

## Chapter 622

# Tests for Two Survival Curves using Cox's Proportional Hazards Model in a Cluster-Randomized Design

---

## Introduction

This procedure calculates power and sample size for testing the equality of survival distributions of two treatment groups using Cox's proportional hazards regression when the data are obtained in a cluster-randomized design. 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, Wang, and Lokhnygina (2018) and Machin, Campbell, Tan, and Tan (2018) which are based on the work of Schoenfeld (1981, 1983).

Cluster-randomized designs are those in which whole clusters of subjects (classes, hospitals, communities, etc.) are randomized into the treatment or control groups. In this design, the survival curves of the two groups, made up of  $K_i$  clusters of  $M_{ij}$  individuals each, are to be tested using a regression coefficient from a Cox regression.

---

## Technical Details

---

### Cox's Proportional Hazards Regression

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

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

where

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

$z$  is a binary indicator variable of treatment group

$t$  is elapsed time

## Tests for Two Survival Curves using Cox's Proportional Hazards Model in a Cluster-Randomized Design

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

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

The two-sided, statistical hypothesis testing survival equality is a test of whether  $b$  is zero. This hypothesis is stated as

$$H_0: b = 0 \quad \text{vs.} \quad H_1: b \neq 0$$

---

## 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 shown next.

### Logrank Test

The logrank test statistic is

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

where  $K$  is the number of deaths,  $Y_{ij}$  is the number of subjects at risk just prior to the  $j^{\text{th}}$  observed event in the  $i^{\text{th}}$  group, and  $I_k$  is a binary variable indicating whether the  $k^{\text{th}}$  event is from group 1 or not.

The distribution of  $L$  is approximately normal with mean  $b\sqrt{P_1 P_2 d N}$  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,  $N_2 = N(P_2)$

$Pe_{v1}$  is probability of the event of interest in the control group

$Pe_{v2}$  is probability of the event of interest in the treatment group

$d$  is the overall probability of an event,  $d = Pe_{v1}P_1 + Pe_{v2}P_2$

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

## Cluster-Randomized Designs

Denote an observation by  $Y_{ijk}$  where  $i = 1, 2$  gives the group,  $j = 1, 2, \dots, K_i$  gives the cluster within group  $i$ , and  $k = 1, 2, \dots, m_{ij}$  denotes an individual in cluster  $j$  of group  $i$ . In this chapter, we will assume that group 1 is the control group and group 2 is the treatment group.

Let  $\rho$  denote the intraclass correlation coefficient (ICC) among individuals from the same cluster. This correlation is the correlation of censor indicator variable. Also let COV denote the coefficient of variation of the cluster sizes. Machin *et al.* (2018) page 101 shows that the number of events,  $e$ , that are needed to obtain a given power of  $1 - \beta$  and a two-sided significance level of  $\alpha$  to detect a hazard ratio of  $HR (h_2 / h_1)$  is given by

$$e = \frac{(1 + r)^2}{r} \left[ \frac{(z_{1-\alpha/2} + z_{1-\beta})^2}{\log^2(HR_1)} \right]$$

where  $r = N_2 / N_1$  and  $z_x = \Phi(x)$  is the standard normal distribution function.

The number of subjects in a regular design can be determined using

$$N = N_1 + N_2 = \frac{e}{F}$$

where

$$F = \frac{(Pev_1 + r(Pev_2))}{1 + r}$$

### Design Effect

When using a cluster-randomized design Machin *et al.* (2018) page 202 show that the above formula is modified using a quantity that is known as the *design effect* (DE). The version of DE that is used in **PASS** is given as formula 12.7 on page 197 of Machin *et al.* (2018) which is

$$DE = 1 + \{[CV^2 + 1]\bar{M} - 1\}\rho$$

where  $\bar{M}$  is the average cluster size of all clusters given by

$$\bar{M} = \frac{K_1 M_1 + K_2 M_2}{K_1 + K_2}$$

$CV$  is the coefficient of variation of cluster sizes of all clusters in the study and  $\rho$  is the ICC as defined above.

The resulting sample size formula in terms of cluster counts and size is

$$N = K_1 M_1 + K_2 M_2 = DE \left( \frac{e}{F} \right)$$

## Power Calculations

The power of the statistical test of  $b$  is given by

$$\Phi(b\sqrt{P_1P_2dN} - z_{1-\alpha/2})$$

or equivalently

$$\Phi(\log(HR_1)\sqrt{P_1P_2dN} - z_{1-\alpha/2})$$

where  $HR_1$  is the actual assumed value of the hazard ratio under the alternative hypothesis and  $N$  is defined as above.

## Example 1 – Showing Impact of ICC and COV on K1 and K2

This example will not only show how to find the sample size but will also show how ICC and COV impact the sample size.

Suppose that a cluster-randomized study is being planned in which  $HR1 = 1.943$ ;  $Pev1 = 0.7$ ;  $Pev2 = 0.5$ ;  $\rho = 0.00, 0.01, 0.05$ ;  $M1$  and  $M2 = 20$ ;  $alpha = 0.05$ ;  $power = 0.90$ ;  $COV = 0.0, 0.6$ ; and  $K1 = K2$ . Sample size (number of clusters) is desired for a two-sided test.

### 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 .....	<b>Number of Clusters</b>
Alternative Hypothesis .....	<b>Two-Sided (H1: HR ≠ 1)</b>
Power.....	<b>0.90</b>
Alpha.....	<b>0.05</b>
M1 (Average Cluster Size).....	<b>20</b>
K2 (Number of Clusters) .....	<b>K1</b>
M2 (Average Cluster Size).....	<b>M1</b>
COV of Cluster Sizes.....	<b>0 0.6</b>
Pev1 (Probability of a Control Event).....	<b>0.7</b>
Pev2 (Probability of a Treatment Event) .....	<b>0.5</b>
HR1 (Actual Hazard Ratio to Detect) .....	<b>1.943</b>
$\rho$ (Intracluster Correlation) .....	<b>0 0.01 0.05</b>

## Output

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

### Numeric Reports

#### Numeric Results

Solve For: Number of Clusters  
 Groups: 1 = Control, 2 = Treatment  
 Hypotheses: H0: HR = 1 vs. H1: HR ≠ 1

Group	Power	Sample Size Ni	Number of Events Ei	Number of Clusters Ki	Cluster Size		Probability of an Event Pevi	Hazard Ratio HR1	Intracluster Correlation ρ	Alpha
					Average Mi	COV				
1: Control		80	56.0	4	20	0.0	0.7		0.00	
2: Treatment	0.90218	80	40.0	4	20	0.0	0.5	1.943	0.00	0.05
Total		160	96.0	8						
1: Control		100	83.3	5	20	0.0	0.7		0.01	
2: Treatment	0.91545	100	59.5	5	20	0.0	0.5	1.943	0.01	0.05
Total		200	142.8	10						
1: Control		160	218.4	8	20	0.0	0.7		0.05	
2: Treatment	0.90915	160	156.0	8	20	0.0	0.5	1.943	0.05	0.05
Total		320	374.4	16						
1: Control		80	56.0	4	20	0.6	0.7		0.00	
2: Treatment	0.90218	80	40.0	4	20	0.6	0.5	1.943	0.00	0.05
Total		160	96.0	8						
1: Control		120	106.0	6	20	0.6	0.7		0.01	
2: Treatment	0.94382	120	75.7	6	20	0.6	0.5	1.943	0.01	0.05
Total		240	181.7	12						
1: Control		200	323.4	10	20	0.6	0.7		0.05	
2: Treatment	0.92296	200	231.0	10	20	0.6	0.5	1.943	0.05	0.05
Total		400	554.4	20						

- Group The group to which the results on this line belong. Group 1 is the control group and group 2 is the treatment group. The first (Total) line shows the total when appropriate.
- Power The probability of rejecting a false null hypothesis for this hazard ratio. This is the power calculated at HR1.
- Ni The number of items (subjects) in the ith group. The total sample size is shown as the last row of each three-row set.
- Ei The number of events in the ith group required to achieve the power indicated.
- Ki The number of clusters in the ith group. The total number of clusters is reported in the last row of each three-row set.
- Mi The average number of items per cluster (or average cluster size) in the ith group.
- COV The coefficient of variation of the cluster sizes of the clusters within the group. This value is used for both groups.
- Pevi The average probability that a subject in the ith group will have an event during the study. This probability includes the impact of various kinds of censoring.
- HR1 The hazard ratio of the treatment group and the control group.  $HR1 = h2 / h1$ . The subscript '1' indicates that this hazard ratio value is used for the alternative hypothesis, H1.
- ρ The intracluster correlation (ICC) between any two observations from within a cluster. This statistic has two common interpretations: (1) the proportion of variation related to the between-cluster variation or (2) the correlation between two subjects from the same cluster.
- Alpha The probability of rejecting a true null hypothesis.

## Tests for Two Survival Curves using Cox's Proportional Hazards Model in a Cluster-Randomized Design

**Summary Statements**

---

A parallel, two-group cluster-randomized design will be used to test whether the Group 2 (treatment) hazard rate is different from the Group 1 (control) hazard rate ( $H_0: HR = 1$  versus  $H_a: HR \neq 1, HR = h_2 / h_1$ ). The comparison will be made using a two-sided, two-sample Cox proportional hazards regression coefficient test with a Type I error rate ( $\alpha$ ) of 0.05. It is anticipated that the probability of observing an event during the course of the study is 0.7 for each member of the control group and 0.5 for each member of the treatment group. The coefficient of variation of the cluster sizes within a group is assumed to be 0. The intracluster correlation coefficient of responses within a particular cluster is assumed to be 0. The calculations are based on the assumption that the hazard ratio is constant throughout the study. To detect a hazard ratio ( $h_2 / h_1$ ) of 1.943, with average cluster sizes of 20 subjects per cluster in Group 1 and 20 subjects per cluster in Group 2, with 90% power, the number of needed clusters is 4 for Group 1 (control) (totaling 80 subjects) and 4 for Group 2 (treatment) (totaling 80 subjects). The corresponding required number of events is 56 in Group 1 and 40 in Group 2 (totaling 96 events).

---

**References**

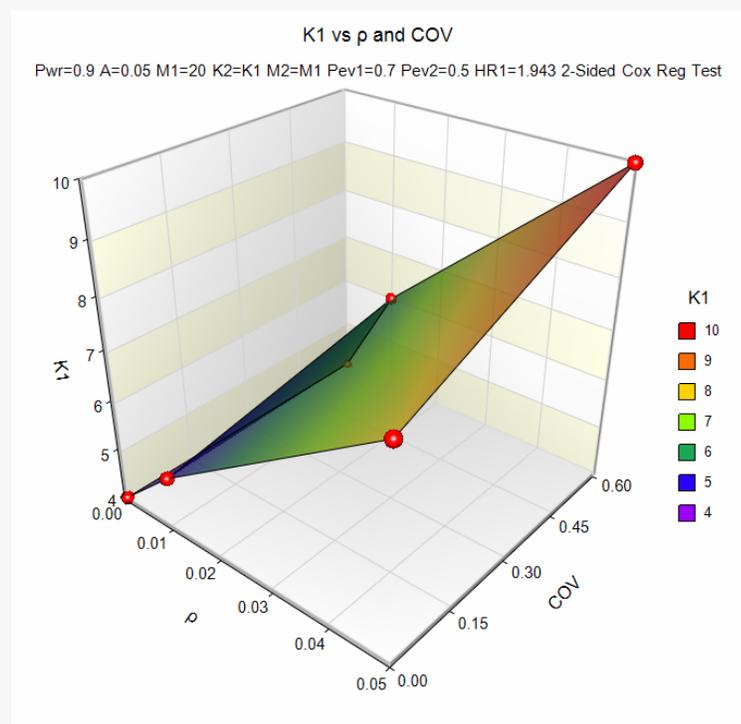
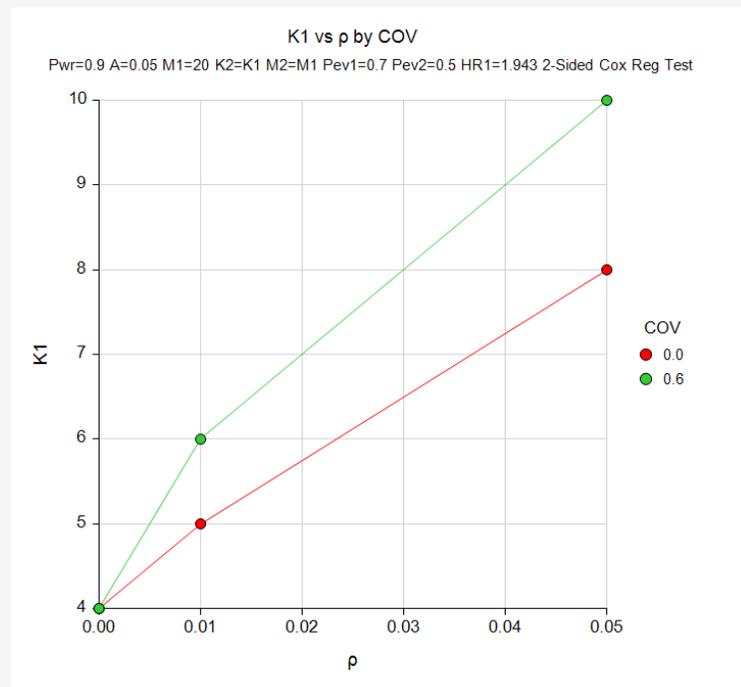
---

- Chow, S.C., Shao, J., Wang, H., and Lokhnygina, Y. 2018. Sample Size Calculations in Clinical Research, 3rd Edition. Chapman & Hall/CRC. Boca Raton, FL. Pages 86-88.
- Donner, A. and Klar, N. 2000. Design and Analysis of Cluster Randomization Trials in Health Research. Arnold. London.
- Machin, D., Campbell, M.J., Tan, S.B, and Tan, S.H. 2018. Sample Sizes for Clinical, Laboratory, and Epidemiology Studies, 4th Edition. Wiley Blackwell.
- Schoenfeld, David A. 1983. 'Sample Size Formula for the Proportional-Hazards Regression Model', Biometrics, Volume 39, Pages 499-503.
- 

This report shows the power for each of the scenarios. Note each scenario requires three report lines.

## Plots Section

### Plots



These plots show the relationship between the number of clusters, COV, and ICC.

## Example 2 – Validation using Chow et al. (2018)

Chow, Shao, Wang, and Lokhnygina (2018) pages 162 - 163 present an example that will be used to validate this procedure. Their example presents the results of a design that uses a regular (not-cluster) randomization pattern. To complete the validation, we will modify their results by hand using the design effect (DE) value.

The parameters of their study is as follows:  $HR1 = 2.0$ ;  $Pev1 = 0.8$ ;  $Pev2 = 0.8$ ;  $\rho = 0.00$ ;  $COV = 0.0, 0.6$ ;  $M1 = M2 = 2.05$ ;  $K1 = K2 = 20$  (so  $N1 = N2 = 41$ ); and  $alpha = 0.05$ . The test is two-sided.

### 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 .....	<b>Number of Clusters</b>
Alternative Hypothesis .....	<b>Two-Sided (H1: HR ≠ 1)</b>
Power.....	<b>0.8</b>
Alpha.....	<b>0.05</b>
M1 (Average Cluster Size).....	<b>2.05</b>
K2 (Number of Clusters).....	<b>K1</b>
M2 (Average Cluster Size).....	<b>M1</b>
COV of Cluster Sizes.....	<b>0</b>
Pev1 (Probability of a Control Event).....	<b>0.8</b>
Pev2 (Probability of a Treatment Event) .....	<b>0.8</b>
HR1 (Actual Hazard Ratio to Detect).....	<b>2</b>
$\rho$ (Intraclass Correlation).....	<b>0</b>

### Output

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

#### Numeric Results

Solve For: [Number of Clusters](#)  
 Groups: 1 = Control, 2 = Treatment  
 Hypotheses: H0: HR = 1 vs. H1: HR ≠ 1

Group	Power	Sample Size Ni	Number of Events Ei	Number of Clusters Ki	Cluster Size		Probability of an Event Pevi	Hazard Ratio HR1	Intraclass Correlation $\rho$	Alpha
					Average Mi	COV				
1: Control	0.80152	41	32.8	20	2.05	0	0.8	2	0	0.05
2: Treatment		41	32.8	20	2.05	0	0.8		0	
Total		82	65.6	40						

PASS calculates the power to be 0.80152 which rounds to 0.80.

## Example 3 – Additional Validation

Continuing with the validation, we will now set  $COV = 0.6$  and  $\rho = 0.05$ . The formula for DE is

$$DE = 1 + \{[CV^2 + 1]\bar{M} - 1\}\rho$$

Plugging in the above values for COV and  $\rho$

$$DE = 1 + \{[0.36 + 1]2.05 - 1\}(0.05) = 1.0894$$

Multiplying 1.0894 times 20 gives 21.788 which rounds up to 22 for K1 and K2.

### 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 .....	<b>Number of Clusters</b>
Alternative Hypothesis .....	<b>Two-Sided (H1: HR ≠ 1)</b>
Power.....	<b>0.8</b>
Alpha.....	<b>0.05</b>
M1 (Average Cluster Size).....	<b>2.05</b>
K2 (Number of Clusters) .....	<b>K1</b>
M2 (Average Cluster Size).....	<b>M1</b>
COV of Cluster Sizes.....	<b>0.6</b>
Pev1 (Probability of a Control Event).....	<b>0.8</b>
Pev2 (Probability of a Treatment Event) .....	<b>0.8</b>
HR1 (Actual Hazard Ratio to Detect).....	<b>2</b>
$\rho$ (Intraclass Correlation).....	<b>0.05</b>

### Output

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

Numeric Results											
Solve For:		Number of Clusters									
Groups:		1 = Control, 2 = Treatment									
Hypotheses:		H0: HR = 1 vs. H1: HR ≠ 1									
Group	Power	Sample Size Ni	Number of Events Ei	Number of Clusters Ki	Cluster Size		Probability of an Event Pevi	Hazard Ratio HR1	Intraclass Correlation $\rho$	Alpha	
					Average Mi	COV					
1: Control		46	40.09	22	2.05	0.6	0.8		0.05		
2: Treatment	0.81292	46	40.09	22	2.05	0.6	0.8	2	0.05	0.05	
Total		92	80.18	44							

PASS also calculates the values of K1 and K2 to be 22. This validates the procedure.