Introduction to Competing Risks

Henrik Ravn

Novo Nordisk

DSBS Course
Survival Analysis in Clinical Trials
January 2018
Survival data

Survival data = \textit{single} outcome.
- time from origin (time zero) to event (death)
- right censoring, delayed entry
- basic quantity is the intensity
  \[ = \text{rate} \]
  \[ = \text{hazard function} \]
  \[ = \lambda(t) \approx \text{Prob(die before } t + dt \mid \text{alive at } t)/dt \]
Survival data

The hazard, $\lambda(t)$, provides a **local** (in time) description of the development.

Models for $\lambda(t)$
- Cox regression
- Poisson regression

The survival function

$$S(t) = \text{Prob(}\text{alive at time } t\text{)}$$

describes the **cumulative** development over time.
Two-state model for survival data

\[ 
\begin{align*}
\lambda(t) & \approx \frac{\text{Prob(state 1 time } t + dt \mid \text{state 0 time } t)}{dt} \\
S(t) & = \text{Prob(state 0 time } t) \\
F(t) & = 1 - S(t) = \text{Prob(state 1 time } t) \text{ is the cumulative probability (”risk”) of death over the interval from 0 to } t
\end{align*}
\]
Relation between rates and risks

We have the relation

\[ F(t) = 1 - \exp(-\int_0^t \lambda(u)du). \]

This means that whenever we have a (Cox/Poisson/...) model for the rate, \( \lambda(t) \), we also have a model for the risk: \( F(t) = 1 - S(t) \).

A **one-to-one** correspondence between hazard and survival (and risk).

In particular, if the rate increases/decreases with a covariate, \( X \), then also the risk increases/decreases with \( X \).
Competing risks

\[ \lambda_1(t) \]

\[ \lambda_k(t) \]

0
Alive

1
Dead, cause 1

\cdot

\cdot

k
Dead, cause k
Competing risks

Cause-specific intensities

\[ \lambda_1(t) \approx \frac{\text{Prob}(\text{state 1 time } t + dt \mid \text{state 0 time } t)}{dt} \]
\[ \lambda_2(t) \approx \frac{\text{Prob}(\text{state 2 time } t + dt \mid \text{state 0 time } t)}{dt} \]
\[ \vdots \]
\[ \lambda_k(t) \approx \frac{\text{Prob}(\text{state } k \text{ time } t + dt \mid \text{state 0 time } t)}{dt} \]

State occupation probabilities include the overall survival function:

\[ S(t) = P(\text{alive time } t). \]

and the \textit{cumulative incidences} ("sub-distribution function") \( j = 1, \ldots, k \):

\[ F_j(t) = P(\text{dead from cause } j \text{ before time } t) \]
Recap: Population and sample

We are used to considering our data as a sample from some (target) population, and the parameters refer to this population.

That is no different in survival analysis, however, it is important to realize that the target population is a complete population, i.e., without censoring.

Our ambition in survival analysis is therefore to draw inference on parameters like the survival function $S(t)$ or the hazard function $\lambda(t)$ from a potentially completely observed population based on incomplete (censored) data.

This is quite ambitious and requires certain assumptions.
For this ambition to be feasible:

1. the complete population should be well-defined
2. censoring should not leave us with a biased sample

Requirement 1 basically tells that the event under study should be able to happen for every one in the population.

A complete population without censoring is not well-defined if, e.g. the event of interest is AE and death is censoring; the event under study should be able to happen for every one in the uncensored population.

We must acknowledge that individuals may die without an AE and inference for AE risks and rates should be made "in the presence of the competing risk of dying".
Recap: Kaplan-Meier and Nelson-Aalen

Censored (independently) sample of times $0 < t_1 < t_2 < \ldots$ with $d(t_1), d(t_2), \ldots$ being the observed numbers of failures.

The Kaplan-Meier (KM) estimator for the survival function

$$\hat{S}(t) = \prod_{t_j \leq t} \left(1 - \frac{d(t_j)}{R(t_j)}\right).$$

The estimator for the risk $F(t)$, is "1–KM".

The Nelson-Aalen estimator for the cumulative hazard

$$\hat{\Lambda}(t) = \sum_{t_j \leq t} \frac{d(t_j)}{R(t_j)}.$$
LEADER – Death from any cause

1–KM was used to estimate risk functions $F(t) = 1 - S(t)$ for the treatment groups.
LEADER – CV death and other causes

\[ \lambda_1(t) \] 
\[ \lambda_2(t) \]
LEADER – Death from cardiovascular causes

1–KM used to estimate risk functions by censoring for other causes of death
Censor the competing cause

Wish to estimate parameters about CV deaths (for the risk or rate) by censoring for other causes of death. Recall that, for this to be feasible:

1. The complete population (i.e., without censoring) should be well-defined
2. Censoring should not leave us with a biased sample

Requirement 2 is still an issue to be debated, however, requirement 1 will be violated when we censor for other causes:

This is because we attempt to make inference for a potentially completely observed population where other causes of death do not exists. Such a population is hypothetical.
In the competing risks model (2 causes of death)

Cause-specific hazards $j = 1, 2$ (“transition intensities”):

$$\lambda_j(t) \approx P(\text{state } j \text{ time } t + dt \mid \text{state 0 time } t)/dt.$$ 

State occupation probabilities include the overall survival function:

$$S(t) = P(\text{alive time } t) = \exp[-(\Lambda_1(t) + \Lambda_2(t))].$$

and the cumulative incidences ("sub-distribution function") $j = 1, 2$:

$$F_j(t) = P(\text{dead from cause } j \text{ before time } t) = \int_0^t S(u)\lambda_j(u)du$$

NB: $S(t) + F_1(t) + F_2(t) = 1$ for all $t$. 
Rates and risks

Note that the risk for cause 1 depends on the rates for both causes 1 and 2.

For the two-state model for (overall) survival, hazard function $\lambda$ and failure function $F$ contain *equivalent* information and one may be obtained from the other.

This one-to-one correspondence is lost for competing risks

Both of the cause-specific hazards, $\lambda_1, \lambda_2$, are needed when computing each of the cumulative incidences, $F_1, F_2$. 
Using the Kaplan-Meier estimator on a single cause

We have the relation:

\[ F_1(t) = P(\text{dead from cause 1 before time } t) = \int_0^t S(u)\lambda_1(u)du. \]

When \( \lambda_2(t) = 0 \), i.e. when the competing event is not present, then

\[ F_1^0(t) = 1 - \exp(-\int_0^t \lambda_1(u)du) = 1 - S_1(t), \text{ say.} \]

That is, “1–KM for cause 1”, \( 1 - \hat{S}_1(t) \), estimates

\[ P(\text{dead from cause 1 before time } t) \quad \text{IF} \quad \lambda_2(t) = 0 \]

That is, if the competing risk does not exist.
The cumulative cause-specific hazard can be estimated using the Nelson-Aalen estimator using only failures from the relevant cause, e.g., cause 1:

\[ \Lambda_1(t) = \int_0^t \lambda_1(u) du \]

may be estimated by

\[ \hat{\Lambda}_1(t) = \sum_{t_j \leq t} \frac{d(t_j) I(\text{cause} = 1)}{R(t_j)}. \]

This is an increasing step function with steps at each observed time of failure from cause 1.
Aalen-Johansen estimator

Non-parametric estimation of risks/probabilities.

The overall survival function $S(t)$ can be estimated by the Kaplan-Meier estimator, $\hat{S}(t)$, *using all failures* (LEADER).

The *cumulative incidences*: $F_1(t)$ and $F_2(t)$ may be estimated by *plugging-in* estimates for $S(t)$, $\Lambda_1(t)$ and $\Lambda_2(t)$. E.g., for cause 1:

$$\hat{F}_1(t) = \sum_{t_j \leq t} \hat{S}(t_{j-1}) \frac{d(t_j)I(\text{cause = 1})}{R(t_j)}.$$

This is often called the *Aalen-Johansen estimator*.

LEADER?
Aalen-Johansen estimator

\[
\hat{F}_1(t) = \sum_{t_j \leq t} \hat{S}(t_{j-1}) \frac{d(t_j) I(\text{cause} = 1)}{R(t_j)}
\]

\[
= \sum_{t_j \leq t} \hat{S}(t_{j-1})(\hat{\Lambda}_1(t_{j-1}) - \hat{\Lambda}_1(t_j))
\]

\[
= \sum \text{KM} \cdot \text{(increments in cause-specific N-Aa)}
\]
1-Kaplan-Meier vs. Aalen-Johansen

We always have:

\[ F_1(t) \leq F_1^0(t) = 1 - S_1(t). \]

The risk is over-estimated by using 1-KM instead of Aalen-Johansen.

The degree of bias depends on the magnitude of the competing risk (cause 2): if there are no competing risks (\( \lambda_2(t) = 0 \)) then they are identical, and the difference between the two increases with \( \lambda_2(t) \).

At best, the simple 1-KM estimator can be considered an approximation to the cumulative incidence that may be used if the competing risk is small. However, the best advice is never to use 1-KM in the presence of competing risks.
Kaplan-Meier vs. Nelson-Aalen

Why does Nelson-Aalen work with competing risks when Kaplan-Meier doesn’t?

This has to do with the fact that the Nelson-Aalen estimates the (cumulative) *rate*, and rates describe the “local” (in time) behavior of the failure process for a given cause. “Therefore”, when assessing the local strength of cause 1, then cause 2 need not be taken into account.

*Risks*, however, cumulate over (long) time periods and the impact of cause 2 must be accounted for when assessing the strength of cause 1.

More on this later
Introduction to Competing Risks

**Inference for cause-specific hazards**

As a consequence, all standard hazard-based models for survival data apply when analyzing *cause-specific hazards*

- non-parametric: estimate \( \Lambda_j(t) = \int_0^t \lambda_j(u) du, j = 1, 2 \) by Nelson-Aalen estimator, compare using, e.g. logrank tests
- Cox regression
- Poisson regression

More on this later today.
Aalen-Johansen estimator using macro %CIF (SAS macro library). The macro requires SAS/IML.

```sas
%CIF(DATA=dropout, TIME=week, STATUS=state, EVENT=1, CENSORED=0, GROUP=drug);
```

The estimates are saved in a data set.

The %CIF macro has been included in PROC LIFETEST in SAS/STAT 14.2.
Aalen-Johansen estimator

```
PROC LIFETEST DATA=dropout PLOTS=CIF;
   TIME week*state(0) / FAILCODE=1;
   STRATA drug;
RUN;
```

Make CIF for all three causes:

```
PROC LIFETEST DATA=dropout PLOTS=CIF;
   TIME week*state(0) / FAILCODE;
   STRATA drug;
RUN;
```
PROC PHREG does not use the Aalen-Johansen estimator but the Fine-Gray model for cumulative incidence regression to estimate CIF.
For the Cox model from standard survival analysis we have

$$\log(-\log(S(t \mid X))) = \log(\Lambda_0(t)) + \beta'X.$$  

The sub-distribution version for competing risks is

$$\log(-\log(1 - F_j(t \mid X))) = \log(\tilde{\Lambda}_{0j}(t)) + \tilde{\beta}'X.$$  

PROC PHREG uses an estimator for $\tilde{\Lambda}_{0j}(t)$ which involves Inverse Probability of Censoring Weights (IPCW) for subjects who have experienced a competing event – more on this later.
Exercise: Drug discontinuations as competing risks

Data on time to drug discontinuation for different reasons in a 1-year RCT (N=559) of active drug vs placebo.

<table>
<thead>
<tr>
<th>week</th>
<th>Time in study in weeks (!)</th>
</tr>
</thead>
<tbody>
<tr>
<td>state</td>
<td>0 = Complete on drug (n=406)</td>
</tr>
<tr>
<td></td>
<td>1 = Adverse event (n=42)</td>
</tr>
<tr>
<td></td>
<td>2 = Withdrew consent (n=62)</td>
</tr>
<tr>
<td></td>
<td>3 = Other (n=49)</td>
</tr>
<tr>
<td>drug</td>
<td>0 = Placebo (n=188)</td>
</tr>
<tr>
<td></td>
<td>1 = Active (n=371)</td>
</tr>
</tbody>
</table>
Drug discontinuations – competing risks

On drug

- \( \lambda_1(t) \) to Adverse Event
- \( \lambda_2(t) \) to Withdrew Consent
- \( \lambda_3(t) \) to Other
Example: Drug discontinuations as competing risks

Data on time to drug discontinuation for different reasons in a 1-year RCT (N=559) of active drug vs placebo.

<table>
<thead>
<tr>
<th>week</th>
<th>Time in study in weeks (!)</th>
</tr>
</thead>
<tbody>
<tr>
<td>state</td>
<td>0 = Complete on drug (n=406)</td>
</tr>
<tr>
<td></td>
<td>1 = Adverse event (n=42)</td>
</tr>
<tr>
<td></td>
<td>2 = Withdrew consent (n=62)</td>
</tr>
<tr>
<td></td>
<td>3 = Other (n=49)</td>
</tr>
<tr>
<td>drug</td>
<td>0 = Placebo (n=188)</td>
</tr>
<tr>
<td></td>
<td>1 = Active (n=371)</td>
</tr>
</tbody>
</table>
Example: Drug discontinuations as competing risks

- Adverse Event
- Withdrew Consent
- Other

On drug

\[ \lambda_1(t) \]

\[ \lambda_2(t) \]

\[ \lambda_3(t) \]
Overall: 1-KM
Placebo: 1-KM
Placebo: CIF
Placebo: Stacked 1-KM
Placebo: Stacked CIF

![Chart showing CIF stacked for different categories over time. The chart includes lines for All, AE, AE+WC, and AE+WC+OT, with the x-axis representing week and the y-axis representing CIF.](chart.png)
In studies of all-cause mortality, risks (probabilities, cumulative incidences) can be computed from rates (hazards) and vice versa - in other words the two functions contain equivalent information.

In studies of events which will not eventually happen for everyone in the population, this is no longer the case and death (and maybe other events) are competing risks which need to be addressed.

In such cases, the risk of a given cause depends on the rates for all competing causes.

Therefore, using ‘1-Kaplan-Meier for a single cause’ as a risk estimator is (upward) biased.

The magnitude of the bias depends on the frequency of the competing events.

A rather simple, unbiased estimator for the risk exists - the ”Aalen-Johansen” estimator.