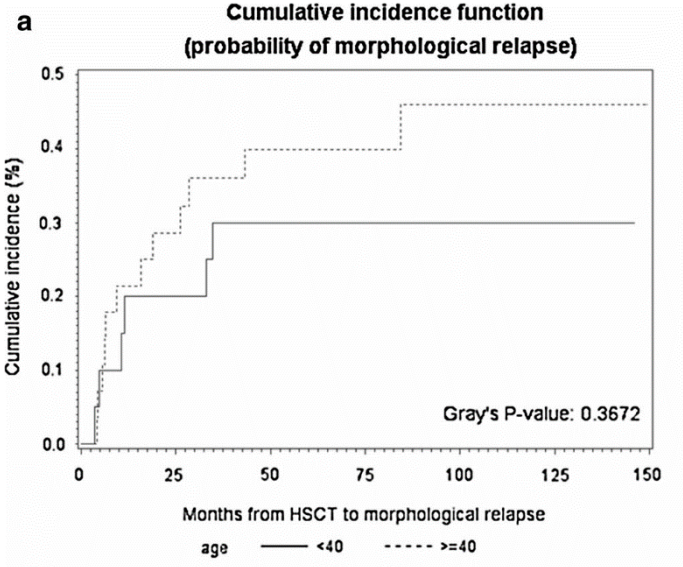
UPPSALA UNIVERSITET

# Analysis of Survival Data

## Lecture 4:
## Summary curves for competing risks
## Hypothesis testing




$$H_0 : h_1(t) = h_2(t) \quad \text{for all } t \le \tau$$

$$H_a : h_1(t) \neq h_2(t) \quad \text{for some } t \le \tau$$

Inger Persson

# Program L4

- **Summary curves for competing risks**
  - Cumulative incidence function
- **Hypothesis testing**
  - K samples tests
    - Log-rank test
    - Gehan's test
    - Gray's test
  - Stratified tests
  - Differences at a fixed point in time

# Competing risks

If a subject may fail due to one of several causes we have **competing risks**.

Example:
$X$ = survival time for patients with a certain heart disease.

Competing risk:
Death from other causes (old age, other diseases, etc.).

If one of the competing risks occurs the event of interest cannot be observed.

UPPSALA UNIVERSITET

# Competing risks vs censoring

Special cases of competing risks can be handled by random censoring.

E.g. when people in the study move abroad, or no longer want to participate.

If the competing events are independent, it is usually okay to denote these observations as censored. In these situations the censoring will be noninformative.

But if the competing events are dependent, the censoring is no longer noninformative.

# Competing risks and Kaplan-Meier

The KM method assumes that all events are independent, and thus, censors events other than the event of interest.

If there are competing risks the KM estimates are biased.

# Competing risks and Kaplan-Meier

**Example**: Studies of bone marrow transplantation and leukemia relapse.

Competing risk: death caused by transplantation complications.

Leukemia relapse and transplantation related mortality are not independent, since both events are likely related to immunologic mechanism following transplantation.

The K-M method is inappropriate for estimating the rate of relapse since it censors transplantation related mortality.

# Summary curves for competing risks

In the case of competing risks the event times and censoring times may not be independent.

Three techniques for summarizing competing risks data:

1) The complement of the Kaplan-Meier estimator, $1-\hat{S}(t)$ Not recommended.
2) The cumulative incidence function
3) Conditional probability for the competing risk

# Complement of the Kaplan-Meier estimator

Probability of experiencing the event before time $t$:

$$1-\hat{S}(t)$$

where occurrences of the competing event(s) are treated as censored observations.

This can be interpreted as the risk of experiencing the event by time $t$ if all competing risks are removed.

This is rarely of clinical interest and this estimator is not recommended.

# Cumulative incidence function

One of the competing risks occurs at $K$ distinct times, $t_1 < t_2 < \cdots < t_K$

$Y_i$ = number of individuals at risk at $t_i$ (number of individuals who have not yet experienced any of the competing risks or experience one of them at $t_i$).

$r_i$ = number of individuals who experience the event of interest at $t_i$.

$d_i$ = number of individuals who experience any of the other events at $t_i$.

# Cumulative incidence function

The **cumulative incidence function** is defined as

$$CI(t) = \begin{cases} 0 & \text{if } t < t_1 \\ \sum_{t_i \le t} \left( \prod_{j=1}^{i-1} 1 - \frac{d_j + r_j}{Y_j} \right) \frac{r_i}{Y_i} & \text{if } t_1 \le t \end{cases}$$

$$= \begin{cases} 0 & \text{if } t < t_1 \\ \sum_{t_i \le t} \hat{S}(t_{i-}) \frac{r_i}{Y_i} & \text{if } t_1 \le t \end{cases}$$

The Kaplan-Meier estimate just before $t_i$, treating any of the competing risks as an event.

# Cumulative incidence function

The cumulative incidence function estimates the probability that the event of interest occurs at or before time *t* and that it occurs before any of the competing risks.

# Example: Time to relapse

| $t_i$ | *Event* | $Y_i$ | $\hat{S}_{KM}(t)$ |
|---|---|---|---|
| 10 | R | 10 | |
| 20+ | | 9 | |
| 35 | R | 8 | |
| 40 | T | 7 | |
| 50+ | | 6 | |
| 55 | R | 5 | |
| 70 | T | 4 | |
| 71 | T | 3 | |
| 80 | R | 2 | |
| 90+ | | 1 | |

$$\hat{S}(t) = \begin{cases} 1 & \text{if } t < t_1 \\ \prod_{t_i \le t} \left(1 - \frac{d_i + r_i}{Y_i}\right) & \text{if } t_1 \le t \end{cases}$$

R = relapse

T = treatment-related death

# Example: Time to relapse

$$\hat{S}(t) = \prod_{t_i \le t} \left(1 - \frac{d_i + r_i}{Y_i}\right) \qquad \hat{S}(10) = \left(1 - \frac{d_{10}}{Y_{10}}\right) = \left(1 - \frac{1}{10}\right) = 0.9$$

$$\hat{S}(35) = \prod_{t_i \le 35} \left(1 - \frac{d_i + r_i}{Y_i}\right) = 0.9\left(1 - \frac{1}{8}\right) = 0.7875$$

$$\hat{S}(40) = \prod_{t_i \le 40} \left(1 - \frac{d_i + r_i}{Y_i}\right) = 0.7875\left(1 - \frac{1}{7}\right) = 0.675$$

$$\hat{S}(55) = \prod_{t_i \le 55} \left(1 - \frac{d_i + r_i}{Y_i}\right) = 0.675\left(1 - \frac{1}{5}\right) = 0.54$$

# Example: Time to relapse

$$CI(t) = \begin{cases} 0 & \text{if } t < t_1 \\ \sum_{t_i \le t} \hat{S}(t_{i-}) \frac{r_i}{Y_i} & \text{if } t_1 \le t \end{cases}$$

$$CI(10) = 1\frac{1}{10} = 0.10$$

$$CI(35) = 0.10 + 0.9\frac{1}{8} = 0.2125$$

# Example: Time to relapse

$$CI(55) = 0.2125 + 0.675\frac{1}{5} = 0.3475$$

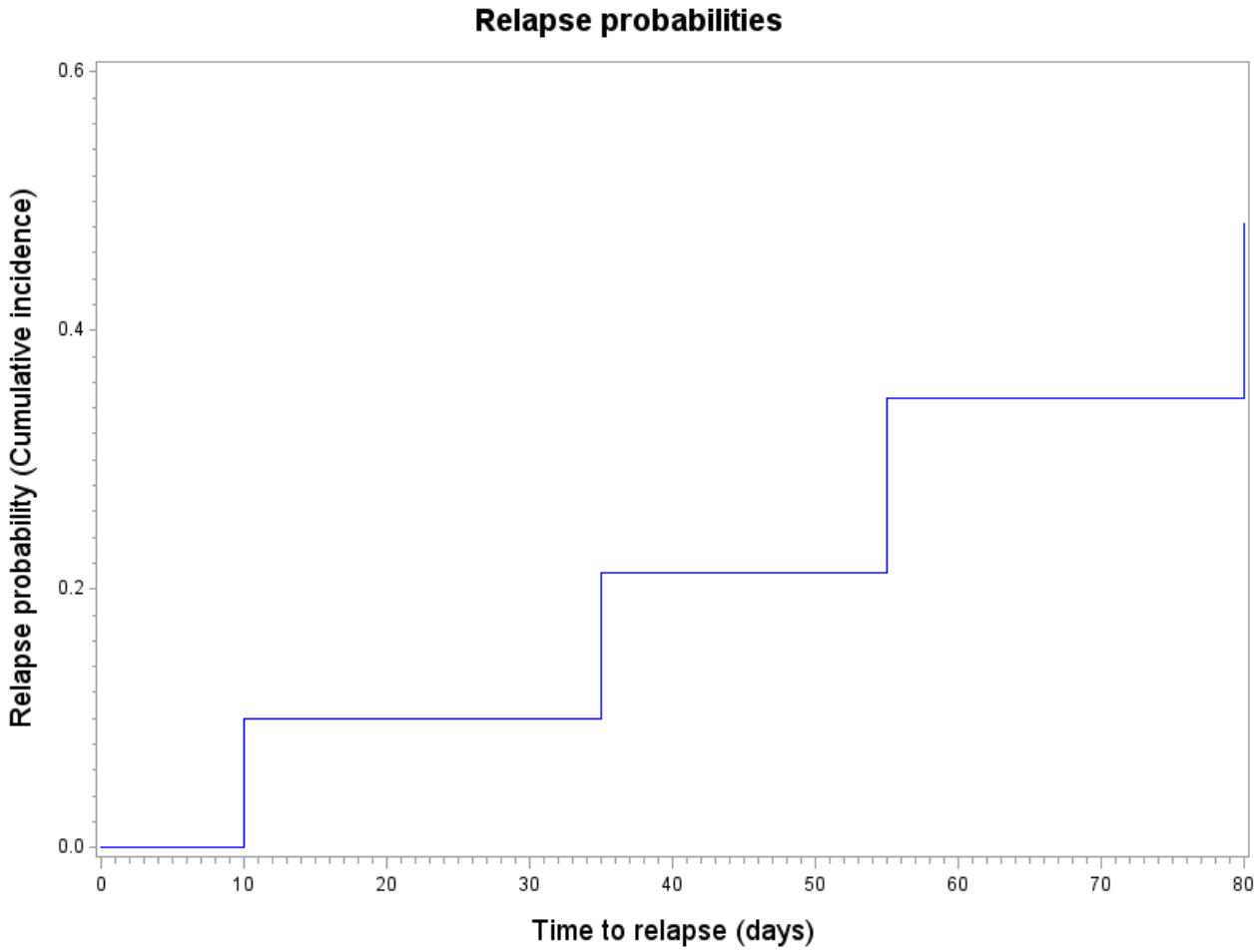
$$CI(80) = 0.3475 + 0.27\frac{1}{2} = 0.4825$$

# Example: Time to relapse

| $t_i$ | *Event* | $Y_i$ | $\hat{S}_{KM}(t)$ | $CI(t)$ |
|---|---|---|---|---|
| 10 | R | 10 | 0.9 | 0.10 |
| 20+ | | 9 | | |
| 35 | R | 8 | 0.7875 | 0.2125 |
| 40 | T | 7 | 0.675 | |
| 50+ | | 6 | | |
| 55 | R | 5 | 0.54 | 0.3475 |
| 70 | T | 4 | 0.405 | |
| 71 | T | 3 | 0.27 | |
| 80 | R | 2 | 0.135 | 0.4825 |
| 90+ | | 1 | | |

R = relapse
T = treatment-related death

# Example: Time to relapse



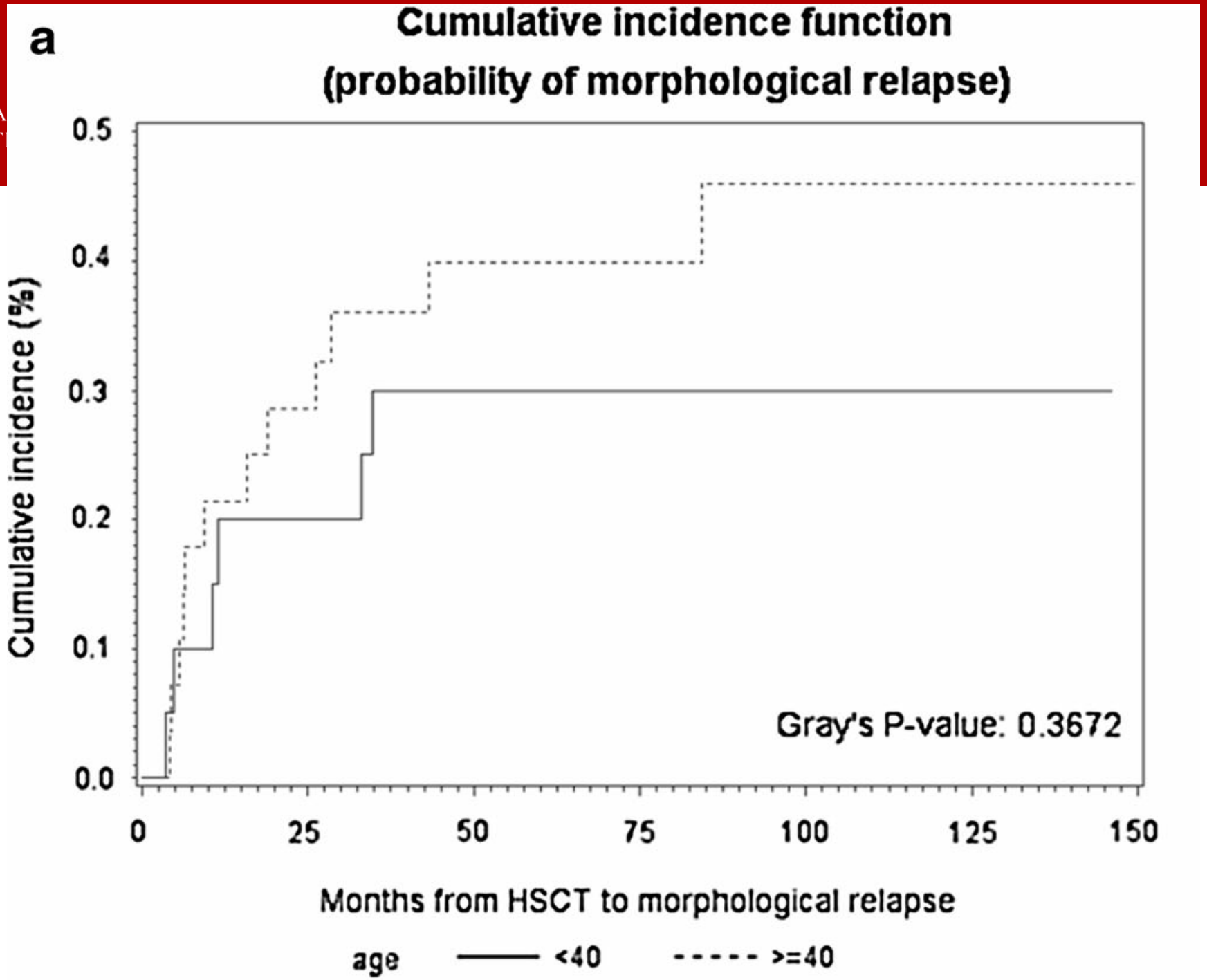
# Example: ALL outcome after stem cell transplantation

Med Oncol (2014) 31:66
DOI 10.1007/s12032-014-0066-9

ORIGINAL PAPER

## Outcome after HSCT in Philadelphia chromosome positive acute lymphoblastic leukemia in Sweden: a population-based study

E. Hulegårdh · H. Hägglund · L. Ahlberg · K. Karlsson · H. Karbach · A. Markuszewska · I. Persson · M. Åström · H. Hallböök


a
Cumulative incidence function
(probability of morphological relapse)
Cumulative incidence (%)
0.0
0.1
0.2
0.3
0.4
0.5
Gray's P-value: 0.3672
0
25
50
75
100
125
150
Months from HSCT to morphological relapse
age
<40
>=40

# Assumptions

Assumptions for the Cumulative Incidence Function (same as for the Kaplan-Meier estimator*):

- random sample(s)
- independent samples (if >1 group)
- noninformative censoring
- right censored data

**Kaplan and Meier, Nonparametric Estimation from Incomplete Observations, JASA, 1958;*

# Conditional probability function for the competing risk

The conditional probability is an estimate of the probability of the occurrence of event $K$ by time $t$, given that none of the competing risks have occurred by that time.

$$CP_K(t) = \frac{CI_K(t)}{1 - CI_{K^c}(t)}$$

Cumulative incidence for risk $K$

Cumulative incidence for all other risks lumped together.

# Program L4

- **Summary curves for competing risks**
  - Cumulative incidence function
- **Hypothesis testing**
  - K samples tests
    - Log-rank test
    - Gehan's test
    - Gray's test
  - Stratified tests
  - Differences at a fixed point in time

# K sample tests (K ≥ 2)

Independent right censored samples (K samples/groups).

$$H_0 : h_1(t) = h_2(t) = \ldots h_k(t) \text{ for all } t \leq \tau$$

$$H_a : \text{At least one of the } h_j(t)'s \text{ different for some } t \leq \tau$$

$\tau$ = largest time at which all of the groups have at least one subject at risk

# K sample tests (K ≥ 2)

Event times $t_1 < t_2 < \ldots < t_D$ in the pooled sample.

$$d_i = \sum_{j=1}^{K} d_{ij}$$

$i$: Time point; $j$: Group

= # patients that experience the event in the pooled sample

$$Y_i = \sum_{j=1}^{K} Y_{ij}$$

= # patients at risk in the pooled sample

# Nelson-Aalen crude estimate of $h(t_i)$

Recap: An estimator of the cumulative hazard rate $H(t)$, is the **Nelson-Aalen estimator**.

$$\tilde{H}(t) = \begin{cases} 0 & \text{if } t < t_1 \\ \sum_{t_i \le t} \left( \frac{d_i}{Y_i} \right) & \text{if } t_1 \le t \end{cases}$$

A crude estimate of the hazard rate at an event time $t_i$ is given by $\frac{d_i}{Y_i}$

# K sample tests (K ≥ 2)

$$P(\text{experience the event at time } t_i) = \frac{d_i}{Y_i} \mid H_0 \text{ true}$$

Reasonable to compare $\frac{d_{ij}}{Y_{ij}}$ to $\frac{d_i}{Y_i}$

Proportion that experience the event in group $j$

Proportion that experience the event in all groups

# K sample tests (K ≥ 2)

Test function:

$$Z_j(\tau) = \sum_{i=1}^{D} W_j(t_i)\left\{\frac{d_{ij}}{Y_{ij}} - \frac{d_i}{Y_i}\right\}$$

Positive weight function
$W_j(t)$ = zero when $Y_{ij}$ is zero.

# K sample tests (K ≥ 2)

All of the commonly used tests have

$W_j(t_i) = Y_{ij} W(t_i)$ where $W(t_i)$ is common for all groups

$$Z_j(\tau) = \sum_{i=1}^{D} W(t_i) \left\{ d_{ij} - Y_{ij} \frac{d_i}{Y_i} \right\}$$

Observed # events in group $j$

Expected # events in group $j$

$$\hat{V}\left(Z_j(\tau)\right) = \sum_{i=1}^{D} W(t_i)^2 \frac{Y_{ij}}{Y_i} \left(1 - \frac{Y_{ij}}{Y_i}\right) \left(\frac{Y_i - d_i}{Y_i - 1}\right) d_i$$

# K sample tests (K ≥ 2)

$$\sum_{i=1}^{K} Z_j(\tau) = 0 \quad \Longrightarrow \quad Z_j\text{'s are linearly dependent}$$

The test statistic is constructed by selecting any $K$-1 of the $Z_j$'s.

For two groups: $$\frac{Z_j(\tau)^2}{\hat{V}\left(Z_j(\tau)\right)} \sim \chi^2$$

$j$ = 1 or 2, doesn't matter which group

# Log-rank and Gehan's test

Two most common weight functions:

1) $W(t) = 1 \text{ for all } t$ gives the **Log-rank test** *(alias Mantel-Cox, or Generalized Savage test)*

2) $W(t_i) = Y_i$ gives **Gehan's test** *(alias Breslow-Gehan, or Wilcoxon test)*

# Assumptions

Assumptions for the Log-rank and Gehan's tests*:

- random, independent samples
- noninformative censoring
- right censored data
- survival probabilities are the same for subjects recruited early and late in the study
- large samples (tests are based on large-sample approximations to the chi-square distribution)

**Log rank test described by Bland & Altman, BMJ. May 1, 2004; 328(7447): 1073.*

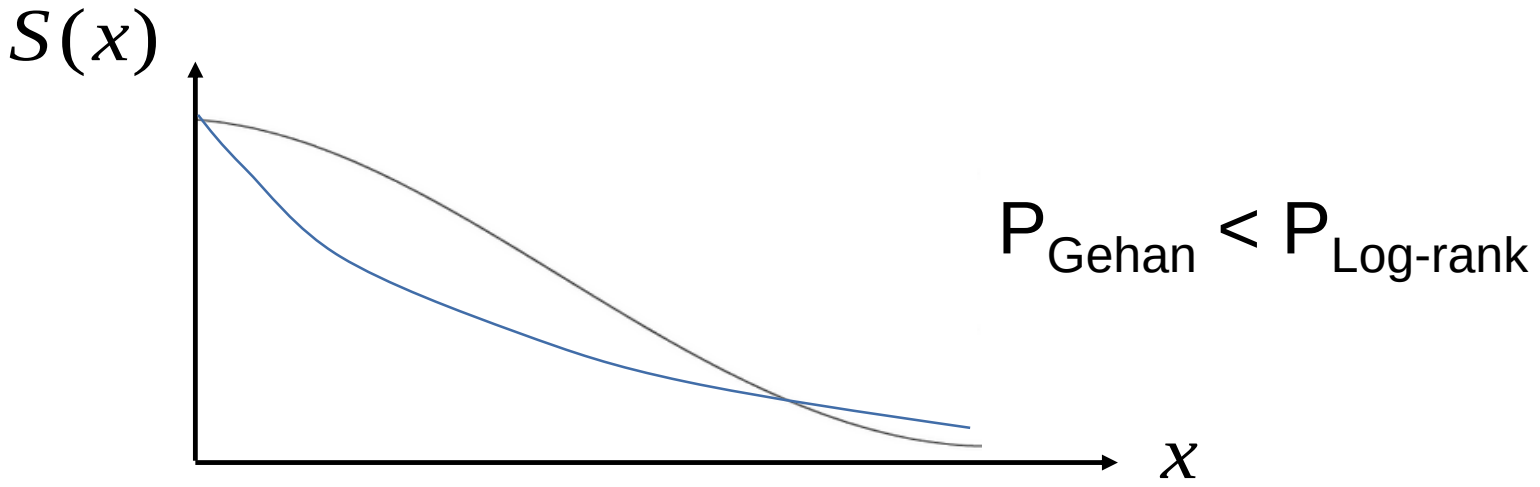
# Log-rank test vs. Gehan’s test

The **Log-rank test** has optimum power to detect differences when the hazard rates in all groups are proportional to each other.

**Gehan’s test** puts larger weight to early differences.

Both test can have misleading results when censoring patterns are different in different groups (Gehan's test more sensitive).

# Log-rank test vs. Gehan's test




Your hypothesis decides which test to use.

# Example: Lifetime of refrigerators

Refrigerators with two different types of components.

$$H_0 : h_1(t) = h_2(t) \quad \text{for all } t \leq \tau$$
$$H_a : h_1(t) \neq h_2(t) \quad \text{for some } t \leq \tau$$

$\tau$ = largest time at which both groups of refrigerators have at least one subject at risk

# Example: Lifetime of refrigerators

Significance level?

Always reflect upon the consequences of wrongly rejecting $H_0$.

Wrongly rejecting $H_0$ in this case means that we would claim that there is a difference in survival between the two types of components, when in fact there is no difference.

# Example: Lifetime of refrigerators

Does this have any serious consequences?

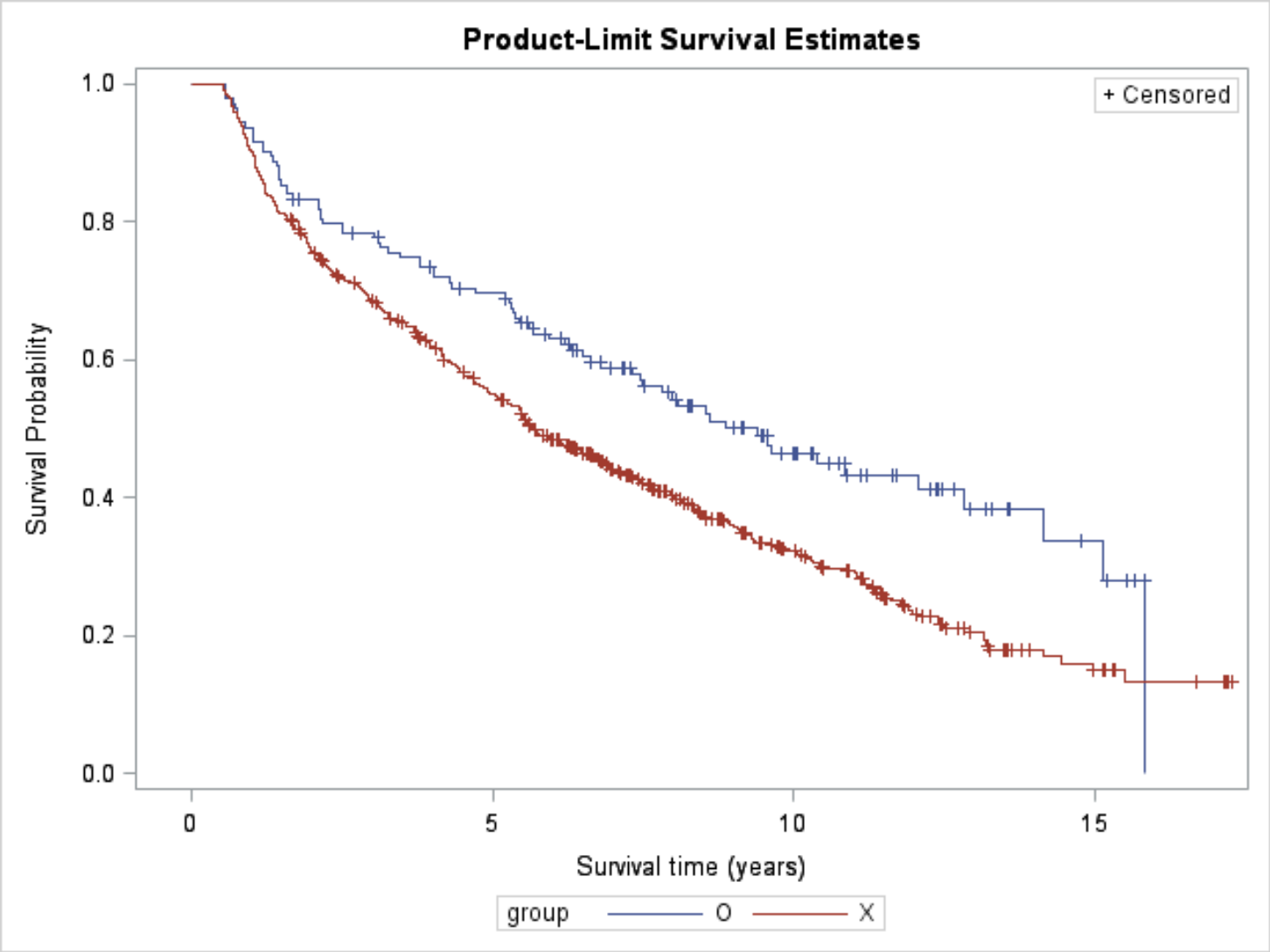
Depends on the circumstances.

One possible consequence of claiming that one component is better than the other, if it isn't, is that the trust/reputation of the company/product may go down.

Economical consequences could also be considered. If one of the components is extremely expensive, we might want to choose a lower significance level to get a "safer" result to base our investment decision on.

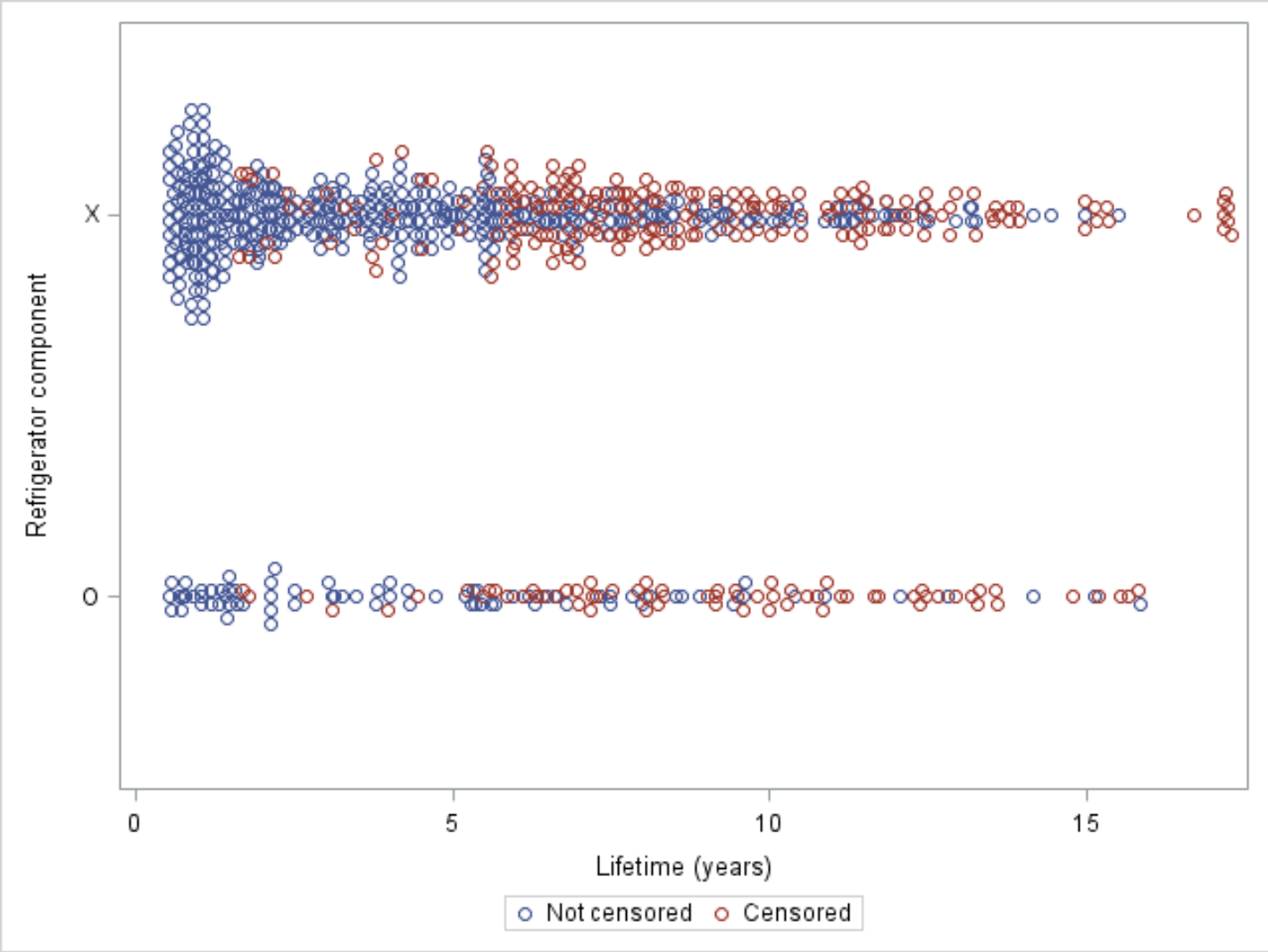
UPPSALA UNIVERSITET

# Example: Lifetime of refrigerators




*In Sas: proc lifetest, strata statement.*

# Example: Lifetime of refrigerators




Similar pattern in both groups, i.e., doesn't contradict noninformative censoring

# Example: Lifetime of refrigerators

| Summary of the Number of Censored and Uncensored Values | | | | | |
|---|---|---|---|---|---|
| Stratum | group | Total | Failed | Censored | Percent Censored |
| 1 | O | 144 | 76 | 68 | 47.22 |
| 2 | X | 610 | 410 | 200 | 32.79 |
| Total | | 754 | 486 | 268 | 35.54 |

No. events

*In Sas: proc lifetest, strata statement.*

# Recap: Assumptions

Assumptions for the Log-rank and Gehan's tests*:

- random, independent samples ✓
- noninformative censoring - *To be discussed with field experts*
- right censored data ✓
- survival probabilities are the same for subjects recruited early and late in the study ✓
- large samples (tests are based on large-sample approximations to the chi-square distribution) ✓

*- To be discussed with field experts (not an issue in this study since all refrigerators are manufactured at the same time)*

**Log rank test described by Bland & Altman, BMJ. May 1, 2004; 328(7447): 1073.*

# Example: Lifetime of refrigerators

**The LIFETEST Procedure**

**Testing Homogeneity of Survival Curves for time over Strata**

| Rank Statistics | | |
|---|---|---|
| group | Log-Rank | Wilcoxon |
| O | -31.719 | -13736 |
| X | 31.719 | 13736 |

$= Z_j(\tau)$

# Example: Lifetime of refrigerators

| Covariance Matrix for the Wilcoxon Statistics | | |
|---|---|---|
| group | O | X |
| O | 20133277 | -2.013E7 |
| X | -2.013E7 | 20133277 |

$$= \hat{V}\left(Z_j(\tau)\right)$$

# Example: Lifetime of refrigerators

| Test of Equality over Strata | | | |
|---|---|---|---|
| Test | Chi-Square | DF | Pr > Chi-Square |
| Log-Rank | 12.0543 | 1 | 0.0005 |
| Wilcoxon | 9.3714 | 1 | 0.0022 |
| -2Log(LR) | 13.0272 | 1 | 0.0003 |

Log-Rank test ---->

Gehan's test ----->

Likelihood ratio test. Based on the exponential distribution which is rarely applicable in a survival model. Can be ignored.

# Other weight functions

$$W(t_i) = \sqrt{Y_i}$$

**Tarone and Ware.**
Gives more weight to differences at time points where there is most data.

$$W(t_i) = \tilde{S}(t_i) \quad \text{where} \quad \tilde{S}(t) = \prod_{t_i \le t} \left(1 - \frac{d_i}{Y_i + 1}\right)$$

**Peto and Peto, Kalbfleish and Prentice.**

$$W(t_i) = \tilde{S}(t_i)\frac{Y_i}{Y_i + 1}$$

**Andersen et al.**

The weights depend on the combined survival experience in the pooled sample.

# Other weight functions

$$W_{p,q}(t_i) = \hat{S}_{KM}(t_{i-1})^p \left(1 - \hat{S}_{KM}(t_{i-1})\right)^q, \quad p \geq 0,\ q \geq 0$$

**Fleming and Harrington.**
A very general class of tests.
E.g. $p$=$q$=0 gives the Log-rank test.

When $q$=0 and $p$>0 most weight is given to early departures,
and when $p$=0 and $q$>0 most weight is given to late departures

# Gray's test for cumulative incidence

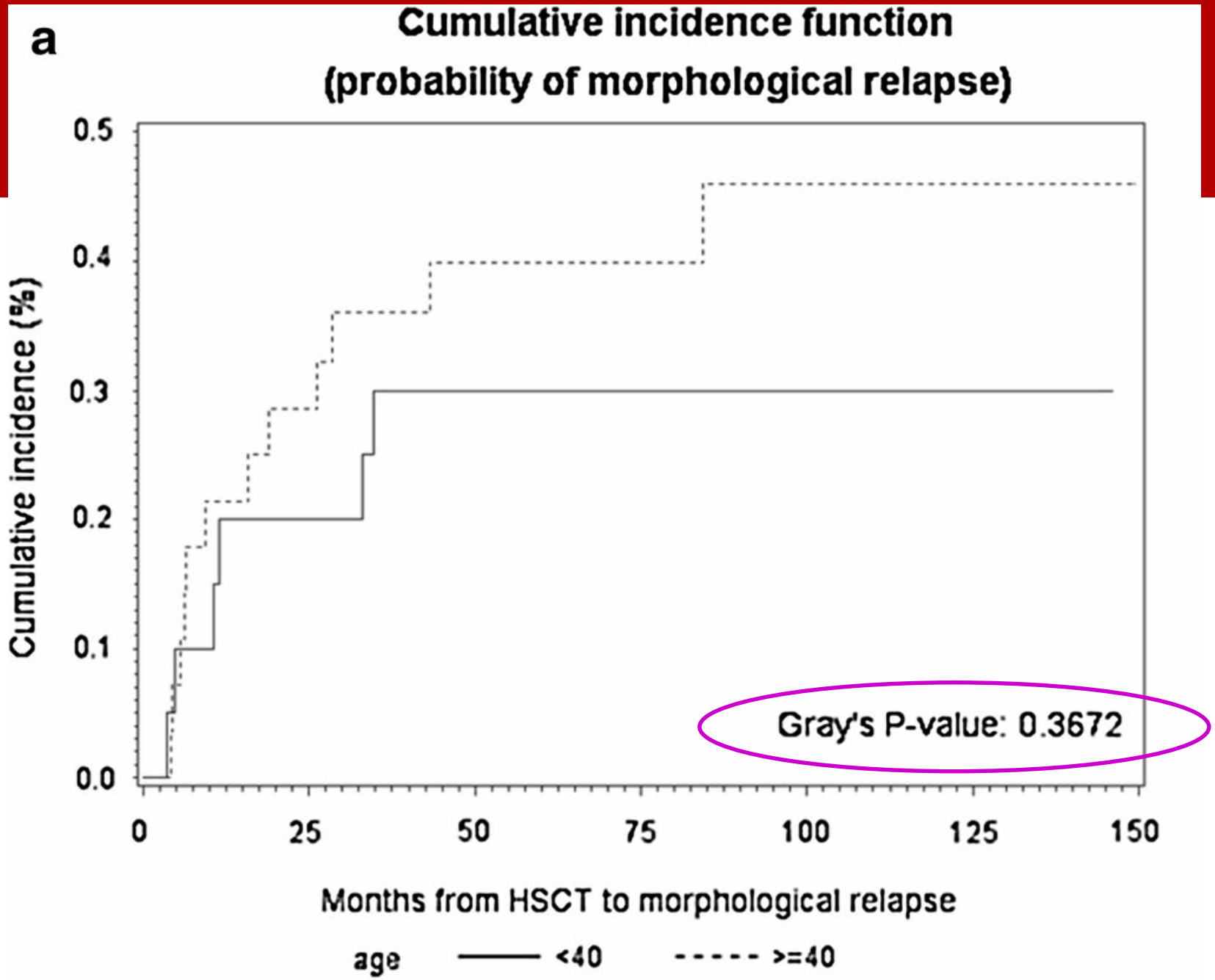
Independent right censored samples (K samples/groups).

$$H_0 : CI_1(t) = CI_2(t) = \ldots CI_k(t) \text{ for all } t \leq \tau$$

$$H_a : \text{At least one of the } CI_j(t)'s \text{ different for some } t \leq \tau$$

$\tau$ = largest time at which all of the groups have at least one subject at risk

*Gray, R. (1988), A Class of K-Sample Tests for Comparing the Cumulative Incidence of a Competing Risk. The Annals of Statistics, 16, 1141–1154.*


a
Cumulative incidence function
(probability of morphological relapse)
Cumulative incidence (%)
0.0
0.1
0.2
0.3
0.4
0.5
0
25
50
75
100
125
150
Months from HSCT to morphological relapse
Gray's P-value: 0.3672
age
<40
>=40

# Assumptions

Assumptions for Gray's test:

- random, independent samples[1]
- noninformative censoring[1]
- right censored data[1]
- time to event and event type independent and identically distributed within each group[2]
- large samples (asymptotic results)[2]

[1] *Same as for the Kaplan-Meier estimator*

[2] *Gray, R. (1988), A Class of K-Sample Tests for Comparing the Cumulative Incidence of a Competing Risk. The Annals of Statistics, 16, 1141–1154.*

# Survival vs. cumulative incidence

"..cause-specific hazards and cumulative incidence curves capture different aspects of the event histories in competing risks data, inference on these metrics may yield different results."

".. the log-rank test correctly detects differences in cause-specific hazards, and, unless there is strong dependence between failure times, is largely unaffected by between-group differences in hazards for other competing events.."

*Dignam JJ, Kocherginsky MN (2008) Choice and interpretation of statistical tests used when competing risks are present. J Clin Oncol 26: 4027-4034.*

# Survival vs. cumulative incidence

"On the other hand, Gray's test correctly detects whether there is a difference in cumulative incidence between groups for a given event, whether that difference is caused by a difference in hazards between the groups for the event itself or by a difference in hazards for the competing events."

"Thus, when choosing and interpreting a statistical test, one must take these properties into account, as well as whether primary interest is in contrasting the relative rate of events or the absolute difference in incidence of events between groups."

# Program L4

- **Summary curves for competing risks**
  - Cumulative incidence function
- **Hypothesis testing**
  - K samples tests
    - Log-rank test
    - Gehan's test
    - Gray's test
  - Stratified tests
  - Differences at a fixed point in time

# Stratified tests

There might be some other covariates that affect the event rates in the $K$ different populations.

**Example**: A breast cancer treatment is being tested.

Two groups are compared, A= new treatment, B= old treatment.

There are more smokers in group A than in group B. It would be wise to adjust for the variable smoking.

This can be done by stratifying a $K$ sample test (e.g. the Log-rank test).

# Stratified tests

Independent right censored samples (K samples/groups, M strata).

$$H_0 : h_{1s}(t) = h_{2s}(t) = \ldots h_{Ks}(t), \quad \text{for s} = 1,\ldots,M, \quad t < \tau$$

$$H_a : \text{At least one of the } h_{js}(t)\text{'s different for some } t < \tau$$

$\tau$ = largest time at which all of the groups have at least one subject at risk

# Stratified tests

$$Z_{js}(\tau) = \sum_{i=1}^{D} W(t_i)\left\{ d_{ij} - Y_{ij}\frac{d_i}{Y_i} \right\} \text{ for stratum } s$$

$$Z_j(\tau) = \sum_{s=1}^{M} Z_{js}(\tau) \text{ for all strata}$$

$$\hat{V}\left(Z_j(\tau)\right) = \sum_{s=1}^{M} \hat{V}\left(Z_{js}(\tau)\right)$$

# Power of stratified tests

Stratified tests have good power against alternative hypotheses that are in the same direction in each stratum.

When this is not the case, stratified tests may have very low power.

In such situations it is better to use separate tests for each stratum.

# Example: Hodgkin's disease

A small study comparing the effectiveness of allogeneic transplants versus autogeneic transplants for Hodgkin's disease or non-Hodgkin's lymphoma is presented in section 1.10.

Allogeneic transplant: from a matching donor

Autogeneic transplant: you own bone marrow is cleansed and returned after a high dose of chemotherapy

Is there a difference in disease-free survival between allogeneic and autogeneic transplants?

# Example: Hodgkin's disease

**Variables:**

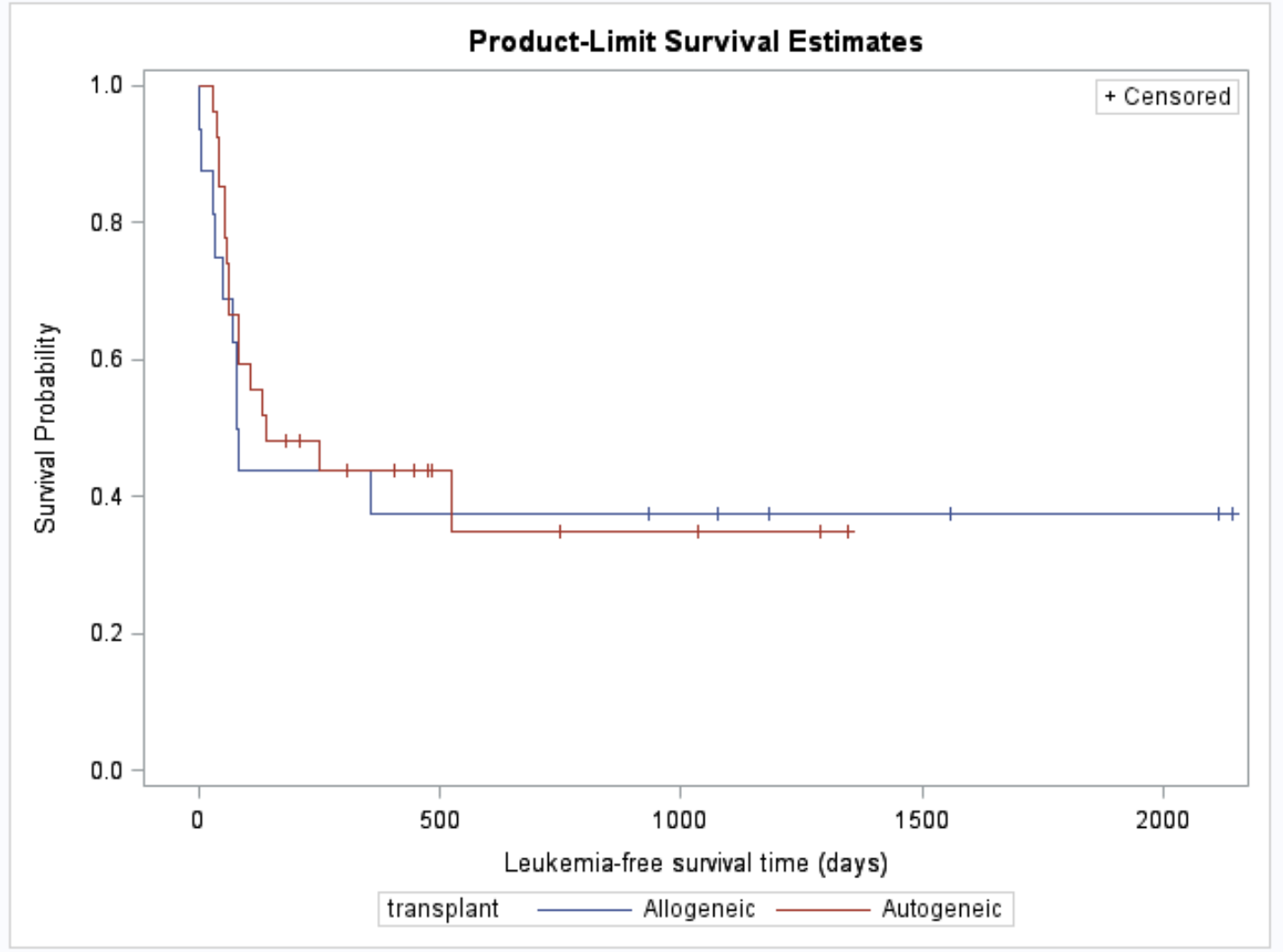
freetime = time to death or relapse (days)
transplant = type of transplant (0=allogeneic, 1=autogeneic)
event = event indicator (1=dead or relapse,
0=alive without relapse)
disease = disease type (0=non-Hodgkin's lymphoma,
1=Hodgkin's disease)

# Example: Hodgkin's disease

To plot the Kaplan-Meier (Product-Limit) survival, use the **strata** command to compare the transplant groups.

```
proc lifetest data=hodgkins;
   time freetime*event(0);
   strata transplant;
run;
```

# Example: Hodgkin's disease



# Example: Hodgkin's disease

To stratify the test, use the **strata** command for the stratifying variable and the /**group=** option for the variable to be tested.

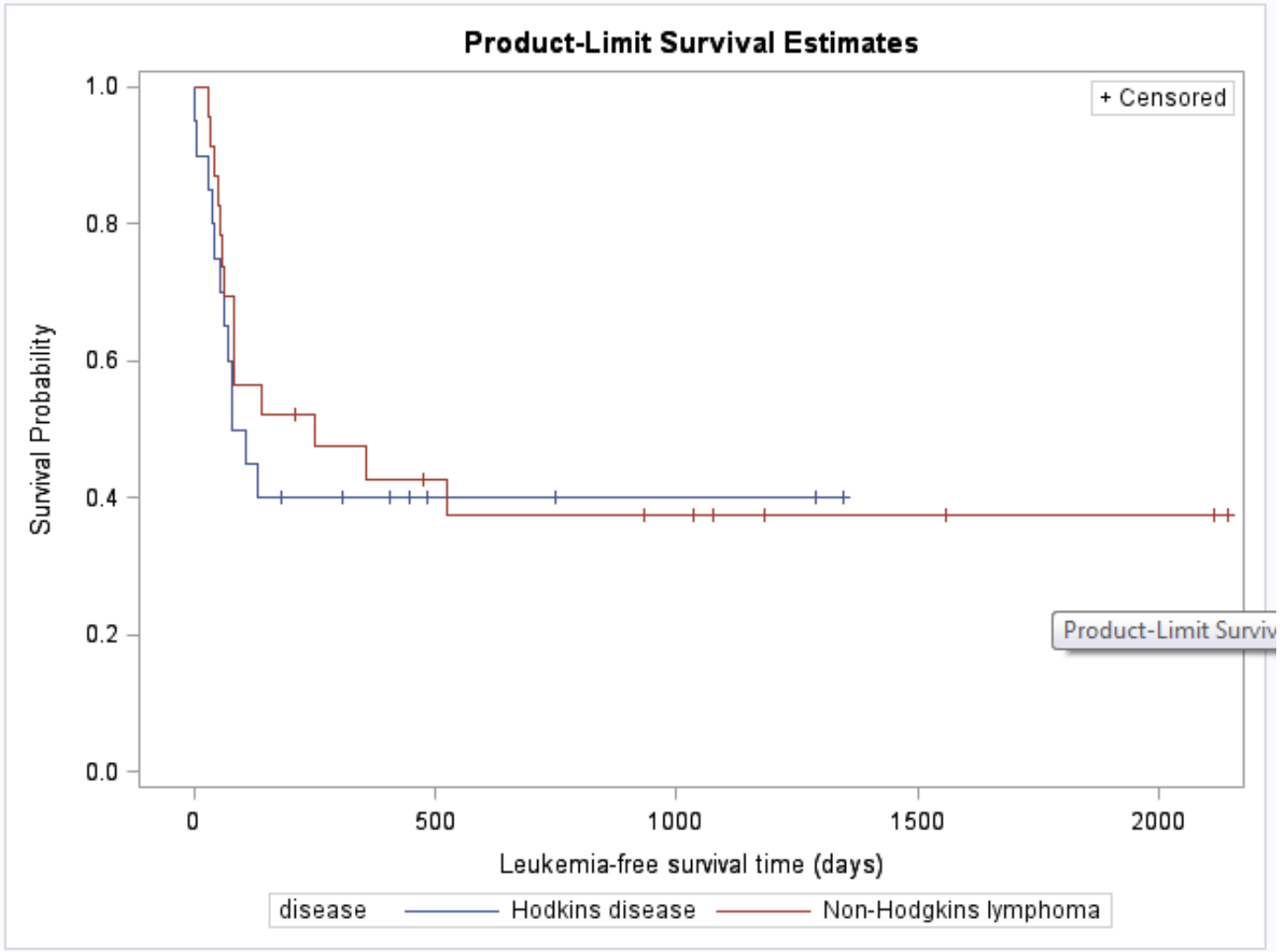
```
proc lifetest data=hodgkins;
   time freetime*event(0);
   strata disease / group=transplant;
run;
```

Stratifying variable

Variable containing groups to be compared

# Example: Hodgkin's disease

Note: Survival curves by the stratification variable only (disease)



# Example: Hodgkin's disease

Group = type of transplant

Stratified by disease

| Stratified Test of Equality over Group | | | |
|---|---|---|---|
| Test | Chi-Square | DF | Pr > Chi-Square |
| Log-Rank | 0.1202 | 1 | 0.7288 |
| Wilcoxon | 0.2942 | 1 | 0.5875 |

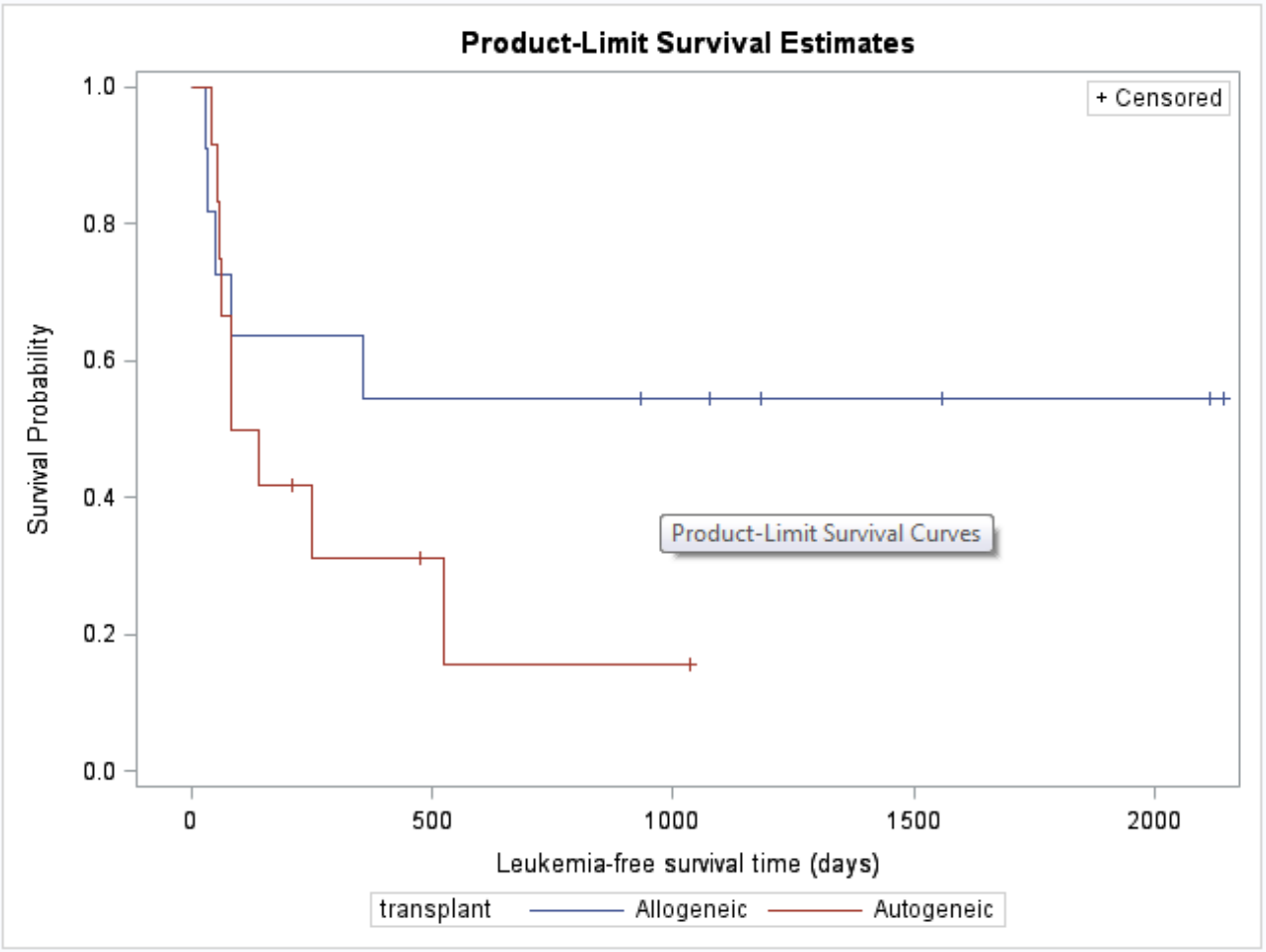
# Example: Hodgkin's disease

Alternative to use if the alternative hypothesis is not in the same direction in the disease groups; testing the difference between types of transplant for each disease group separately:

```
proc lifetest data=hodgkins;
   time freetime*event(0);
   strata transplant;
   by disease;
run;
```

# Example: Hodgkin's disease

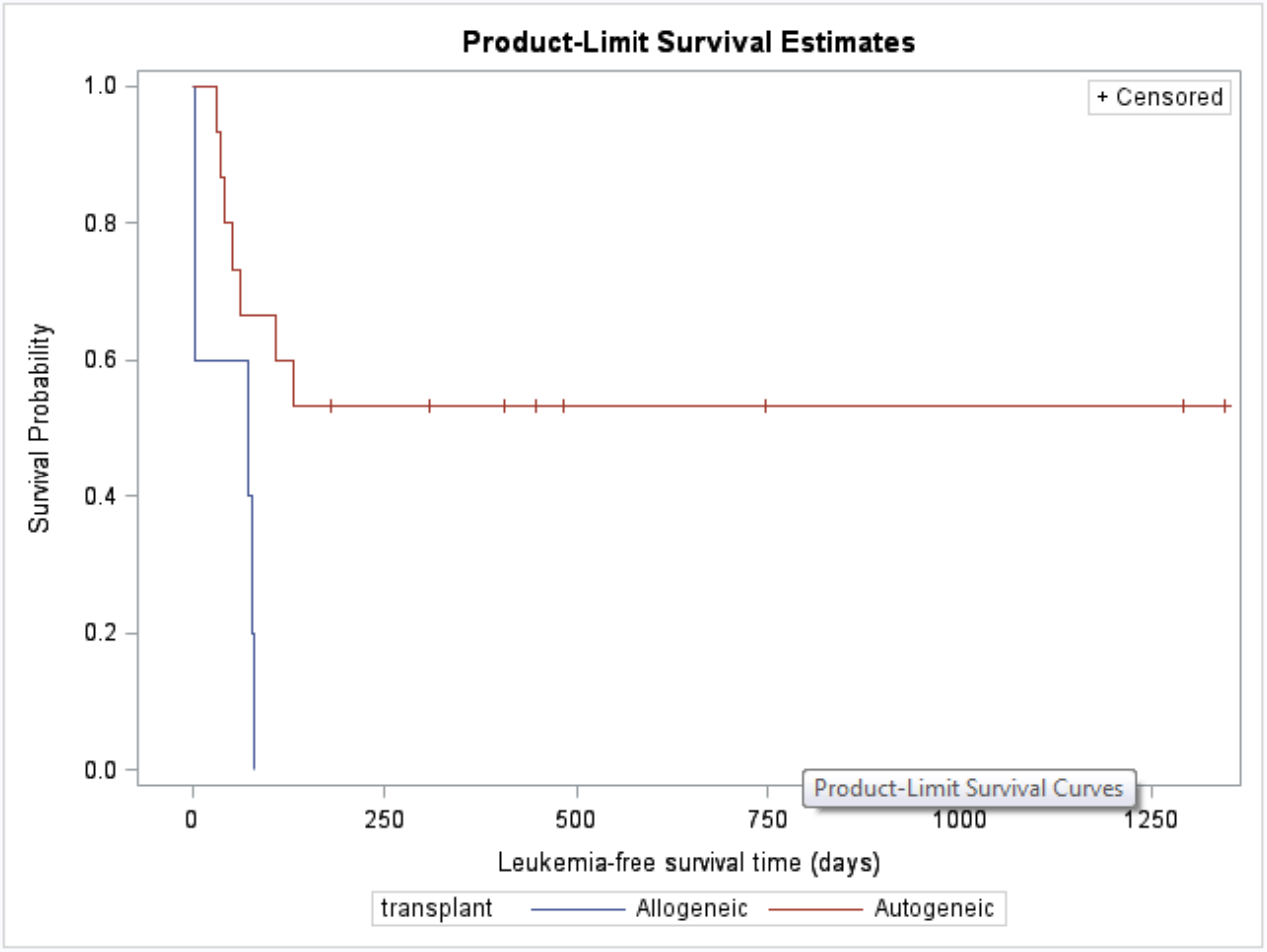
**disease=Non-Hodgkins lymphoma**




| Test of Equality over Strata | | | |
|---|---|---|---|
| Test | Chi-Square | DF | Pr > Chi-Square |
| Log-Rank | 1.6552 | 1 | 0.1983 |
| Wilcoxon | 0.6447 | 1 | 0.4220 |
| -2Log(LR) | 10.1967 | 1 | 0.0014 |

# Example: Hodgkin's disease

**disease=Hodkins disease**




| Test of Equality over Strata | | | |
|---|---|---|---|
| Test | Chi-Square | DF | Pr > Chi-Square |
| Log-Rank | 6.3574 | 1 | 0.0117 |
| Wilcoxon | 5.1923 | 1 | 0.0227 |
| -2Log(LR) | 16.5414 | 1 | <.0001 |

# Program L4

- **Summary curves for competing risks**
  - Cumulative incidence function
- **Hypothesis testing**
  - K samples tests
    - Log-rank test
    - Gehan's test
    - Gray's test
  - Stratified tests
  - Differences at a fixed point in time

# Differences in outcome at a fixed point in time

Sometimes you might want to compare $K$ survival curves or cumulative incidence curves at a predetermined fixed point in time, $t_0$.

Important to select $t_0$ before examining the data.

Test statistics are special cases of tests for contrasts between a set of parameters.

# Tests for differences in outcome at a fixed point in time

$$H_0 : S_1(t_0) = S_2(t_0) = \ldots = S_K(t_0)$$

$H_a$ : At least one of the $S_j(t_0)$'s is different, for predetermined $t_0$

or

$$H_0 : CI_1(t_0) = CI_2(t_0) = \ldots = CI_K(t_0)$$

$H_a$ : At least one of the $CI_j(t_0)$'s is different, for predetermined $t_0$

# Tests for differences in outcome at a fixed point in time

Kaplan-Meier (Product-Limit) estimates of the survival function, or cumulative incidence estimates are used, together with the estimated variances of these statistics.

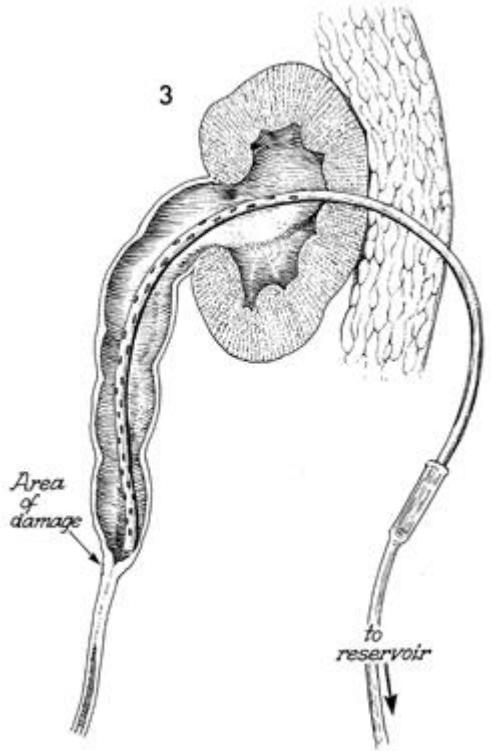
Test statistic when comparing the survival in two groups:

$$Z = \frac{\hat{S}_1(t_0) - \hat{S}_2(t_0)}{\sqrt{\hat{V}\left(\hat{S}_1(t_0)\right) + \hat{V}\left(\hat{S}_2(t_0)\right)}}$$

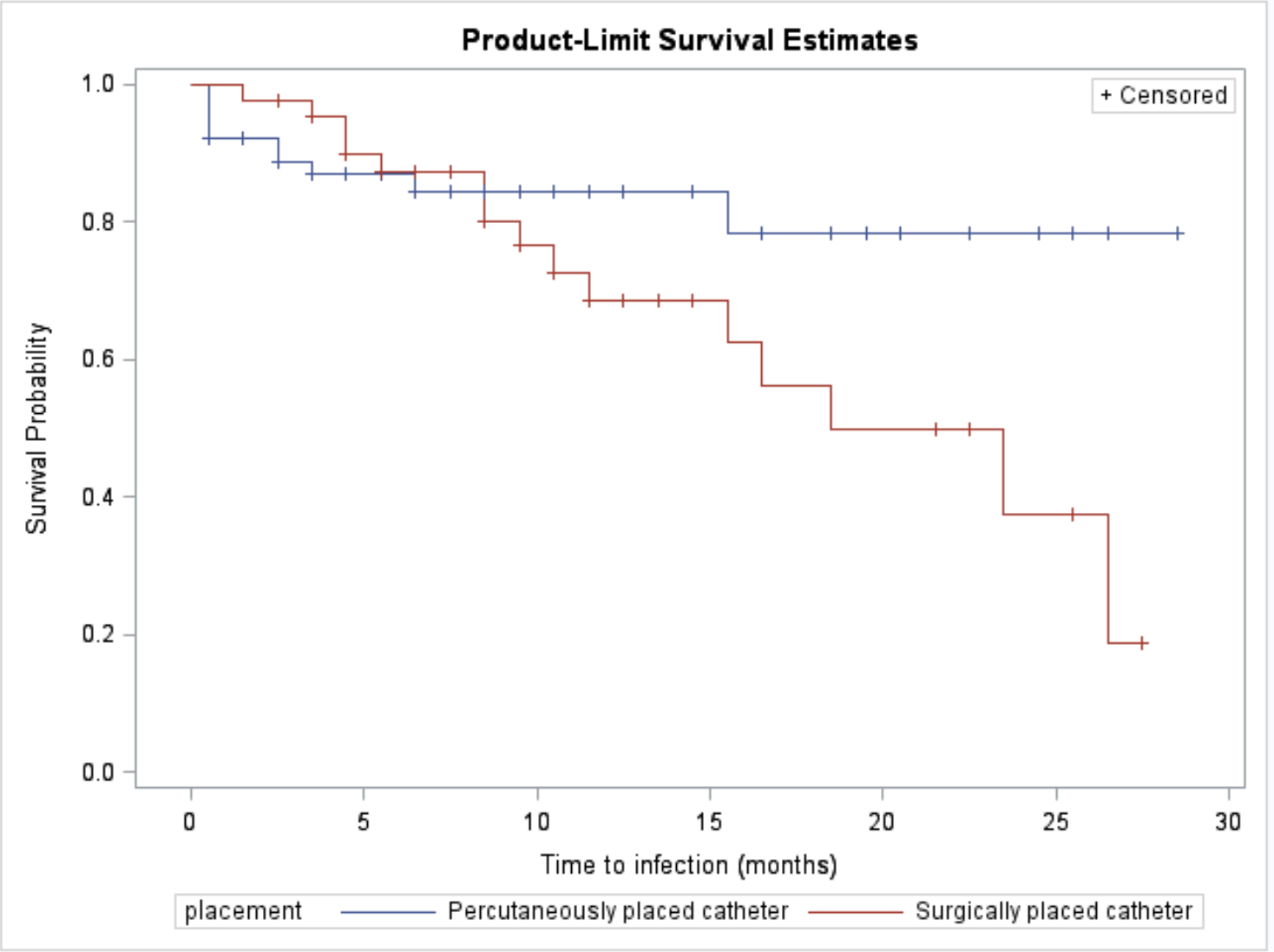
# Example: Renal catheters

In section 1.4 the times to first exit-site infection (in months) of patients with renal insufficiency was reported.

In the study 43 patients had a surgically placed catheter (Group 1) and 76 patients had a percutaneous placement of their catheter (Group 0).



# Example: Renal catheters



# Example: Renal catheters

To compare the survival functions at e.g. 3 months:

$$H_0 : S_1(3) = S_0(3)$$
$$H_a : S_1(3) \neq S_0(3)$$

Test statistic:

$$Z = \frac{\hat{S}_1(3) - \hat{S}_0(3)}{\sqrt{\hat{V}\left(\hat{S}_1(3)\right) + \hat{V}\left(\hat{S}_0(3)\right)}}$$

# Example: Renal catheters

Stratum 1: placement = Percutaneously placed catheter

Product-Limit Survival Estimates

| time | | Survival | Failure | Survival Standard Error | Number Failed | Number Left |
|---|---|---|---|---|---|---|
| 0.0000 | | 1.0000 | 0 | 0 | 0 | 76 |
| 0.5000 | | . | . | . | 1 | 75 |
| 0.6000 | | | | | 2 | 74 |
| ... | | | | | | |
| 2.5000 | | . | . | . | 7 | 55 |
| 2.5000 | | 0.8882 | 0.1118 | 0.0376 | 8 | 54 |
| 2.5000 | * | . | . | . | 8 | 53 |
| 2.5000 | * | . | . | . | 8 | 52 |
| 2.5000 | * | . | . | . | 8 | 51 |
| 2.5000 | * | . | . | . | 8 | 50 |
| 2.5000 | * | . | . | . | 8 | 49 |
| 3.5000 | | 0.8700 | 0.1300 | 0.0409 | 9 | 48 |
| 3.5000 | * | . | . | . | 9 | 47 |

$$\hat{S}_0(3) = 0.8882$$

$$\hat{V}\left(\hat{S}_0(3)\right) = 0.0376^2 = 0.00141$$

# Example: Renal catheters

Stratum 2: placement = Surgically placed catheter

| Product-Limit Survival Estimates | | | | | | |
|---|---|---|---|---|---|---|
| time | | Survival | Failure | Survival Standard Error | Number Failed | Number Left |
| 0.0000 | | 1.0000 | 0 | 0 | 0 | 43 |
| 1.5000 | | 0.9767 | 0.0233 | 0.0230 | 1 | 42 |
| 2.5000 | * | . | . | . | 1 | 41 |
| 2.5000 | * | . | . | . | 1 | 40 |
| 3.5000 | | 0.9523 | 0.0477 | 0.0329 | 2 | 39 |
| 3.5000 | * | | | | 2 | 38 |

$$\hat{S}_1(3) = 0.9767$$

$$\hat{V}\left(\hat{S}_1(3)\right) = 0.0230^2$$

$$= 0.00053$$

# Example: Renal catheters

To compare the survival functions at e.g. 3 months:

$$H_0 : S_1(3) = S_0(3)$$
$$H_a : S_1(3) \neq S_0(3)$$

Test statistic:

$$Z = \frac{\hat{S}_1(3) - \hat{S}_0(3)}{\sqrt{\hat{V}\left(\hat{S}_1(3)\right) + \hat{V}\left(\hat{S}_0(3)\right)}} = 2.01$$