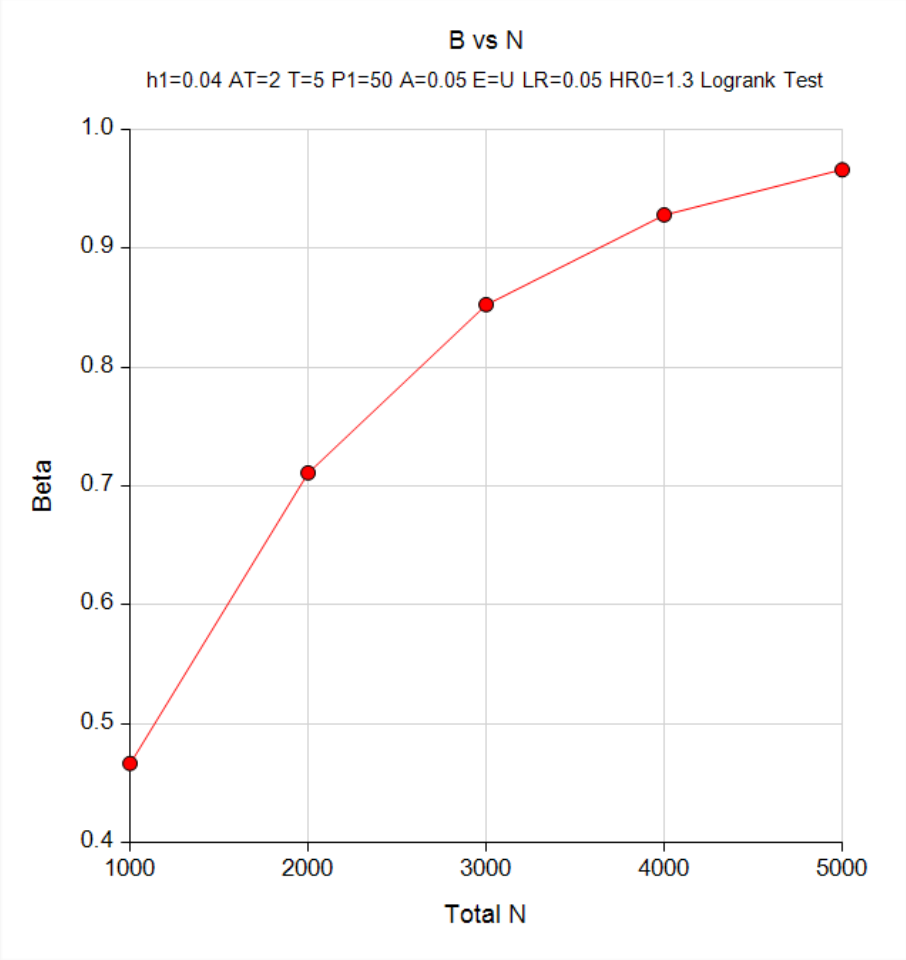
Jung,Sin-Ho; Kang, Sun J.; McCall, Linda M.; Blumenstein, Brent. 2005. 'Sample Sizes Computation for Two-Sample Noninferiority Log-Rank Test', J. of Biopharmaceutical Statistics, Volume 15, pages 969-979.

Lakatos, Edward. 1988. 'Sample Sizes Based on the Log-Rank Statistic in Complex Clinical Trials', Biometrics, Volume 44, March, pages 229-241.

Finally, a scatter plot of the results is displayed.

Plots Section

Plots



This plot shows the relationship between sample size and. Note that for 90% power, a total sample size of about 4000 is required. The exact number will be found in Example 2.

Example 2 – Finding the Sample Size

Continuing with the previous example, the researcher wants to investigate the sample sizes necessary to achieve 80% and 90% power. All other parameters will remain the same as in Example 1.

Setup

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

Design Tab

Solve For	**Sample Size**
Power	**0.80 0.90**
Alpha	**0.05**
Group Allocation	**Equal (N1 = N2)**
HR0 (Non-Inferiority Hazard Ratio)	**1.3**
h1 (Hazard Rate of Reference Group)	**0.04**
Accrual Time (Integers Only)	**2**
Accrual Pattern	**Uniform or Equal**
Total Time (Integers Only)	**5**
References Lost	**0.05**
References Switch to Treatment	**0.0**
Treatments Lost	**0.05**
Treatments Switch to Reference	**0.0**

Output

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

Numeric Results (Sample Size)

Solve For: Sample Size
Test Type: Two-Sample Logrank Test
Groups: 1 = Reference, 2 = Treatment
Hypotheses: H0: HR ≥ HR0 vs. H1: HR < HR0

	Sample Size			Hazard Ratio		Ref. Group Hazard		Time		Proportion Lost*		Proportion Switching Groups*		
Power	N1	N2	N	Non-Inferiority HR0	Actual HR1	Rate h1	Accrual Pattern	Accrual	Total	Ref.	Treat.	Ref. to Treat.	Treat. to Ref.	Alpha
0.8	1344	1345	2689	1.3	1	0.04	Equal	2	5	0.05	0.05	0	0	0.05
0.9	1865	1866	3731	1.3	1	0.04	Equal	2	5	0.05	0.05	0	0	0.05

* The reported proportions are during a single time period.

Example 3 – Validation using Jung (2005)

Jung et al. (2005) pages 974-975 present an example that will be used to validate this procedure. In this article, an 8.8-year trial is presented in which patient accrual occurs the first 3.8 years. The baseline hazard rate is 0.0446. The value of *HR*0 is 1.3 and the value of *HR* is 1.0. Equal allocation between groups is used and uniform accrual is assumed. The significance level is 0.05 and the desired power is 0.90. Given these values, the number of events is found to be 499 and the sample size is 1891.

Since this procedure using integer values for the accrual and trial time, the accrual time and total time will be set to 4 and 9 years, respectively.

Setup

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

Design Tab	
Find (Solve For)	**Sample Size**
Power	**0.90**
Alpha	**0.05**
Group Allocation	**Equal (N1 = N2)**
HR0	**1.3**
h1	**0.0446**
Accrual Time	**4**
Accrual Pattern	**Uniform or Equal**
Total Time	**9**
References Lost	**0.0**
References Switch to Treatment	**0.0**
Treatments Lost	**0.0**
Treatments Switch to Reference	**0.0**

Output

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

Numeric Results (Sample Size)

Solve For: Sample Size
Test Type: Two-Sample Logrank Test
Groups: 1 = Reference, 2 = Treatment
Hypotheses: H0: HR ≥ HR0 vs. H1: HR < HR0

	Sample Size			Hazard Ratio		Ref. Group Hazard		Time		Proportion Lost*		Proportion Switching Groups*		
Power	N1	N2	N	Non-Inferiority HR0	Actual HR1	Rate h1	Accrual Pattern	Accrual	Total	Ref.	Treat.	Ref. to Treat.	Treat. to Ref.	Alpha
0.9	933	933	1866	1.3	1	0.0446	Equal	4	9	0	0	0	0	0.05

* The reported proportions are during a single time period.

Numeric Results (Events)

Solve For: Sample Size
Test Type: Two-Sample Logrank Test
Groups: 1 = Reference, 2 = Treatment
Hypotheses: H0: HR ≥ HR0 vs. H1: HR < HR0

	Number of Events			Hazard Ratio		Ref. Group Hazard		Time		Proportion Lost*		Proportion Switching Groups*		
Power	E1	E2	E	Non-Inferiority HR0	Actual HR1	Rate h1	Accrual Pattern	Accrual	Total	Ref.	Treat.	Ref. to Treat.	Treat. to Ref.	Alpha
0.9	249.3	249.3	498.6	1.3	1	0.0446	Equal	4	9	0	0	0	0	0.05

* The reported proportions are during a single time period.

Note that the number of events (499) matches Jung's results exactly. The sample size of 1866 is slightly less than Jung's 1891. This difference occurs because these results were obtained for 4 years of accrual, not 3.8, and because we used Lakatos' method for transforming the number of events into the sample size.

Example 4 – Inputting Time-Dependent Hazard Rates from a Spreadsheet

Time-dependent parameters (hazard rates, losses to follow-up, etc) may be entered. See Example 4 of Chapter 715 (Logrank Tests) for an extensive example of how this is done for the logrank test.