Chapter 560

Cumulative Incidence

Introduction

This routine calculates nonparametric, maximum-likelihood estimates and confidence limits of the probability of failure (the *cumulative incidence*) for a particular cause in the presence of other causes. This is sometimes called the problem of *competing risks*. This procedure also calculates Gray's Test and the Pepe and Mori Test for comparing groups while accounting for competing risks.

An often used, though incorrect, approach is to treat all failures from causes other than that of interest as censored observations and estimate the cumulative incidence using 1 - *KM* (Kaplan-Meier estimate). The problem with this approach is that it makes the incorrect assumption that the probability of failing prior to time *t* from other causes is zero. This leads to overestimation of the cumulative incidence. This overestimation can be quite substantial if there are many failures from other causes in the data.

Technical Details

Estimation and Confidence Intervals

The following results are summarized from Marubini and Valsecchi (1996). Suppose that one of *K* mutually exclusive events may occur to a subject. These events may be failure, death, etc. When an event occurs, the time until it occurred *T* and the type of event *k* is recorded. The experiment may be terminated before any event occurs for some subjects in which case they are called *censored* observations and the time until censoring is recorded. The cause-specific hazard functions are defined as

$$h_k(t) = \lim_{\Delta t \to 0^+} \frac{\Pr\{t \le T < t + \Delta t,\, K = k | T \ge t\}}{\Delta t}, \quad k = 1, 2, ..., K$$

The overall survival function is denoted by

$$S(t) = \Pr(T > t)$$

and the probability of failing for cause *k* is denoted by

$$I_k(t) = \Pr(T \le t, k = K), \quad k = 1, 2, ..., K$$

These cause-specific failure probabilities, also known as the cumulative incidence functions, are defined as

$$I_k(t) = \int_0^t h_k(u) S(u) du, \quad k = 1, 2, ..., K$$

where

$$S(u) = \exp\left\{ - \int_0^u \left(\sum_{k=1}^{k} h_k(y) \right) dy \right\}$$

This makes use of the assumption that a subject must be event free up to time *u* to then fail of cause *k* at time *u*.

If we let d_{kj} denote the number of subjects having event *k* and time t_j, $d_j = \sum_{k=1}^{K} d_{kj}$, and n_j denote the number of subjects at risk at time t_j, the likelihood may be written as

$$L = \prod_{j=1}^{J}\left[\left(\prod_{k=1}^{K} h_{kj}^{d_{kj}}\right)\left(1-\sum_{k=1}^{K} h_{kj}\right)^{n_j - d_j}\right]$$

The ML estimate of the cause-specific hazard is

$$\hat{h}_{kj} = \frac{d_{kj}}{n_j}$$

and the estimated crude cumulative incidence for event *k* is

$$\hat{I}_k(t) = \sum_{j|t_j \le t} \hat{S}(t_{j-1})\frac{d_{kj}}{n_j}$$

where $\hat{S}(t_{j-1})$ is the usual Kaplan-Meier estimate of survival until time t_{j-1}.

The variance of the crude cumulative incidence is estimated by

$$Var\left[\hat{I}_k(t_j)\right] = \sum_{i=1}^{j}\left\{\left[\hat{I}_k(t_j) - \hat{I}_k(t_i)\right]^2 \frac{d_i}{n_i(n_i - d_i)}\right\} + \sum_{i=1}^{j}\left[\hat{S}(t_{i-1})\right]^2\left(\frac{n_i - d_{ki}}{n_i}\right)\left(\frac{d_{ki}}{n_i^2}\right)$$
$$- 2\sum_{i=1}^{j}\left\{\left[\hat{I}_k(t_j) - \hat{I}_k(t_i)\right]\hat{S}(t_{i-1})\frac{d_{ki}}{n_i^2}\right\}$$

Finally, using the above estimate cumulative incidence and its estimated variance, approximate $100(1-\alpha)\%$ confidence intervals may be calculated using

$$\exp\left\{\log\left[\hat{I}_k(t)\right] \pm z_{1-\alpha/2}\frac{\sqrt{Var\left[\hat{I}_k(t)\right]}}{\hat{I}_k(t)}\right\}$$

This expression guarantees that the resulting values will be between zero and one.

Hypothesis Tests

The following results are summarized from Pintilie (2006). In the absence of competing risks, the logrank test is often used to compare the 2 or more groups with respect to a particular outcome. If competing risks are present, the logrank test can still be used to compare the cause-specific hazards, but it does not account for competing risks. The logrank test for cause-specific hazards is conducted by assigning a "1" to the event of interest and a "0" to all competing risk events and censored observations. Thus, the logrank test effectively ignores the competing risks and focuses only on the event of interest. This can result in very misleading results if not interpreted correctly. Please see the Kaplan-Meier (Logrank tests) procedure and documentation for more information about the logrank test; it will not be repeated here.

In the presence of competing risks, you should use Gray's test for *K* samples or Pepe and Mori's test for 2 samples as a direct approach to compare groups with respect to the event of interest while accounting for the effects of the other competing risk events. Gray's test compares the weighted averages of the subdistribution hazards across groups for the event of interest. Pepe and Mori's test compares the cumulative incidence functions (CIF's) directly.

Both are appropriate for comparing groups with respect to an outcome of interest while accounting for competing risks.

Gray's Test

Gray's test is a *K*-sample test that was introduced by Gray (1988). It compares the weighted averages of the subdistribution hazards across groups for the event of interest. The null hypothesis is that there is no difference among the *K* groups. The test is based on the $K-1$ score statistics

$$Z_K(\tau) = \int_{t=0}^{\tau} W_K(t)\{\gamma_K(t) - \gamma_0(t)\}dt,$$

where τ is taken to be the largest observed time in the study and $\gamma(t)$ is the subdistribution hazard with

$$\gamma_K(t) = \frac{i_K(t)}{1 - I_K(t)}.$$

In these equations $i_K(t)$ is the subdensity and $I_K(t)$ is the subdistribution function in group *K* for the event of interest, $W_K(t)$ is a weight function, and $\gamma_0(t)$ is the subdistribution hazard for all groups taken together.

If we let *Z* represent the vector of $K-1$ score statistics and represent the variance-covariance matrix for *Z*, the test statistic, Q, is given by

$$Q_{Gray} = Z'\Sigma^{-1}Z$$

For large samples, Q_{Gray} is approximately distributed as a chi-squared random variable with $K-1$ degrees of freedom. The formula for Σ is complex and is not given here. See Pintilie (2006) and Gray (1988) for more information about this test.

Pepe and Mori's Test

Pepe and Mori's test is a 2-sample test that was introduced by Pepe and Mori (1993). This test compares the cumulative incidence functions (CIF's) directly for the event of interest. The null hypothesis is that there is no difference between the 2 groups.

For groups 1 and 2, respectively, define $\hat{I}_1(t)$ and $\hat{I}_2(t)$ as the cumulative incidence function estimates for the event of interest, $\hat{I}_{cr,1}(t)$ and $\hat{I}_{cr,2}(t)$ as the cumulative incidence function estimates for competing risks, $\hat{C}_1(t)$ and $\hat{C}_2(t)$ as the Kaplan-Meier estimate of survival where the events are defined as observations that are either competing risk events or censored, and N_1 and N_2 as the total sample sizes. For both groups, also define $n_1(t_j)$ and $n_2(t_j)$ as the number at risk at time t_j, $d_1(t_j)$ and $d_2(t_j)$ as the number of events of interest or competing risk events at time t_j, and $d_{ev,1}(t_j)$ and $d_{ev,2}(t_j)$ as the number of events of interest at time t_j. Further define $W(t)$ as a weight function that decreases with time.

For ordered unique times representing either events or censored observations, $t_1, t_2, \ldots, t_n, ,$ Pepe and Mori's test statistic is given by

$$Q_{PM} = \frac{s^2}{\hat{\sigma}^2},$$

where

$$s = \sqrt{\frac{N_1 N_2}{N_1+N_2} \sum_{all\ t_j} W(t_j)\left[\hat{I}_1(t_j) - \hat{I}_2(t_j)\right](t_{j+1} - t_j)}$$

$$\hat{\sigma}^2 = \frac{N_1 N_2(\hat{\sigma}_1^2 + \hat{\sigma}_2^2)}{N_1+N_2}$$

with

$$W(t_j) = \frac{(N_1+N_2)\hat{C}_1(t_{j-1})\hat{C}_2(t_{j-1})}{N_1\hat{C}_1(t_{j-1}) + N_2\hat{C}_2(t_{j-1})}.$$

$$\hat{\sigma}_K^2 = \sum_{all\ t_j} \frac{\left[v_{1,K}(t_j) - \hat{I}_{cr,K}(t_j)v_{2,K}(t_j)\right]^2 d_{ev,K}(t_j) + v_{3,K}(t_j)^2\left(d_K(t_j) - d_{ev,K}(t_j)\right)}{n_K(t_j)(n_K(t_j)-1)}$$

$$v_{1,K}(t_j) = \sum_{t_k \geq t_j} W(t_k)\,(t_{k+1} - t_k)\left(1 - \hat{I}_K(t_k)\right)$$

$$v_{2,K}(t_j) = \sum_{t_k \geq t_j} W(t_k)\,(t_{k+1} - t_k)$$

$$v_{3,K}(t_j) = \sum_{t_k \geq t_j} W(t_k)\,(t_{k+1} - t_k)\hat{I}_K(t_k)$$

For large samples, Q_{PM} is approximately distributed as a chi-squared random variable with 1 degree of freedom. See Pintilie (2006) and Pepe and Mori (1993) for more information about this test.

Data Structure

This routine requires at least two variables: one containing the elapsed time values and another containing the type of event. Optional variables include a group identification variable and a frequency variable.

Marubini and Valsecchi (1996) include an example of hypothetical data consisting of two treatment groups and two events. The events are local relapse (1) and distant metastases (2). Censored observations are represented with a zero. Information is available on 35 subjects in each group. These data are stored in the Marubini database. The table below shows the data.

Marubini dataset

Time	Treatment	Event
1	A	1
13	A	1
17	A	1
30	A	1
34	A	1
41	A	1
78	A	1
100	A	1
119	A	1
169	A	1
1	A	2
6	A	2
8	A	2
13	A	2
13	A	2
15	A	2
.	.	.
.	.	.
.	.	.

Time	Treatment	Event
7	B	1
16	B	1
16	B	1
20	B	1
39	B	1
49	B	1
56	B	1
73	B	1
93	B	1
113	B	1
1	B	2
2	B	2
4	B	2
6	B	2
8	B	2
9	B	2
.	.	.
.	.	.
.	.	.

Procedure Options

This section describes the options available in this procedure.

Variables Tab

This panel specifies the variables used in the analysis.

Time Variable

Time Variable

This variable contains the times for each subject. These are elapsed times. If your data are event dates, you must subtract a starting date so that the values are elapsed time. You can scale your data in days, months, or years by dividing by the appropriate constant.

Note that negative time values are treated as missing values. Zero time values are replaced by the value in the Zero Time Replacement option.

Zero Replacement Value

Under normal conditions, a respondent beginning the study is alive and cannot die until after some small period of time has elapsed. Hence, a time value of zero is not defined and is ignored (treated as a missing value). If a zero time value does occur in the database, it is replaced by this positive amount. If you do not want zero time values replaced, enter a “0.0” here.

This option would be used when a “zero” on the database does not actually mean zero time. Instead, it means that the response occurred before the first reading was made and so the actual survival time is only known to be less than one.

Event (Censor) Type Variable

Event (Censor) Variable

Specify the variable containing an identifier for the event type of each observation. If the observation was censored, indicate that using another identifier.

The meaning of these event-type identifiers is designated in the Event Types and Censored Types boxes.

Event Types

This box lists the values of the Type Variable that are to be designated as events. These values may be letters or numbers. For a competing risks analysis, at least two events must be present on the database.

The event may be any occurrence of interest such as failure, death, or recovery. For example, in heart surgery, the events might be death because of heart failure or death for other reasons (accident, cancer, etc.). In this case, two events would be used.

Censor Types

This box lists the values of the Type Variable that are to be designated as being censored. These values may be letters or numbers. Usually, at least one censor value is used.

All of the censor-type values are interpreted as meaning that the observation is right censored. A right censored observation is withdrawn from the study before an event occurs. Hence, you know that the event will occur after the given length of time, but you do not know when it will occur. For example, the study may end before the patient has died.

Group Variable (Optional)

Group Variable

An optional categorical grouping variable may be specified. Separate curves are calculated for each group. This variable is required for hypothesis testing.

Frequency Variable (Optional)

Frequency Variable

This optional variable gives the count, or frequency, of the time value. Frequency values must be positive integers. When omitted, each row has a frequency of one.

This variable is often used to indicate the number of CENSORED values at the end of a study. It may also be used to indicate TIES for event-type data.

Reports Tab

The following options control which reports are displayed and the format of those reports.

Events to Report

Events Reported On

This option limits the reports to those event-types that are of primary interest. Separate sets of reports and plots are generated for each event type listed here. Enter ALL to have all event types reported on.

Note that this option does not change the analysis, just the reports.

Descriptive Reports

Data Summary Report - Cumulative Survival Summary Report

These options specify whether to display the corresponding reports.

Summary Report Times

Specify a list of times that are reported on in the summary reports. The regular reports contain output for each time on the database. These may not represent the most useful time values. For example, in a study that lasts two years, you may want a summary report for every six months.

To specify the time values, use numbers separated by commas or blanks. You may specify a sequence of values with a colon, putting the increment inside parentheses. For example: 5:25(5) means 5,10,15,20,25. Avoid 0 and negative numbers.

Use (10) alone to specify ten, equal-spaced values between zero and the maximum (zero not included).

Confidence Level

This value is used in all confidence interval calculations. Typical confidence levels are 90%, 95%, and 99%, with 95% being the most common.

Hypothesis Tests

Alpha

This is the significance level used in all hypothesis tests. A value of 0.05 is most commonly used.

Hypothesis Tests – Compare Groups While Accounting for Competing Risks

Gray's Test (K Groups)

This test compares *K* groups while accounting for competing risks. Gray's test compares the weighted average of the subdistribution hazards across groups for the event of interest. This test is appropriate if you wish to compare groups and account for competing risks. A separate test is conducted for each event type.

Pepe and Mori's Test (K Groups)

This test compares two groups while accounting for competing risks. Pepe and Mori's test compares the cumulative incidence functions (CIF's) directly between the two groups for the event of interest. This test is appropriate if you wish to compare groups and account for competing risks. A separate test is conducted for each event type.

Hypothesis Tests – Compare the Cause-Specific Hazards among Groups

Logrank Test (K Groups)

This test compares K groups without accounting for competing risks. This test is used to compare the cause-specific hazards. A separate test is conducted for each event type.

All competing risk events are considered censored by this test. This is the same logrank test that is calculated using the Kaplan-Meier (Logrank Tests) procedure if the event of interest is coded as failures (e.g. "1") and censored observations and competing risks are all coded as being censored (e.g. "0").

Report Options Tab

These options control the format of reports.

Report Options

Variable Names

This option lets you select whether to display only variable names, variable labels, or both.

Value Labels

This option lets you select whether to display only values, value labels, or both. Use this option if you want to automatically attach labels to the values of the group variable (like 1=Yes, 2=No, etc.). See the section on specifying *Value Labels* elsewhere in this manual.

Report Options – Decimal Places

Times – P-Values

These options specify the number of decimal places shown for the corresponding values. These do not affect the precision of the calculations… only the output formatting.

Report Options – Title

Report Title

This option specifies a title to appear at the top of each page.

Plots Tab

The following options control which plots are displayed and the format of those plots.

Select Plots

Individual Incidence Plots – Combined Survival Plots

Specify whether to display each of these plots.

Plot Format Buttons

Click to change the format of the plots.

Select Plots – Plot Arrangement

Two Plots Per Line

Specify whether to display one or two plots per line. Choosing two plots forces the plots to be smaller so that two will fit across a page.

Storage Tab

These options control the storage of information back to the dataset for further use.

Data Storage Columns

Group - Product Limit

Each of these options let you specify columns on the dataset to which the corresponding data are automatically stored. Warning: existing data are replaced, so make sure that the columns you select are empty.

Note that no attempt is made to store the time values in their original order. That's why you have to store the Group and Event Type to identify the incidence values.

Example 1 – Cumulative Incidence

This section presents an example of how to generate cumulative incidence reports. The data used were shown above and are found in the Marubini dataset.

You may follow along here by making the appropriate entries or load the completed template **Example 1** by clicking on Open Example Template from the File menu of the Cumulative Incidence window.

1 Open the Marubini dataset.

- From the **File** menu of the NCSS Data window, select **Open Example Data**.
- Click on the file **Marubini.NCSS**.
- Click **Open**.

2 Open the Cumulative Incidence window.

- Using the Analysis menu or the Procedure Navigator, find and select the **Cumulative Incidence** procedure.
- On the menus, select **File**, then **New Template**. This will fill the procedure with the default template.

3 Specify the variables.

- On the Cumulative Incidence window, select the **Variables tab**.
- Enter **Time** in the **Time Variable** box.
- Enter **Event** in the **Event (Censor) Variable** box.
- Enter **1 2** in the **Event Types** box.
- Enter **0** in the **Censor Types** box.
- Enter **Treatment** in the **Group Variable** box.

4 Specify the reports.

- On the Cumulative Incidence window, select the **Reports tab**.
- Enter **1** in the **Events Reported On** box.
- Leave all **Descriptive Reports** checkboxes checked.
- Enter **25 50 75 100 125 150 175 200** in the **Summary Report Times** box.
- Enter **95** in the **Confidence Level** box.
- Leave **Gray's Test** and **Pepe and Mori's Test** boxes checked.
- Check **Logrank Test**

5 Specify the plots.

- On the Cumulative Incidence window, select the **Plots tab**.
- Check the **Confidence Limits** box after clicking on the format button for each plot.
- Check the **Two Plots Per Line** box.

6 Run the procedure.

- From the Run menu, select **Run Procedure**. Alternatively, just click the green **Run** button.

Data Summary Section

Data Summary Section. Treatment = A

Type of Observation	Values	Count	Minimum	Maximum
Censored	5	5	34	207
Event = 1	10	10	1	169
Event = 2	18	20	1	240
Total	33	35	1	240

This section of the report displays information about the database. It is especially useful to allow you to check for obvious data-entry errors. The Values column gives the number of unique time values. The Count column uses the values in the Frequency Variable when it was used.

Cumulative Incidence Detail Report

Cumulative Incidence Detail Report for Event = 1 and Treatment = A

Time	Number At Risk	Events of Type 1	Events of All Types	Cumulative Incidence	Lower 95% C.L. Cum. Inc.	Upper 95% C.L. Cum. Inc.	Standard Error of Cum. Inc.	1 - Product Limit
1.0	35	1	2	0.0286	0.0041	0.1972	0.0282	0.0571
6.0	33	0	1	0.0286	0.0041	0.1972	0.0282	0.0857
8.0	32	0	1	0.0286	0.0041	0.1972	0.0282	0.1143
13.0	31	1	3	0.0571	0.0149	0.2195	0.0392	0.2000
15.0	28	0	1	0.0571	0.0149	0.2195	0.0392	0.2286
.	.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.	.
240.0	1	0	1	0.3144	0.1871	0.5285	0.0833	1.0000

This report displays the cumulative incidence values along with their confidence intervals and standard errors.

Time

This is the time value, t_j, being reported on. These values are from the dataset being analyzed. Note that tied values are combined.

Number at Risk

This is the number of individuals at risk, n_j, just before time t_j.

Events of Type 1

This is the number of events of the type indicated (in this report, the type is 1), d_{kj}, that occurred at t_j.

Events of All Types

This is the number of events of all types, d_{kj}, that occurred at time t_j. Note that censored observations have a zero in this column.

Cumulative Incidence

This is the cumulative incidence, $\hat{I}_k(t_j)$. This is the cumulative probability of event *k* up through the current time value, accounting for all other events. The formula used to calculate this value was given in the Technical Details section earlier in this chapter.

Lower and Upper 95% C.L. Cum. Inc.

These are confidence limits for the above cumulative incidence.

Standard Error of Cum. Inc.

This is the estimated standard error of the cumulative incidence. It is calculated using $\sqrt{Var\left[\hat{I}_k\left(t_j\right)\right]}$.

1 – Product Limit

This is one minus the Kaplan-Meier product limit estimate. This is a cumulative incidence measure calculated assuming that there are no other possible events other than the event of interest.

Cumulative Survival Detail Report

Cumulative Survival Detail Report for Event = 1 and Treatment = A

Time	Number At Risk	Events of Type 1	Events of All Types	Cumulative Survival	Lower 95% C.L. Cum. Surv.	Upper 95% C.L. Cum. Surv.	Standard Error of Cum. Surv.	Product Limit
1.0	35	1	2	0.9714	0.8028	0.9959	0.0282	0.9429
6.0	33	0	1	0.9714	0.8028	0.9959	0.0282	0.9143
8.0	32	0	1	0.9714	0.8028	0.9959	0.0282	0.8857
13.0	31	1	3	0.9429	0.7805	0.9851	0.0392	0.8000
15.0	28	0	1	0.9429	0.7805	0.9851	0.0392	0.7714
.	.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.	.
240.0	1	0	1	0.6856	0.4715	0.8129	0.0833	0.0000

This report displays the cumulative survival values along with their confidence intervals and standard errors. The cumulative survival values are equal to one minus the cumulative incidence values.

Cumulative Incidence Summary Report

Cumulative Incidence Summary Report for Event = 1 and Treatment = A

Time	Number At Risk	Events of Type 1	Events of All Types	Cumulative Incidence	Lower 95% C.L. Cum. Inc.	Upper 95% C.L. Cum. Inc.	Standard Error of Cum. Inc.	1 - Product Limit
25.0	26	0	0	0.0857	0.0290	0.2529	0.0473	0.2571
50.0	18	0	0	0.1727	0.0834	0.3578	0.0642	0.4623
75.0	15	0	0	0.1727	0.0834	0.3578	0.0642	0.4940
100.0	10	1	1	0.2402	0.1308	0.4411	0.0745	0.6964
125.0	8	0	0	0.2739	0.1564	0.4797	0.0783	0.7301
150.0	5	0	0	0.2739	0.1564	0.4797	0.0783	0.7976
175.0	3	0	0	0.3144	0.1871	0.5285	0.0833	0.8786
200.0	2	0	0	0.3144	0.1871	0.5285	0.0833	0.9190

This report displays the cumulative incidence values at designated time values. All definitions are the same as in the Cumulative Incidence report. The number at risk and the numbers of events are for the last actual time value in the data before the current time value. For example, the values shown for the time of 75.0 are actually the values for the time of 63.0 on the database since 63.0 is the largest time value before 75.0.

Cumulative Survival Summary Report

Cumulative Survival Summary Report for Event = 1 and Treatment = A

Time	Number At Risk	Events of Type 1	Events of All Types	Cumulative Survival	Lower 95% C.L. Cum. Surv.	Upper 95% C.L. Cum. Surv.	Standard Error of Cum. Surv.	Product Limit
25.0	26	0	0	0.9143	0.7471	0.9710	0.0473	0.7429
50.0	18	0	0	0.8273	0.6422	0.9166	0.0642	0.5377
75.0	15	0	0	0.8273	0.6422	0.9166	0.0642	0.5060
100.0	10	1	1	0.7598	0.5589	0.8692	0.0745	0.3036
125.0	8	0	0	0.7261	0.5203	0.8436	0.0783	0.2699
150.0	5	0	0	0.7261	0.5203	0.8436	0.0783	0.2024
175.0	3	0	0	0.6856	0.4715	0.8129	0.0833	0.1214
200.0	2	0	0	0.6856	0.4715	0.8129	0.0833	0.0810

This report displays the survival values at designated time values. All definitions are the same as in the Cumulative Survival report. The number at risk and the numbers of events are for the last actual time value in the data before the current time value. For example, the values shown for the time of 75.0 are actually the values for the time of 63.0 on the database since 63.0 is the largest time value before 75.0.

Gray’s Test Report

Gray's Test
This test compares K groups while accounting for competing risks using the subdistribution hazards.

Event Type	Chi-Square Value	DF	P-Value	Reject H0 at α = 0.05?
Event = 1	0.2225	1	0.63711	No

This report displays the results of Gray’s test, which accounts for competing risks. With a p-value of 0.63711, we would fail to reject the null hypothesis of no difference between the two groups for event 1.

Pepe and Mori’s Test Report

Pepe and Mori’s Test
This test compares 2 groups while accounting for competing risks using the cumulative incidence functions (CIF's).

Event Type	Chi-Square Value	DF	P-Value	Reject H0 at α = 0.05?
Event = 1	0.0575	1	0.81043	No

This report displays the results of Pepe and Mori’s test, which accounts for competing risks. With a p-value of 0.81043, we would fail to reject the null hypothesis of no difference between the two groups for event 1. This is consistent with the results of Gray’s test.

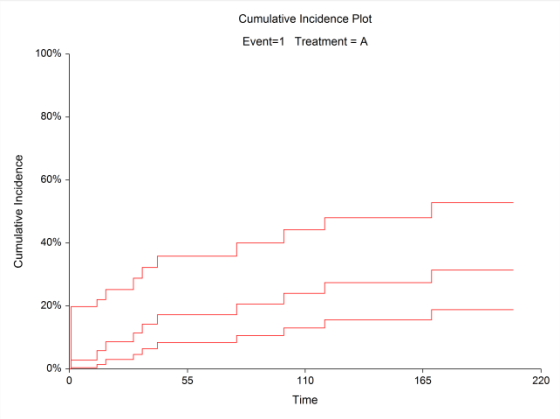
Logrank Test Report

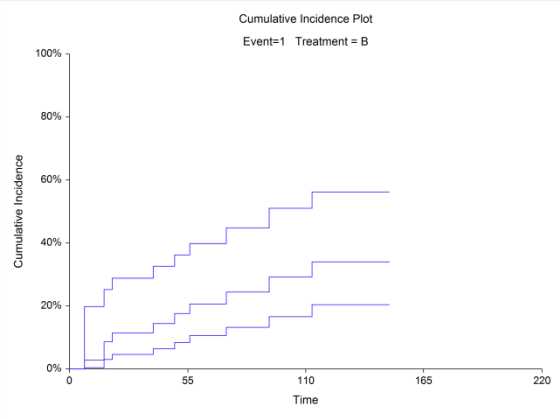
Logrank Test
This test compares the cause-specific hazards for K groups without accounting for competing risks. All competing risk events are considered censored by this test.

Event Type	Chi-Square Value	DF	P-Value	Reject H0 at α = 0.05?
Event = 1	2.0392	1	0.15329	No

This report displays the results of the logrank test, which does not account for competing risks. With a p-value of 0.15329, we would fail to reject the null hypothesis of no difference between the two groups for event 1.

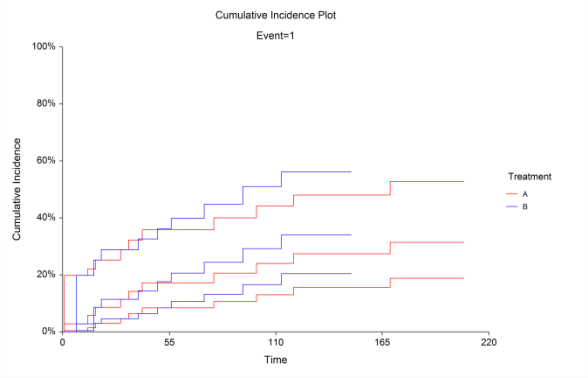
Individual Cumulative Incidence Plots





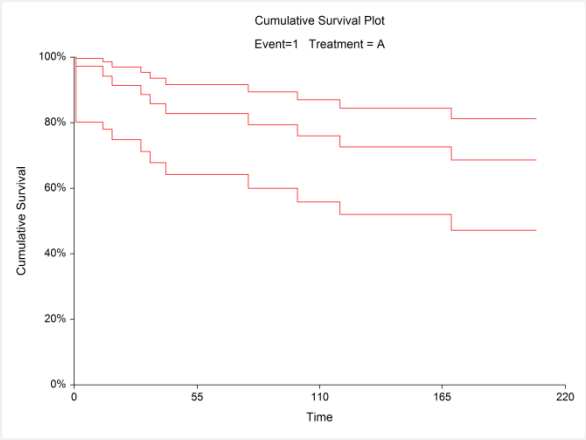
These plots show the cumulative incidence curve surrounded by the 95% confidence interval.

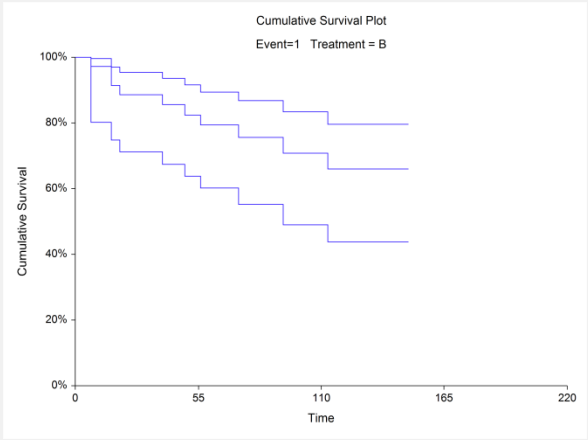
Combined Cumulative Incidence Plot



This plot shows the cumulative incidence for each of the two groups.

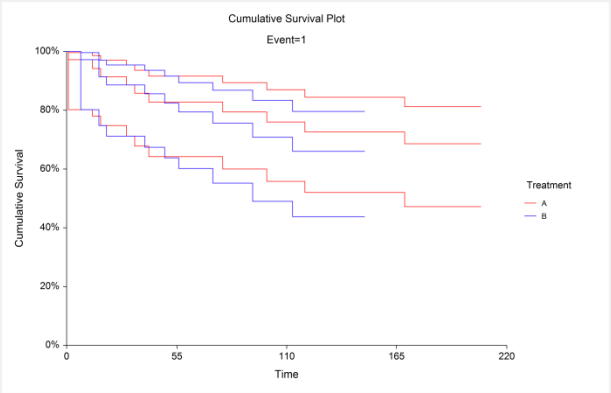
Individual Cumulative Survival Plots





These plots display the cumulative survival values and confidence intervals for each group, one group per plot.

Combined Cumulative Survival Plot



This plot displays the cumulative survival values for both groups.

Example 2 – Validation of Cumulative Incidence Estimation Values and Gray's Test using Marubini and Valsecchi (1996)

Marubini and Valsecchi (1996) reported cumulative incidence values calculated for the data in the Marubini dataset. For event 1 and group B at Time = 16 they calculated the cumulative incidence as 0.08571, with a standard error of 0.04732 and confidence limits of 0.02905 to 0.25292. They also report the Gray's test statistic for event 1 to be 0.222 with a p-value of 0.634.

You may follow along here by making the appropriate entries or load the completed template **Example 2** by clicking on Open Example Template from the File menu of the Cumulative Incidence window.

1 Open the Marubini dataset.

- From the **File** menu of the NCSS Data window, select **Open Example Data**.
- Click on the file **Marubini.NCSS**.
- Click **Open**.

2 Open the Cumulative Incidence window.

- Using the Analysis menu or the Procedure Navigator, find and select the **Cumulative Incidence** procedure.
- On the menus, select **File**, then **New Template**. This will fill the procedure with the default template.

3 Specify the variables.

- On the Cumulative Incidence window, select the **Variables tab**.
- Enter **Time** in the **Time Variable** box.
- Enter **Event** in the **Event (Censor) Variable** box.
- Enter **1 2** in the **Event Types** box.
- Enter **0** in the **Censor Types** box.
- Enter **Treatment** in the **Group Variable** box.

4 Specify the reports.

- On the Cumulative Incidence window, select the **Reports tab**.
- Enter **1** in the **Events Reported On** box.
- Check only the **Cumulative Incidence Detail Report**.
- Enter **95** in the **Confidence Level** box.
- Check only **Gray's Test**.

5 Specify the plots.

- On the Cumulative Incidence window, select the **Plots tab**.
- Deselect all plots.

6 Run the procedure.

- From the Run menu, select **Run Procedure**. Alternatively, just click the green **Run** button.

Output

Cumulative Incidence Detail Report for Event = 1 and Treatment = B

Time	Number At Risk	Events of Type 1	Events of All Types	Cumulative Incidence	Lower 95% C.L. Cum. Inc.	Upper 95% C.L. Cum. Inc.	Standard Error of Cum. Inc.	1 - Product Limit
1.0	35	0	1	0.0000	0.0000	0.0000	0.0000	0.0286
2.0	34	0	1	0.0000	0.0000	0.0000	0.0000	0.0571
4.0	33	0	1	0.0000	0.0000	0.0000	0.0000	0.0857
6.0	32	0	1	0.0000	0.0000	0.0000	0.0000	0.1143
7.0	31	1	1	0.0286	0.0041	0.1972	0.0282	0.1429
8.0	30	0	1	0.0286	0.0041	0.1972	0.0282	0.1714
9.0	29	0	1	0.0286	0.0041	0.1972	0.0282	0.2000
10.0	28	0	1	0.0286	0.0041	0.1972	0.0282	0.2286
13.0	27	0	1	0.0286	0.0041	0.1972	0.0282	0.2571
16.0	26	2	2	**0.0857**	**0.0290**	**0.2529**	**0.0473**	0.3143
17.0	24	0	3	0.0857	0.0290	0.2529	0.0473	0.4000
18.0	21	0	2	0.0857	0.0290	0.2529	0.0473	0.4571
20.0	19	1	1	0.1143	0.0454	0.2874	0.0538	0.4857
27.0	18	0	1	0.1143	0.0454	0.2874	0.0538	0.5143
29.0	17	0	1	0.1143	0.0454	0.2874	0.0538	0.5429
34.0	16	0	0	0.1143	0.0454	0.2874	0.0538	0.5429
39.0	15	1	2	0.1448	0.0643	0.3259	0.0599	0.6038
49.0	13	1	1	0.1752	0.0847	0.3627	0.0650	0.6343
50.0	12	0	1	0.1752	0.0847	0.3627	0.0650	0.6648
56.0	11	1	1	0.2057	0.1063	0.3983	0.0693	0.6952
60.0	10	0	0	0.2057	0.1063	0.3983	0.0693	0.6952
63.0	9	0	0	0.2057	0.1063	0.3983	0.0693	0.6952
69.0	8	0	1	0.2057	0.1063	0.3983	0.0693	0.7333
73.0	7	1	1	0.2438	0.1329	0.4473	0.0755	0.7714
76.0	6	0	1	0.2438	0.1329	0.4473	0.0755	0.8095
78.0	5	0	0	0.2438	0.1329	0.4473	0.0755	0.8095
93.0	4	1	1	0.2914	0.1666	0.5098	0.0832	0.8571
110.0	3	0	1	0.2914	0.1666	0.5098	0.0832	0.9048
113.0	2	1	1	0.3390	0.2045	0.5621	0.0874	0.9524
149.0	1	0	0	0.3390	0.2045	0.5621	0.0874	0.9524

Gray's Test

This test compares K groups while accounting for competing risks using the subdistribution hazards.

Event Type	Chi-Square Value	DF	P-Value	Reject H0 at α = 0.05?
Event = 1	**0.2225**	1	**0.63711**	No

The cumulative incidence value (0.0857), the standard error (0.0473), and confidence limits (0.0290,0.2529) for Time = 16 match those reported by Marubini and Valsecchi (1996). The test statistic (0.2225) and p-value (0.63711) match as well with slight differences due to rounding.

Example 3 – Validation of Gray's Test, Pepe and Mori's Test, and the Logrank Test using Pintilie (2006)

Pintilie (2006) uses the R package "cmprsk" (written by Bob Gray, the author of Gray (1998)) to compute Gray's test for a Follicular cell lymphoma dataset. The author reports the test statistics and p-values obtained from R "cmprsk" to be 2.631747 and 1.047464e-01, respectively, for event 1 and 39.309879 and 3.616071e-10, respectively, for event 2.

For event 1, Pintilie (2006) further computes Pepe and Mori's test (test statistic = 2.617415, p-value = 0.1056965) and the logrank test (test statistic = 6.941658, p-value = 0.008421103).

This example will demonstrate that **NCSS** gets the same results.

You may follow along here by making the appropriate entries or load the completed template **Example 3** by clicking on Open Example Template from the File menu of the Cumulative Incidence window.

1 Open the Follic dataset.

- From the **File** menu of the NCSS Data window, select **Open Example Data**.
- Click on the file **Follic.NCSS**.
- Click **Open**.

2 Open the Cumulative Incidence window.

- Using the Analysis menu or the Procedure Navigator, find and select the **Cumulative Incidence** procedure.
- On the menus, select **File**, then **New Template**. This will fill the procedure with the default template.

3 Specify the variables.

- On the Cumulative Incidence window, select the **Variables tab**.
- Enter **dftime** in the **Time Variable** box.
- Enter **event** in the **Event (Censor) Variable** box.
- Enter **1 2** in the **Event Types** box.
- Enter **0** in the **Censor Types** box.
- Enter **group** in the **Group Variable** box.

4 Specify the reports.

- On the Cumulative Incidence window, select the **Reports tab**.
- Enter **ALL** in the **Events Reported On** box.
- Uncheck all **Descriptive Reports** checkboxes.
- Check **Gray's Test**, **Pepe and Mori's Test**, and **Logrank Test**.

5 Specify the plots.

- On the Cumulative Incidence window, select the **Plots tab**.
- Deselect all plots.

6 Run the procedure.

- From the Run menu, select **Run Procedure**. Alternatively, just click the green **Run** button.

Output

Gray's Test
This test compares K groups while accounting for competing risks using the subdistribution hazards.

Event Type	Chi-Square Value	DF	P-Value	Reject H0 at α = 0.05?
event = 1	**2.6317**	1	**0.10475**	No
event = 2	**39.3470**	1	**0.00000**	Yes

Pepe and Mori's Test
This test compares 2 groups while accounting for competing risks using the cumulative incidence functions (CIF's).

Event Type	Chi-Square Value	DF	P-Value	Reject H0 at α = 0.05?
event = 1	**2.6174**	1	**0.10570**	No
event = 2	17.7670	1	0.00002	Yes

Logrank Test
This test compares the cause-specific hazards for K groups without accounting for competing risks. All competing risk events are considered censored by this test.

Event Type	Chi-Square Value	DF	P-Value	Reject H0 at α = 0.05?
event = 1	**6.9417**	1	**0.00842**	Yes
event = 2	66.1370	1	0.00000	Yes

The test statistics and p-values reported by **NCSS** all match Pintilie (2006) except for Gray's Test for event 2… there is a slight difference there. As was stated in the introduction, Pintilie (2006) uses the R package "cmprsk" to compute Gray's test. We reran the example from the book using the most recent version of "cmprsk", version 2.2-7, and found that the results from "cmprsk" in R match the values reported by **NCSS**.