Chapter 731

Non-Inferiority Tests for Two Survival Curves Using Cox's Proportional Hazards Model

Introduction

A clinical trial may be employed to test the non-inferiority of a treatment over a control in regards to their survival distributions. Because survival times are not normally distributed and because some survival times are censored, Cox proportional-hazards regression is often used to analyze the data. The formulation for testing the significance of a Cox regression coefficient is identical to the standard logrank test. Thus the power and sample size formulas for one analysis also work for the other.

The Cox Regression model has the added benefit over the exponential model that it does not assume that the hazard rates are constant, but only that they are proportional. That is, that the hazard ratio remains constant throughout the experiment, even if the hazard rates vary.

This procedure is documented in Chow, Shao, and Wang (2008) which summarizes the work of Schoenfeld (1981, 1983). Note that there was an error in Chow, Shao, and Wang (2008) page 179 which caused the sample size to be doubled. This error has been corrected in this edition.

Technical Details

Cox's Proportional Hazards Regression

Cox's proportional hazards regression is widely used for survival data. The regression model is

$$h(t \mid z) = h(t \mid 0)\exp(bz)$$

where

b is the regression coefficient which is equal to $\log\left[h(t\,|\,1)/h(t\,|\,0)=\log(HR)\right]$

z is a binary indicator variable of treatment group

t is elapsed time

$h(t/z)$ is the hazard rate at time t, given covariate z

HR is the hazard ratio $h(t\,|\,1)/h(t\,|\,0)$

Non-Inferiority Hypothesis

Lower Hazards Better

Assuming that lower hazard rates are better, non-inferiority means that the treatment hazard rate is at most, only slightly higher than the control hazard rate. We find it more convenient to state the hypotheses in terms of the hazard ratio, HR, rather than the Cox regression coefficient, b. Remembering that $b = \log(HR)$ and assuming that $HR_{ni} > 1$, non-inferiority requires that $HR < HR_{ni}$. Here, HR_{ni} is the boundary of clinical insignificance or the non-inferiority boundary.

The statistical hypotheses that results in the conclusion of non-inferiority when the null hypothesis is rejected is

$$H_0 : \log(HR) \geq \log(HR_{ni}) \quad \text{versus} \quad H_a : \log(HR) < \log(HR_{ni})$$

Higher Hazards Better

Assuming that higher hazard rates are better, non-inferiority means that the treatment hazard rate is at most, only slightly lower than the control hazard rate. We find it more convenient to state the hypotheses in terms of the hazard ratio, HR, rather than the Cox regression coefficient, b. Remembering that $b = \log(HR)$ and assuming that $HR_{ni} < 1$, non-inferiority requires that $HR > HR_{ni}$. Here, HR_{ni} is called the boundary of clinical insignificance or the non-inferiority boundary.

The statistical hypotheses that results in the conclusion of non-inferiority when the null hypothesis is rejected is

$$H_0 : \log(HR) \leq \log(HR_{ni}) \quad \text{versus} \quad H_a : \log(HR) > \log(HR_{ni})$$

Test Statistic

It can be shown that the test of b based on the partial likelihood method of Cox (1972) coincides with the common logrank test statistic. When testing non-inferiority, you can use b from a Cox regression or calculate the modified logrank statistic as follows

Logrank Test

The logrank test statistic is

$$L = \frac{\sum_{k=1}^{K}\left(I_k - \dfrac{Y_{1i}HR_{ni}}{Y_{1i}HR_{ni} + Y_{2i}}\right)}{\left[\sum_{k=1}^{K}\left(\dfrac{Y_{1i}Y_{2i}HR_{ni}}{\left(Y_{1i}HR_{ni} + Y_{2i}\right)^2}\right)\right]^{-1/2}}$$

Where K is the number of deaths, Y_{ij} is the number of subjects at risk just prior to the jth observed event in the ith group, and I_k is is a binary variable indicating whether the kth even is from group 1 or not.

The distribution of L is approximately normal with mean $\left(\log(HR) - \log(HR_{ni})\right)\sqrt{P_1 P_2 dN}$ and unit variance, where

P_1 is the proportion of N that is in the control group

P_2 is the proportion of N that is in the treatment group

N is the total sample size

N_1 is the sample size from the control group, $N_1 = N(P_1)$

N_2 is the sample size from the treatment group, $N_2 = N(P_2)$

Pev_1 is probability of the event of interest in the control group

Pev_2 is probability of the event of interest in the treatment group

d is the overall probability of an event, $d = Pev_1 P_1 + Pev_2 P_2$

b is the Cox regression coefficient, $b = \log(HR)$

Power Calculations

The power of this test is given by

$$\Phi\left(\left(\log(HR) - \log(HR_{ni})\right)\sqrt{P_1 P_2 dN} - z_{1-\alpha}\right)$$

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 *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.

Test Direction

Alternative Hypothesis

In most studies, a lower hazard rate is better since it indicates a longer life-time. However, in some situations, a higher hazard rate is desired. For example, if the length of life of disease organisms is being studied. This option lets you specify which option you have.

Each option results in different null and alternative hypotheses.

- **Ha: HR < HRni (Lower Hazard Better)**

  In this case lower is better. The alternative hypothesis is that hazard ratio is lower than a clinically significant value called the clinical superiority ratio, *HRni*.

- **Ha: HR > HRni (Higher Hazard Better)**

  In this case higher is better. The alternative hypothesis is that hazard ratio is higher than a clinically significant value called the clinical superiority ratio, *HRni*.

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 (When Solving for Sample Size)

Group Allocation

Select the option that describes the constraints on *N1* or *N2* or both.

The options are

- **Equal (N1 = N2)**

  This selection is used when you wish to have equal sample sizes in each group. Since you are solving for both sample sizes at once, no additional sample size parameters need to be entered.

- **Enter R = N2/N1, solve for N1 and N2**

  For this choice, you set a value for the ratio of *N2* to *N1*, and then PASS determines the needed *N1* and *N2*, with this ratio, to obtain the desired power. An equivalent representation of the ratio, *R*, is

  $$N2 = R * N1.$$

- **Enter percentage in Group 1, solve for N1 and N2**

  For this choice, you set a value for the percentage of the total sample size that is in Group 1, and then PASS determines the needed *N1* and *N2* with this percentage to obtain the desired power.

R (Group Sample Size Ratio)

This option is displayed only if Group Allocation = "Enter R = N2/N1, solve for N1 and N2."

R is the ratio of *N2* to *N1*. That is,

$$R = N2 / N1.$$

Use this value to fix the ratio of *N2* to *N1* while solving for *N1* and *N2*. Only sample size combinations with this ratio are considered.

N2 is related to *N1* by the formula:

$$N2 = [R \times N1],$$

where the value *[Y]* is the next integer $\geq Y$.

For example, setting *R* = 2.0 results in a Group 2 sample size that is double the sample size in Group 1 (e.g., *N1* = 10 and *N2* = 20, or *N1* = 50 and *N2* = 100).

R must be greater than 0. If *R* < 1, then *N2* will be less than *N1*; if *R* > 1, then *N2* will be greater than *N1*. You can enter a single or a series of values.

Percent in Group 1

This option is displayed only if Group Allocation = "Enter percentage in Group 1, solve for N1 and N2."

Use this value to fix the percentage of the total sample size allocated to Group 1 while solving for *N1* and *N2*. Only sample size combinations with this Group 1 percentage are considered. Small variations from the specified percentage may occur due to the discrete nature of sample sizes.

The Percent in Group 1 must be greater than 0 and less than 100. You can enter a single or a series of values.

Sample Size (When <u>Not</u> Solving for Sample Size)

Group Allocation

Select the option that describes how individuals in the study will be allocated to Group 1 and to Group 2.

The options are

- **Equal (N1 = N2)**

  This selection is used when you wish to have equal sample sizes in each group. A single per group sample size will be entered.

- **Enter N1 and N2 individually**

  This choice permits you to enter different values for *N1* and *N2*.

- **Enter N1 and R, where N2 = R * N1**

  Choose this option to specify a value (or values) for *N1*, and obtain *N2* as a ratio (multiple) of *N1*.

- **Enter total sample size and percentage in Group 1**

  Choose this option to specify a value (or values) for the total sample size (*N*), obtain *N1* as a percentage of *N*, and then *N2* as *N - N1*.

Sample Size Per Group

This option is displayed only if Group Allocation = "Equal (N1 = N2)."

The Sample Size Per Group is the number of items or individuals sampled from each of the Group 1 and Group 2 populations. Since the sample sizes are the same in each group, this value is the value for *N1*, and also the value for *N2*.

The Sample Size Per Group must be ≥ 2. You can enter a single value or a series of values.

N1 (Sample Size, Group 1)

This option is displayed if Group Allocation = "Enter N1 and N2 individually" or "Enter N1 and R, where N2 = R * N1."

N1 is the number of items or individuals sampled from the Group 1 population.

N1 must be ≥ 2. You can enter a single value or a series of values.

N2 (Sample Size, Group 2)

This option is displayed only if Group Allocation = "Enter N1 and N2 individually."

N2 is the number of items or individuals sampled from the Group 2 population.

N2 must be ≥ 2. You can enter a single value or a series of values.

R (Group Sample Size Ratio)

This option is displayed only if Group Allocation = "Enter N1 and R, where N2 = R * N1."

R is the ratio of *N2* to *N1*. That is,

$$R = N2/N1$$

Use this value to obtain *N2* as a multiple (or proportion) of *N1*.

N2 is calculated from *N1* using the formula:

$$N2=[R \times N1],$$

where the value *[Y]* is the next integer ≥ *Y*.

For example, setting *R* = 2.0 results in a Group 2 sample size that is double the sample size in Group 1.

R must be greater than 0. If *R* < 1, then *N2* will be less than N1; if *R* > 1, then *N2* will be greater than *N1*. You can enter a single value or a series of values.

Total Sample Size (N)

This option is displayed only if Group Allocation = "Enter total sample size and percentage in Group 1."

This is the total sample size, or the sum of the two group sample sizes. This value, along with the percentage of the total sample size in Group 1, implicitly defines *N1* and *N2*.

The total sample size must be greater than one, but practically, must be greater than 3, since each group sample size needs to be at least 2.

You can enter a single value or a series of values.

Percent in Group 1

This option is displayed only if Group Allocation = "Enter total sample size and percentage in Group 1."

This value fixes the percentage of the total sample size allocated to Group 1. Small variations from the specified percentage may occur due to the discrete nature of sample sizes.

The Percent in Group 1 must be greater than 0 and less than 100. You can enter a single value or a series of values.

Sample Size – Proportion of Events Observed

Pev1 (Probability of a Control Event)

Enter one or more values for *Pev1*, the average probability of an event (death) for each individual in the control group. The number of events in the control group is equal to *Pev1* x *N1*.

Range

$0 < Pev1 < 1$

Pev2 (Probability of a Treatment Event)

Enter one or more values for *Pev2*, the average probability of an event (death) for each individual in the control group. The number of events in the control group is equal to *Pev2* x *N2*.

Range

$0 < Pev2 < 1$

Hazard Ratios

HR (Hazard Ratio = h2/h1)

Specify one or more hazard ratios, *HR*, at which the power is calculated. *HR* is defined as the treatment hazard rate (*h2*) divided by the control hazard rate (*h1*). *HR* represents the assumed value at which the power of rejecting H0 is calculated. In non-inferiority tests, this ratio is usually assumed to be 1.0

The null hypothesis is H0: HR=1. The alternative hypothesis is Ha: HR<HR_{ni} or Ha: HR>HR_{ni}.

Range

If Ha: HR < HR_{ni} [Lower Hazard Better] then the range is $0 < HR < HR_{ni}$. Typical values are between 0.9 and 2.0. Usually, HR = 1.

If Ha: HR > HR_{ni} [Higher Hazard Better] then the range is HR > HR_{ni}. Typical values are from 0.25 up to 1.1. Usually, HR = 1.

Press the "Survival Parameter Conversion Tool" button to estimate reasonable values based on median survival times, hazard rates, or the proportion surviving.

HRni (Non-Inferiority Hazard Ratio)

Specify one or more values for the non-inferiority ratio, HR_{ni}. If lower/higher is better, HR_{ni} is the largest/smallest that HR can be that still concludes non-inferiority.

The one-sided alternative hypotheses are Ha: HR < HR_{ni} or Ha: HR > HR_{ni}.

Range

If Ha: HR < HR_{ni} [Lower Hazard Better], the range is $0 < HR < HR_{ni}$. Typical values are between 1.01 and 4.0.

If Ha: HR > HR_{ni} [Higher Hazard Better], the range is $HR_{ni} < HR$. Typical values are from 0.25 up to 0.99.

Example 1 – Finding the Sample Size

A researcher is planning a non-inferiority trial using a parallel, two-group, equal sample allocation design to compare the survivability of a new treatment with that of the current treatment. The non-inferiority ratio is 1.2. The desired power is 0.90 and significance level is 0.05. HR will be between 0.5 and 1. The probability of observing an event is 0.50 in the control group and 0.30 in the treatment group.

Setup

This section presents the values of each of the parameters needed to run this example. First, from the PASS Home window, load the **Non-Inferiority Tests for Two Survival Curves Using Cox's Proportional Hazards Model** procedure window by expanding **Survival**, then **Two Survival Curves**, then clicking **Non-Inferiority**, and then clicking on **Non-Inferiority Tests for Two Survival Curves Using Cox's Proportional Hazards Model**. 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**.

Option	Value
Design Tab	
Solve For	**Sample Size**
Alternative Hypothesis	**Ha: HR < HRni [Lower Hazard Better]**
Power	**0.90**
Alpha	**0.05**
Group Allocation	**Equal (N1 = N2)**
Pev1 (Event Probability in Group 1)	**0.5**
Pev2 (Event Probability in Group 2)	**0.3**
HR (Actual Hazard Ratio = h2/h1)	**0.5 to 1 by 0.1**
HRni (Non-Inferiority Hazard Ratio)	**1.2**

Annotated Output

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

Numeric Results

Numeric Results with Ha: HR < HRni

Power	Total Sample Size N	Control Sample Size N1	Trtmnt Sample Size N2	Prop'n Control N1/N P1	Hazard Ratio h2/h1 HR	Non Inf Hazard Ratio HRni	Control Prob Event Pev1	Trtmnt Prob Event Pev2	Control Events E1	Trtmnt Events E2	Alpha
0.9006	112	56	56	0.500	0.500	1.200	0.500	0.300	28.0	16.8	0.050
0.9007	179	89	90	0.500	0.600	1.200	0.500	0.300	44.5	27.0	0.050
0.9011	296	148	148	0.500	0.700	1.200	0.500	0.300	74.0	44.4	0.050
0.9005	522	261	261	0.500	0.800	1.200	0.500	0.300	130.5	78.3	0.050
0.9003	1036	518	518	0.500	0.900	1.200	0.500	0.300	259.0	155.4	0.050
0.9000	2577	1288	1289	0.500	1.000	1.200	0.500	0.300	644.0	386.7	0.050

References
Chow, S.C., Shao, J., Wang, H. 2008. Sample Size Calculations in Clinical Research, 2nd Edition. Chapman & Hall/CRC.
Schoenfeld, David A. 1983. 'Sample Size Formula for the Proportional-Hazards Regression Model', Biometrics, Volume 39, Pages 499-503.

Report Definitions
Power is the probability of rejecting a false null hypothesis. Power should be close to one.
N is the total sample size.
N1 and N2 are the sample sizes of the control and treatment groups.
P1 is the proportion of the total sample that is in the control group, group 1.
HR is the hazard ratio: h2/h1.
HRni is the non-inferiority hazard ratio.
Pev1 and Pev2 are the probabilities of an event in the control and the treatment groups.
E1 and E2 are the number of events required in the control and the treatment groups.
Alpha is the probability of a type one error: rejecting a true null hypothesis.

Summary Statements
A non-inferiority test of whether the hazard ratio with an overall sample size of 112 subjects
(of which 56 are in the control group and 56 are in the treatment group) achieves 90% power at
a 0.050 significance level when the hazard ratio is actually 0.500. The non-inferiority ratio
is 1.200. The number of events required to achieve this power is 44.8. It is anticipated that
the proportion of subjects observed with the event during the study is 0.500 for the control
group and 0.300 for the treatment group. These results assume that the hazard ratio is constant
throughout the study and that Cox proportional hazards regression or the non-inferiority
logrank test is used to analyze the data.

These reports show the values of each of the parameters, one scenario per row.

Plots Section
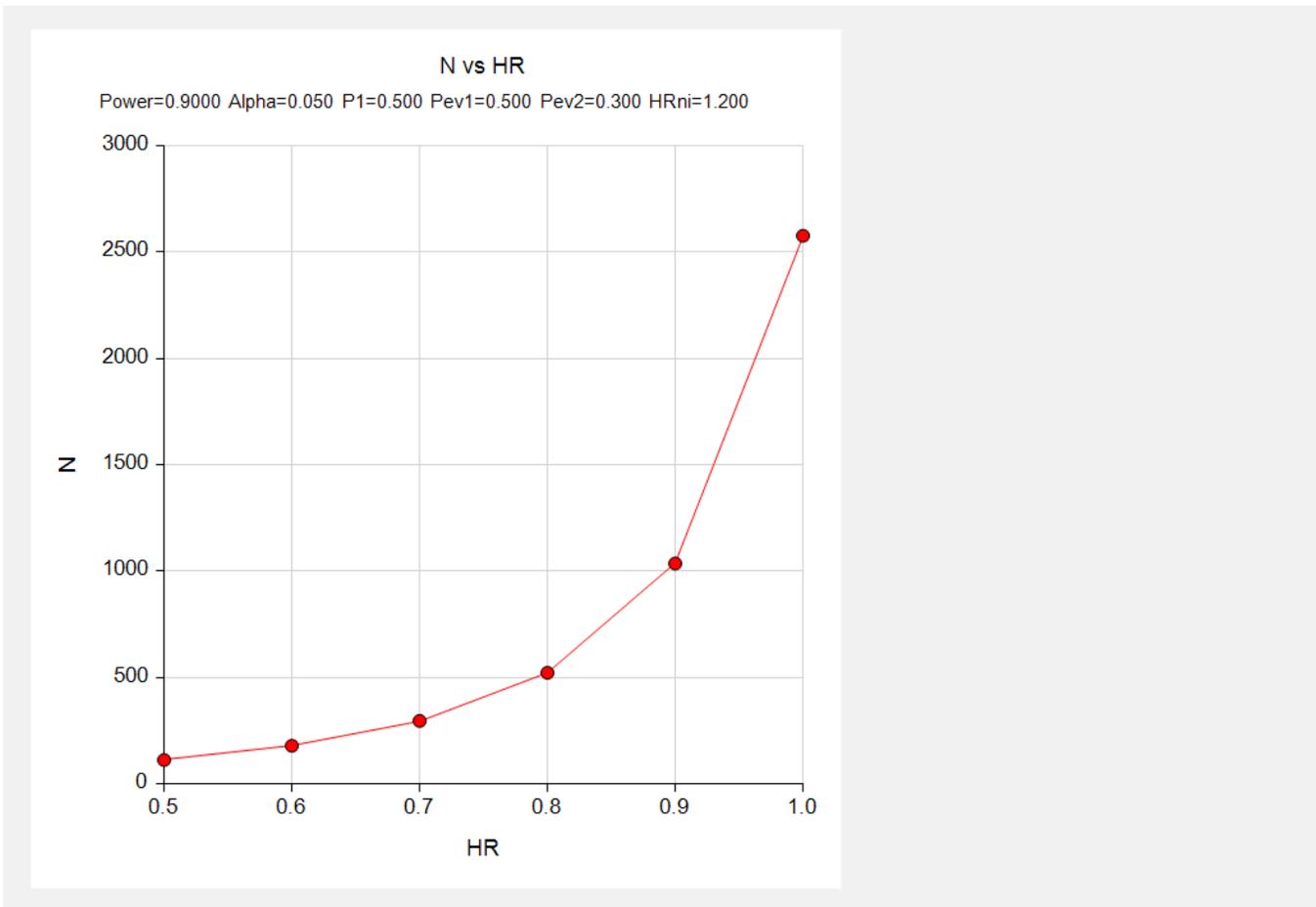
This plot shows the relationship between HR and N.

Example 2 – Validation using Chow et al. (2008)

Chow et al. (2008) page 179 present an example that we will use for validation. In their example, HR = 2, log(HRni) = 0.3, Pev1 = Pev2 = 0.8, P1 = 0.5, alpha = 0.05, and power = 0.8. They obtained a value of about 100 per group (this is a correction from the 200 that they originally printed in error).

Actually, this example is for a superiority test, but if HR and HRni are switched, it can be used to validate the non-inferiority 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 **Non-Inferiority Tests for Two Survival Curves Using Cox's Proportional Hazards Model** procedure window by expanding **Survival**, then **Two Survival Curves**, then clicking **Non-Inferiority**, and then clicking on **Non-Inferiority Tests for Two Survival Curves Using Cox's Proportional Hazards Model**. 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**.

Option	Value
Design Tab	
Solve For	**Sample Size**
Alternative Hypothesis	**Ha: HR < HRni [Lower Hazard Better]**
Power...	**0.80**
Alpha..	**0.05**
Group Allocation	**Equal (N1 = N2)**
Pev1 (Probability of a Control Event)	**0.8**
Pev2 (Probability of a Treatment Event)..............	**Pev1**
HR (Hazard Ratio = h2/h1)	**1.35** (which is exp(0.3))
HRni (Non-Inferiority Hazard Ratio)....................	**2**

Annotated Output

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

Numeric Results

Numeric Results with Ha: HR < HRni

Power	Total Sample Size N	Control Sample Size N1	Trtmnt Sample Size N2	Prop'n Control N1/N P1	Hazard Ratio h2/h1 HR	Non Inf Hazard Ratio HRni	Control Prob Event Pev1	Trtmnt Prob Event Pev2	Control Events E1	Trtmnt Events E2	Alpha
0.8015	201	100	101	0.500	1.350	2.000	0.800	0.800	80.0	80.8	0.050

PASS also calculates the value of N = 201 which is within rounding of the 200 that Chow calculated.