XVII Summer School of the Master's degree in Statistics and Operations Research

Multi-state models: Rates, risks, and pseudo-values

I: **Introduction to multi-state models; non-parametric estimation**

Per Kragh Andersen and Henrik Ravn

https://multi-state-book.github.io/barcelona2024/

Overview of course

I Introduction to multi-state models; non-parametric estimation

II Regression models for intensities (Cox)

III Direct regression models for marginal parameters (Cox, Fine-Gray, Ghosh-Lin, ELOS)

IV Pseudo-values (1)

V Pseudo-values (2)

The course is based on the book 'Models for Multi-State Survival Data: Rates, Risks, and Pseudo-Values' by PKA and HR.

I: Introduction to multi-state models

- Multi-state models: marginal parameters (state occupation probabilities, expected length of stay in a state, 'ELOS') and transition intensities.
  - Two-state model for survival data, competing risks
  - Recurrent events, illness-death model
- Observations: counting processes and at risk processes; censoring
- Non-parametric inference for intensities: the Nelson-Aalen estimator
- Plug-in for marginal parameters: Kaplan-Meier and Aalen-Johansen estimators, ELOS, expected number of recurrent events before time t

A multi-state model

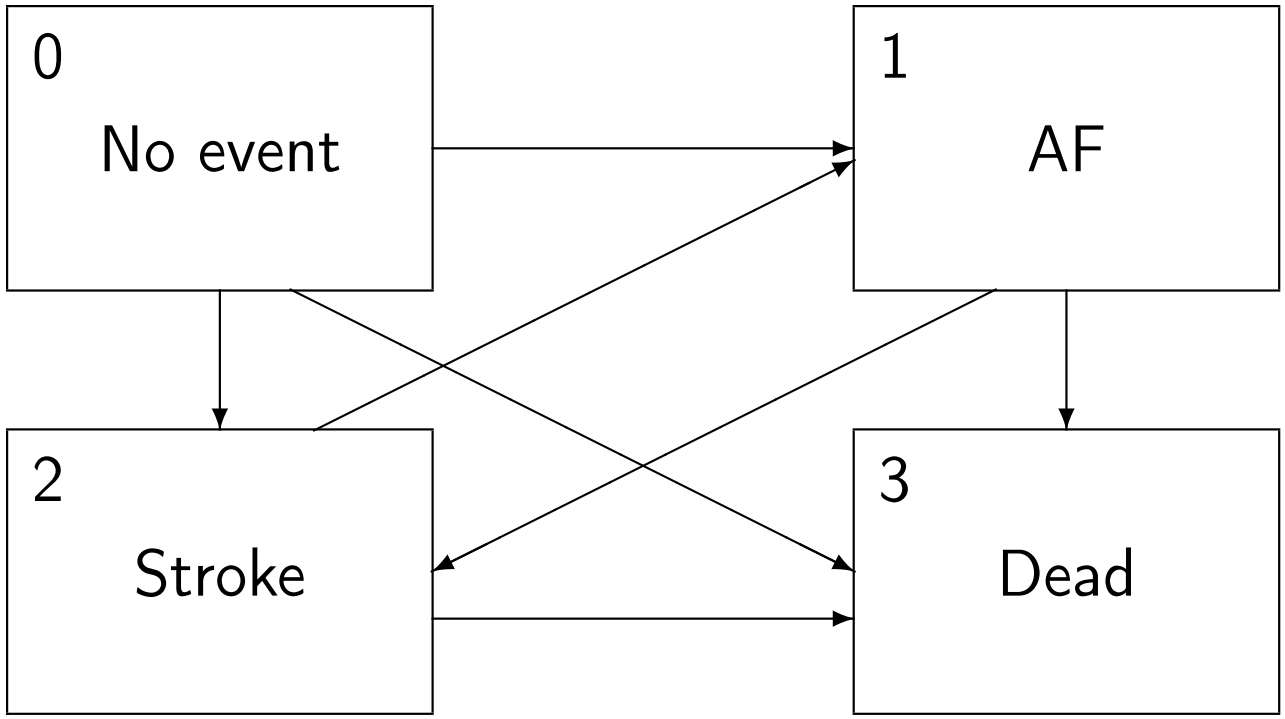


Figure 1: Copenhagen Holter study: States and transitions (AF: Atrial fibrillation, a serious heart arrhythmia).

The Copenhagen Holter study

- Larsen et al. (2015, *J. Amer. College Cardiol.*)
- 678 subjects from Copenhagen aged 55, 60, 65, 70 or 75 with 0, 1, > 1 cardiovascular risk factors were recruited 1998-2000
- Initial examination including a 48-hour continuous ECG monitoring ('Holter')
- From Holter monitoring, presence of ESVEA (excessive supra-ventricular ectopic activity, a certain heart arrhythmia) was ascertained
- Register-based follow up until 2013 with respect to atrial fibrillation (AF, 77), stroke (73), and death (261)
- Purpose: examine the impact of ESVEA on AF and stroke

Parameters in multi-state models

- $V(t) \in \mathcal{S} = \{0, 1, \ldots, k\}$ *state* occupied at time t

  Marginal parameters:

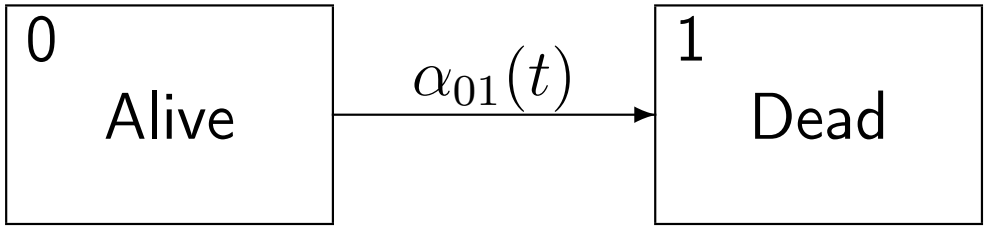
- $Q_h(t) = P(V(t) = h) = E(I(V(t) = h)), h \in \mathcal{S}$ state occupation probabilities
- $\varepsilon_h(\tau) = E(\int_0^\tau I(V(t) = h)dt) = \int_0^\tau Q_h(t)dt$ *expected length of stay* in state $h \in \mathcal{S}$ in $[0, \tau]$

  Conditional parameters:

- $P_{hj}(s, t) = P(V(t) = j \mid V(s) = h,$ past information in $[0, s))$ transition probabilities
- $\alpha_{hj}(t) = \lim_{dt \to 0} P_{hj}(t, t + dt)/dt$ transition intensities

The transition intensities (hazards, rates) are the basic building blocks for multi-state models, but marginal parameters often have more direct interpretations.

Two-state model for survival data



Transition intensity: *hazard function*
$\alpha(t) = \alpha_{01}(t) \approx P(\text{state } 1 \text{ time } t + dt \mid \text{state } 0 \text{ time } t)/dt.$

State occupation probabilities: *survival function*,
$Q_0(t) = S(t) = P(\text{state 0 time } t) = P(T > t),$

and *cumulative probability of death before time* t, Eq. (1.2):

$Q_1(t) = 1 - Q_0(t) = F(t) = P(\text{state 1 time } t) =$
$1 - \exp(-\int_0^t \alpha_{01}(u)du).$

The PBC-3 trial in liver cirrhosis

- Lombard et al. (1993, *Gastroenterol.*)
- Multi-centre randomized trial in patients with primary biliary cirrhosis.
- Patients ($n = 349$) recruited 1 Jan, 1983 - 1 Jan, 1987 from six European hospitals and randomized to CyA (176) or placebo (173).
- Followed until death or liver transplantation (no longer than 31 Dec, 1989); CyA: 30 died, 14 were transplanted; placebo: 31 died, 15 were transplanted; 4 patients were lost to follow-up before 1989.
- Primary outcome variable: time to death, incompletely observed, due to: liver transplantation, loss to follow-up, alive 31 Dec, 1989.
- In some analyses, the outcome is defined as 'time to failure of medical treatment', i.e., time to the composite end-point of either death or liver transplantation, whatever came first.

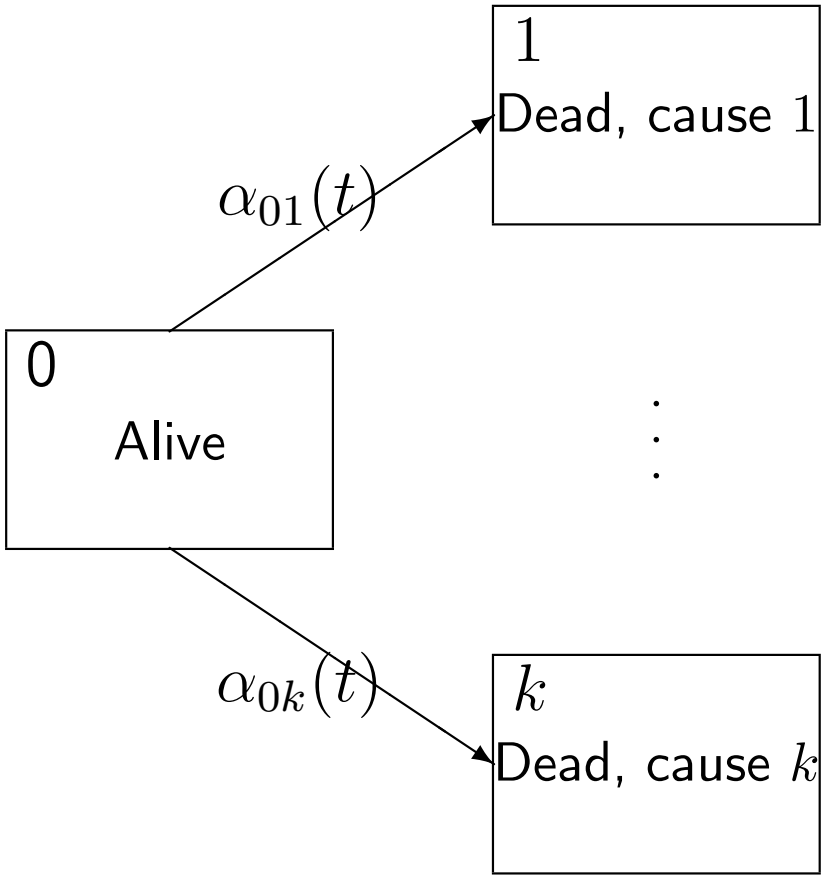
The PBC-3 trial in liver cirrhosis

Variable	Description
`id`	patient id
`unit`	hospital
`days`	follow-up time in days
`status`	0 = censoring, 1 = transplantation, 2 = death without transplantation
`tment`	0 = placebo, 1 = CyA
`sex`	0 = female, 1 = male
`age`	age (years)
`bili`	bilirubin (micromoles/L)
`alb`	albumin (g/L)
`stage`	disease stage: 2 = I-II, 3 = III, 4 = IV

Examples of survival data

- Copenhagen Holter study: stroke-free survival (i.e., disregarding AF)
- PBC-3 trial: composite end-point

Competing risks model



Competing risks model

Transition intensities: *cause-specific hazards* $h = 1, \ldots, k$:

$$\alpha_h(t) = \alpha_{0h}(t) \approx P(\text{state } h \text{ time } t + dt \mid \text{state } 0 \text{ time } t)/dt.$$

State occupation probabilities: *overall survival function*:

$Q_0(t) = S(t) = P(\text{alive time } t)$,

$= \exp(-\int_0^t (\alpha_{01}(u) + \cdots + \alpha_{0k}(u))du)$

and *cumulative incidences* $h = 1, \ldots, k$, Eq. (1.3):

$$Q_h(t) = F_h(t) = P(\text{dead from cause } h \text{ before time } t) = \int_0^t S(u)\alpha_{0h}(u)du.$$

Examples of competing risks data

- Copenhagen Holter study: stroke and death without stroke (i.e., disregarding AF)
- PBC-3 trial: liver transplantation and death without liver transplantation

Expected length of stay

In both models, the t_0-restricted mean survival time (RMST), $\varepsilon_0(t_0) = E(\min(T, t_0))$ (T is the survival time), Eq. (1.10), is:

$$\varepsilon_0(t_0) = \int_0^{t_0} S(t)dt.$$

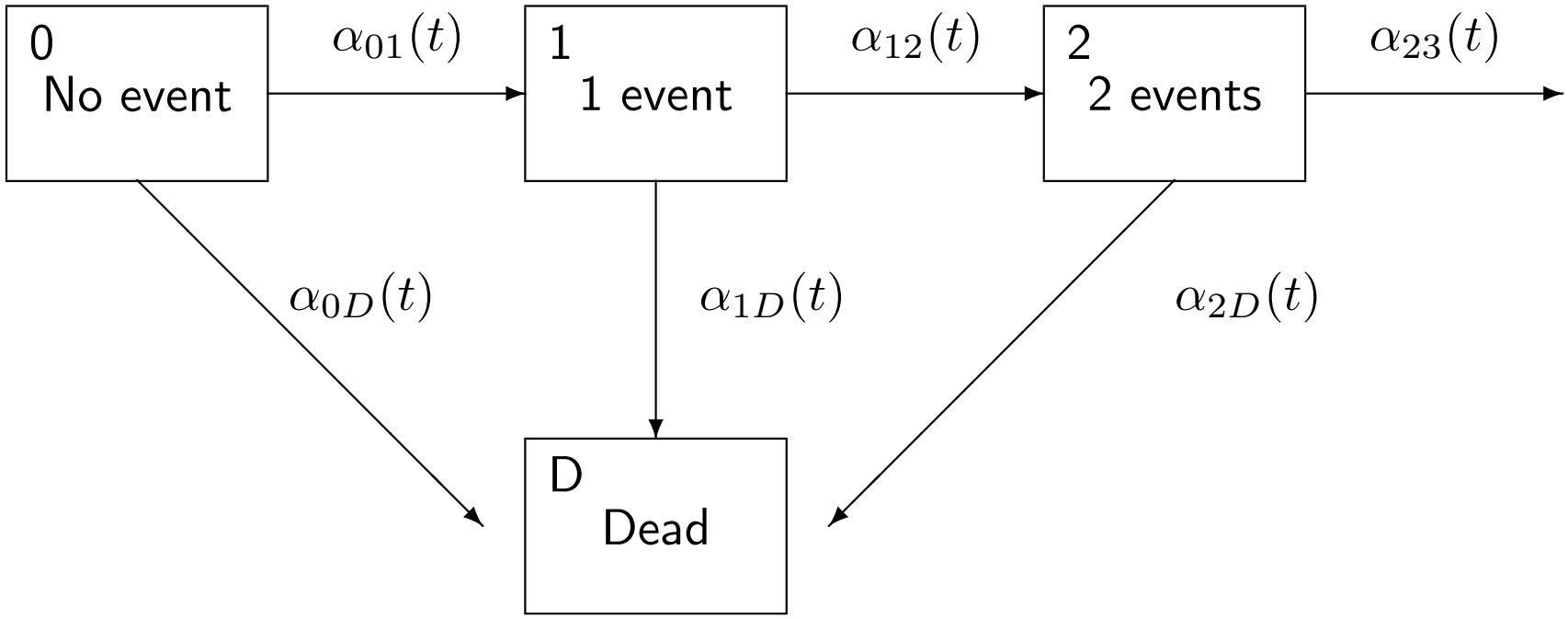
In the two-state survival model: $\varepsilon_1(t_0) = \int_0^{t_0} Q_1(t)dt$ is the expected number of years lost (YL) before time t_0, i.e. $t_0 - E(\min(T, t_0))$.

In the competing risks model,

$$\varepsilon_h(t_0) = \int_0^{t_0} Q_h(t)dt$$

is the expected number of years lost 'due to cause h' before time t_0.

Recurrent events (with competing risks)



Recurrent events: marginal parameters

The most important marginal parameter is

$$\mu(t) = E(N(t)) = \int_0^t S(u)\alpha^*(u)du,$$

with $N(t) =$ number of events in $[0, t]$, and $\alpha^*(\cdot)$ the *marginal rate function given survival*

$$\alpha^*(t) \approx E(dN(t) \mid T > t)/dt.$$

In the model without the final death state, $\mu(t) = \int_0^t \alpha^*(u)du$ with $\alpha^*(\cdot)$ now being the marginal rate function $\alpha^*(t) \approx E(dN(t))/dt$.

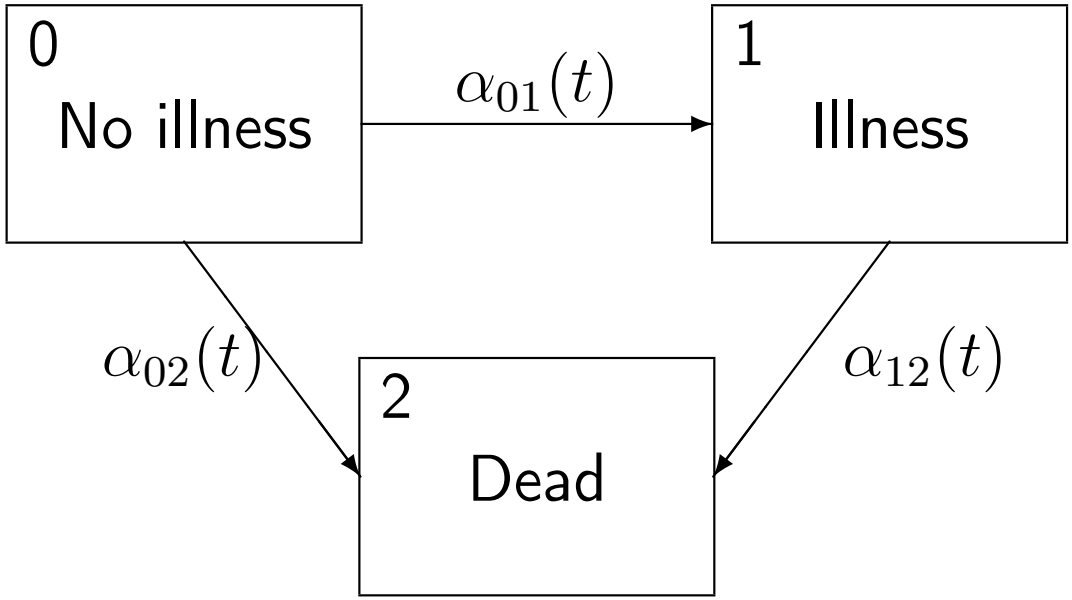
Example: Recurrent episodes in affective disorder

- Kessing et al. (2004, *Acta Psych. Scand.*)
- 119 patients with 'unipolar' (depressive, 98) or 'bipolar' (manic-depressive, 21) disorder had their first episode recorded 1959-63 at hospital in Zürich, Switzerland
- Followed up until 1985 w.r.t. new episodes (on average 5.6) and death (78)
- Purpose: study how repeated episodes depend on initial diagnosis (unipolar vs. bipolar)

Recurrent episodes in affective disorders

Variable name	Description
`id`	patient id
`episode`	number of affective episodes
`state`	Status at time `start`:
	0 = no current affective episode, 1 = current affective episode
`start`	start time in state (months)
`stop`	last time seen in state (months)
`status`	status at time `stop`:
	0 = transition to state 0
	1 = transition to state 1
	2 = transition to death
	3 = censoring
`prev`	'start' of time to next transition to state 1, even if in state 1
`bip`	0 = unipolar, 1 = bipolar
`sex`	0 = female, 1 = male
`age`	age (years)
`year`	year of initial episode

The irreversible illness-death model



Examples of illness-death model data

- Copenhagen Holter study: no event, stroke, death with or without stroke (i.e., still disregarding AF)
- PBC-3 trial: no event, liver transplantation, death with or without liver transplantation ('in principle' – information after liver transplantation is not available)

Observations

Observation of

$$(V_i(t), t \in [0, \tau_i], i = 1, \ldots, n),$$

(where τ_i is either the time when $V_i(\cdot)$ reaches an *absorbing state*, or a time C_i of *right-censoring*) can be represented by *counting processes*:

$$N_{hji}(t) = \text{ number of direct } h \to j \text{ transitions } (h \neq j)$$

$$\text{observed in } [0, t] \text{ for subject } i = 1, \ldots, n,$$

and *at risk processes*

$$Y_{hi}(t) = \text{ indictor for } i \text{ being observed in state } h \text{ at time } t - .$$

Independent censoring

We will assume throughout that censoring is *independent*, i.e.,

$$\frac{P(V(t+dt)=j \mid V(t)=h, \text{ past for } s<t \textbf{ and } \boldsymbol{C>t})}{dt} \approx \alpha_{hj}(t)$$

(Eq. (1.6)). In other words, the additional knowledge that, at time t, a subject is not only at risk for a $h \to j$ transition, but also uncensored should not alter the intensities

Censoring by liver transplantation in the PBC-3 trial?

The Nelson-Aalen estimator

If we assume that the intensity $\alpha_{hj}(t)$ is the same for all subjects and independent on the past (Markov assumption), then a natural estimator for

$$\alpha_{hj}(t)dt \approx P(V(t+dt) = j \mid V(t) = h)$$

is

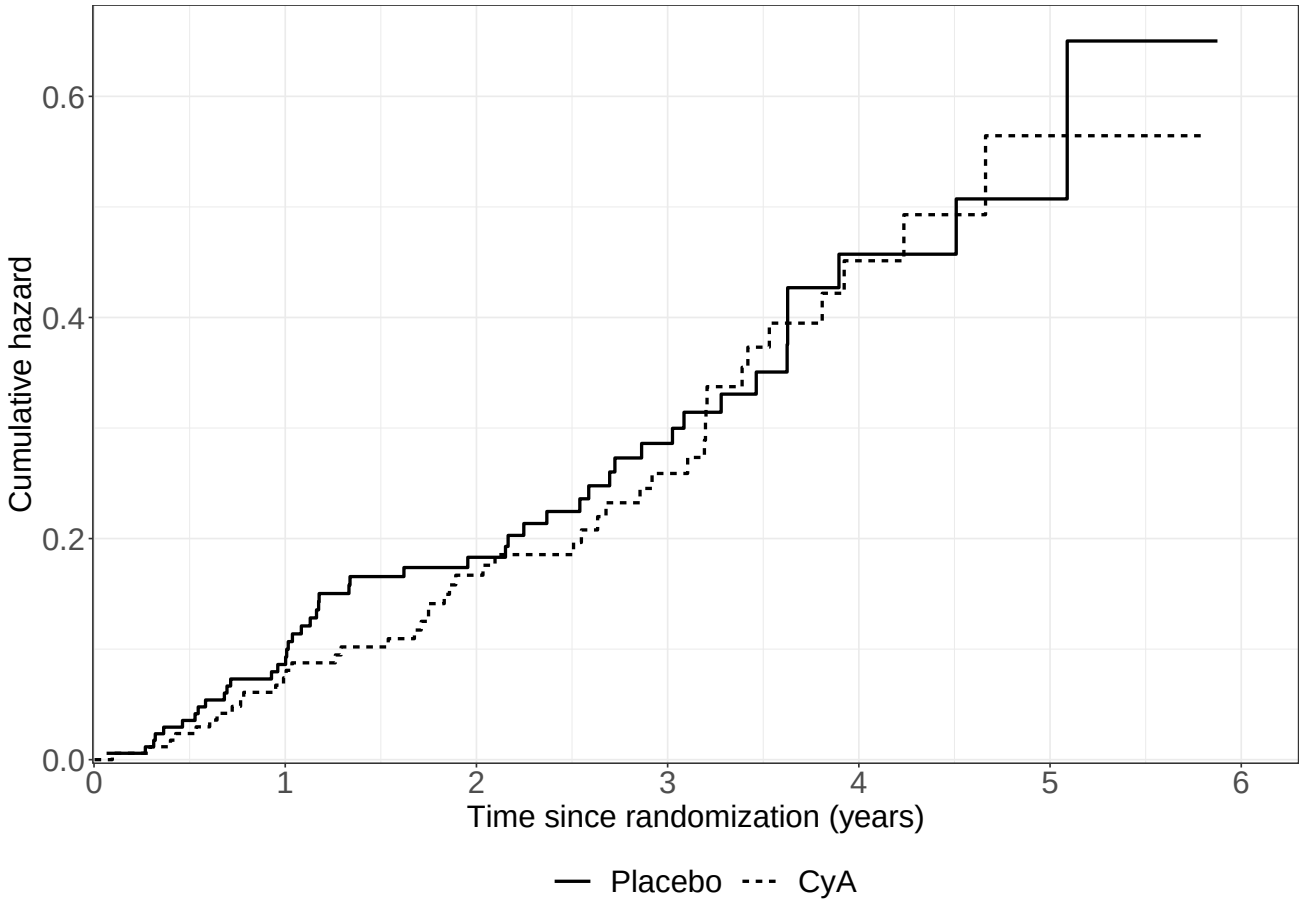
$$\widehat{\alpha_{hj}(t)dt} = \frac{\sum_i dN_{hji}(t)}{\sum_i Y_{hi}(t)},$$

leading to the *Nelson-Aalen* estimator, Eq. (3.10), for the cumulative intensity $A_{hj}(t) = \int_0^t \alpha_{hj}(u)du$:

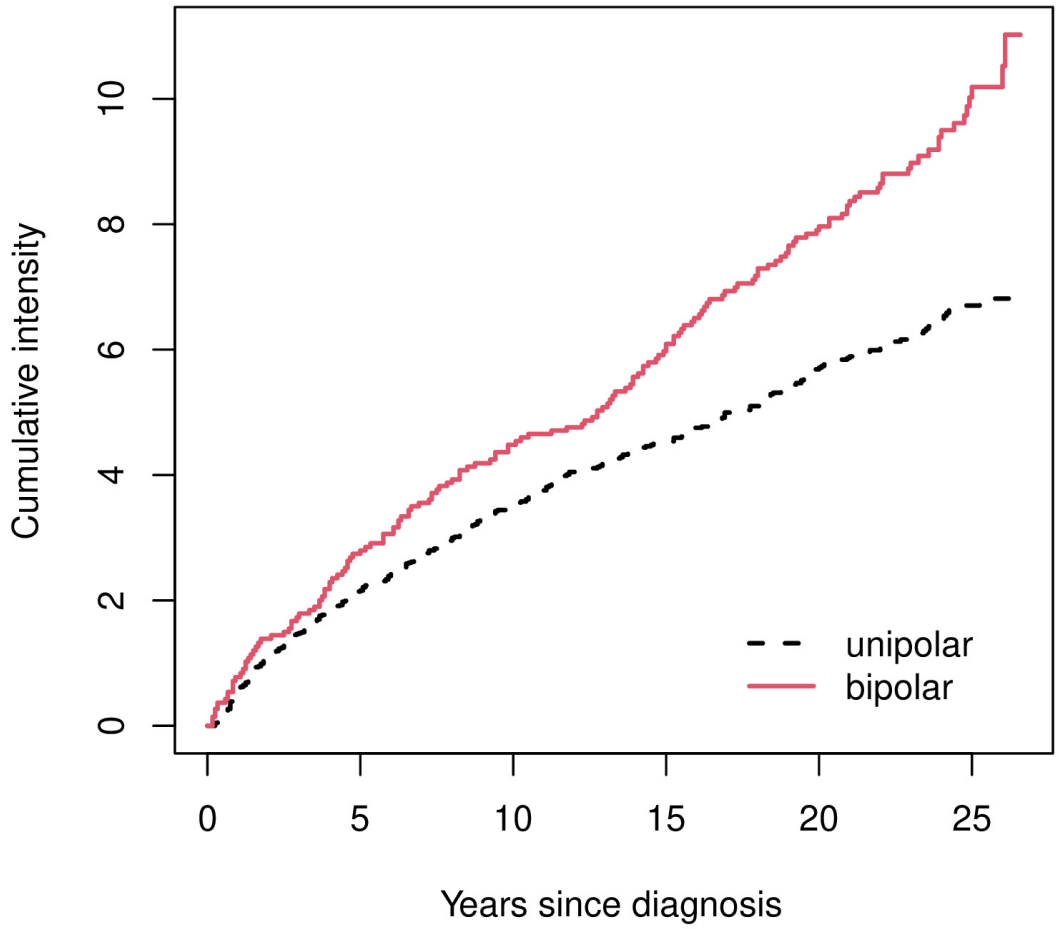
$$\widehat{A}_{hj}(t) = \int_0^t \frac{\sum_i dN_{hji}(u)}{\sum_i Y_{hi}(u)},$$

an increasing step function with steps at observed $h \to j$ transition times. The local slope estimates the intensity. The estimator has a maximum likelihood interpretation.

The PBC-3 trial, composite end-point



Repeated episodes in affective disorder



R code, cumulative hazards

```
library(survival)
pbcna <- survfit(Surv(days, status != 0) ~ tment, data = pbc3)
plot(pbcna,fun="cumhaz")

affna <- survfit(Surv(start, stop, status==1) ~ bip,
           data = subset(affective,state==0))
plot(affna,fun="cumhaz")
```

This code gives crude versions of the figures that may be improved upon by adding options to `plot` etc.

Estimation of marginal parameters: Two-state model

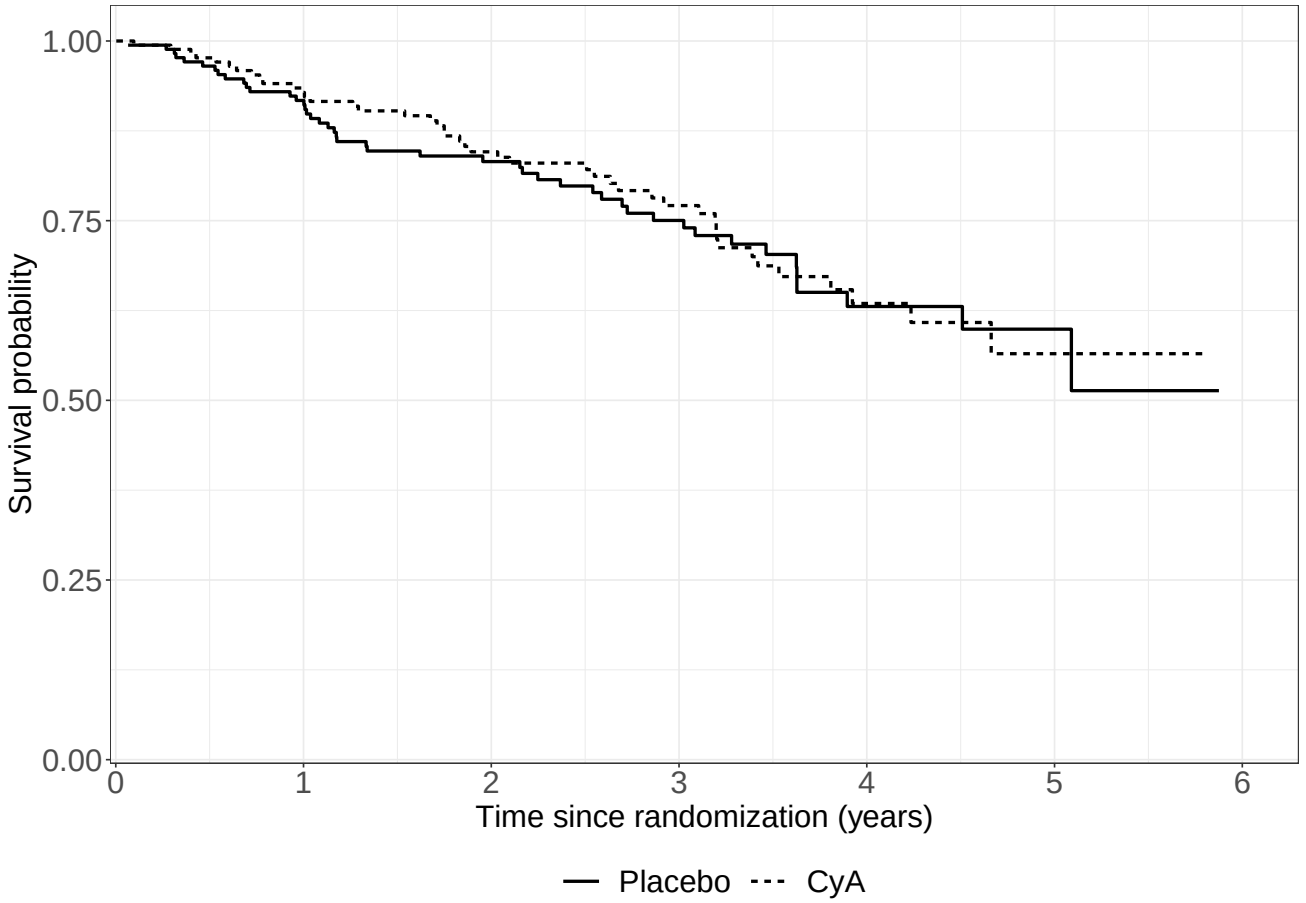
The intensities are the basic parameters in multi-state models and marginal parameters, such as state occupation probabilities $Q_h(t)$ may be estimated by *plug-in* (or by simulation).

For the two-state model, this leads to the Kaplan-Meier (1958) estimator, Eq. (4.3), for $S(t) = Q_0(t)$:

$$\widehat{S}(t) = \prod_{u \leq t} \left(1 - \frac{\sum_i dN_{01i}(u)}{\sum_i Y_{0i}(u)} \right).$$

Note that we do not use $\exp(-\widehat{A}_{01}(t))$. This is because the 'exponential formula' $S(t) = \exp(-A_{01}(t))$ only holds for *continuous distributions* while, in general, the relationship is given by a 'product-integral'.

The PBC-3 trial, composite end-point



Estimation of marginal parameters: competing risks

The cause-h cumulative incidence is $Q_h(t) = \int_0^t S(u)\alpha_{0h}(u)du$, and the plug-in estimator is the *Aalen-Johansen* estimator, Eq. (4.9):

$$\widehat{Q}_h(t) = \int_0^t \widehat{S}(u-)d\widehat{A}_{0h}(u),$$

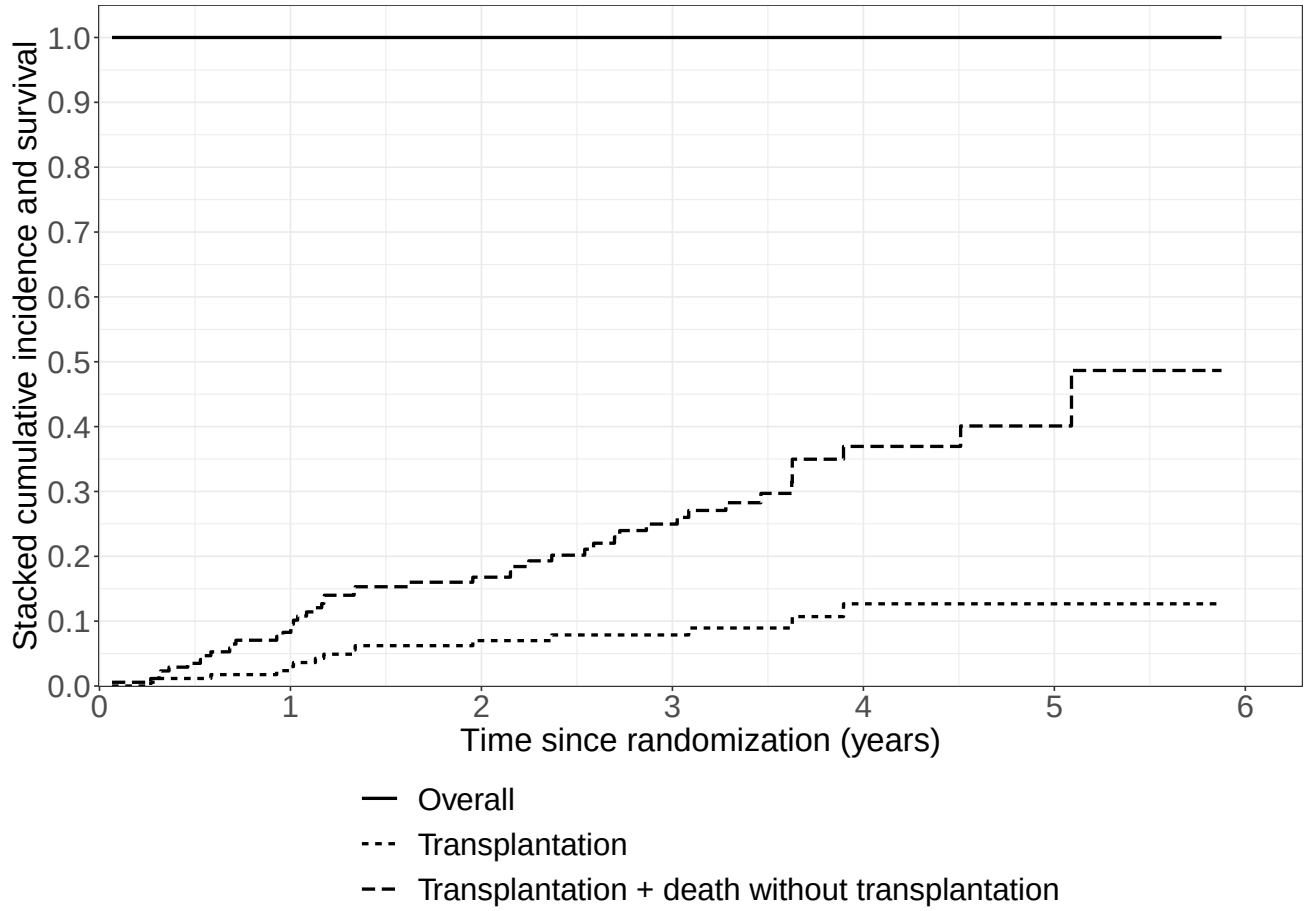
with $\widehat{S}$, the Kaplan-Meier estimator for overall survival (i.e., counting all deaths) and $\widehat{A}_{0h}$ the Nelson-Aalen estimator for the cumulative cause-h-specific hazard.

The quantity $Q_h(t_0)$ is the t_0-year *risk* of cause h from which t_0-year *risk differences* or *risk ratios* may be estimated. We will study inference for these contrasts later in the course using *pseudo-values*.

Note that a naive '1-Kaplan-Meier estimator', say $1 - \widehat{S}_h(t)$, counting only cause-h events (and censoring for the competing risks) is upwards biased

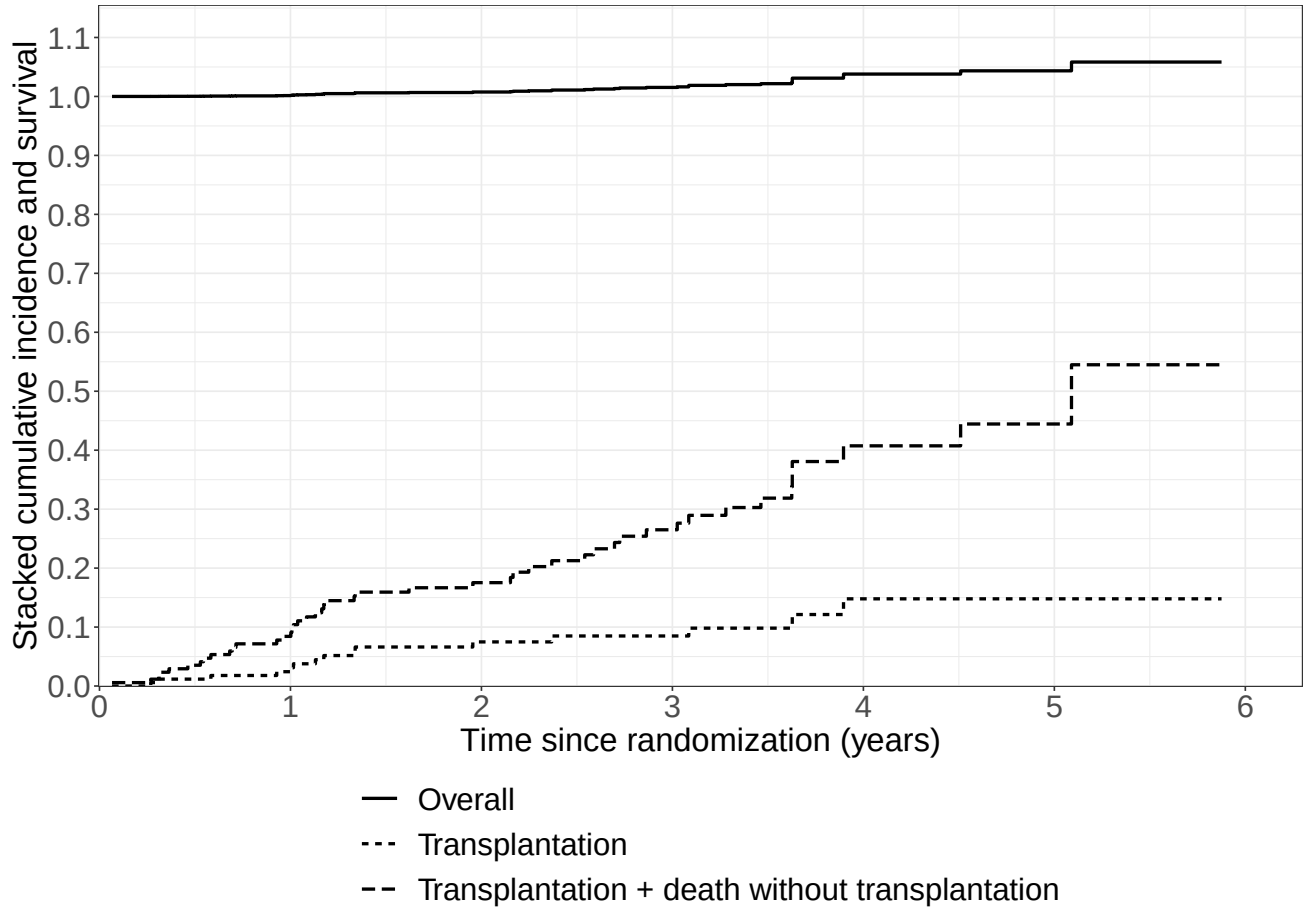
$$\widehat{Q}_h(t) \leq 1 - \widehat{S}_h(t).$$

The PBC-3 trial, death and transplantation



Aalen-Johansen estimator (placebo) – correct!

The PBC-3 trial, death and transplantation



1 minus Kaplan-Meier estimator (placebo) – biased!

Estimation of general state occupation probabilities

Both the Kaplan-Meier and Aalen-Johansen estimators are special cases of a general plug-in estimator of *transition probabilities* in Markov processes due to Aalen and Johansen (1978).

Datta and Satten (2001) showed that for general multi-state processes (i.e., also non-Markov), this estimator is consistent for *state occupation probabilities*.

ELOS

$$\varepsilon_h(t_0) = \int_0^{t_0} Q_h(t)dt$$

may also be estimated by plug-in.

The PBC-3 trial, death and transplantation

	Placebo	CyA
RMST, 3 years	2.606	2.678
Years lost, 3 years		
Transplantation	0.143	0.086
Death without transpl.	0.251	0.236
	3.000	3.000

Estimates are areas under Kaplan-Meier, resp. Aalen-Johansen estimates at t_0 =3 years.

R code, K-M, Aa-J, ELOS

```
library(survival)
pbcna <- survfit(Surv(days, status != 0) ~ tment, data = pbc3)
#NB: same as for N-Aa

plot(pbcna)

newstatus<-as.factor$status
ajfit <- survfit(Surv(days, newstatus) ~ tment, data = pbc3)
plot(ajfit)

print(ajfit,rmean=3*365)
```

Again, the code gives a crude (non-stacked) version of the correct figure that may be improved upon.

Estimation of marginal parameters: recurrent events

For recurrent events with competing risks, plug-in leads to the Cook-Lawless (1997) estimator, Eq. (4.13):

$$\widehat{\mu}(t) = \int_0^t \widehat{S}(u-) d\widehat{A}^*(u)$$

($\widehat{S}, \widehat{A}^*$: Kaplan-Meier and Nelson-Aalen, respectively).

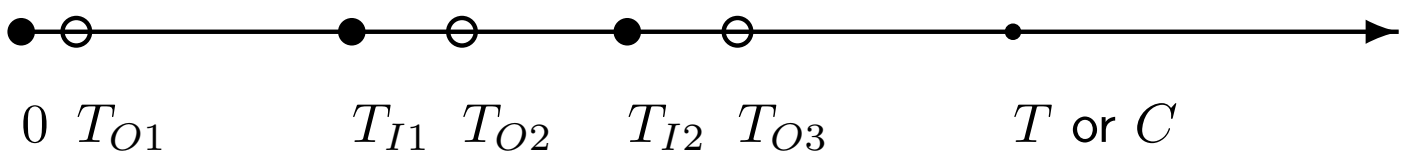
Without competing risks (i.e., $S(t) = 1$), the Nelson-Aalen estimator estimates $\mu(t)$.

Note that, as it was the case for competing risks, treating deaths as censorings (and estimating $\mu(t)$ by Nelson-Aalen), an upwards biased estimator is obtained.

Recurrent events: periods not at risk ('episodes')

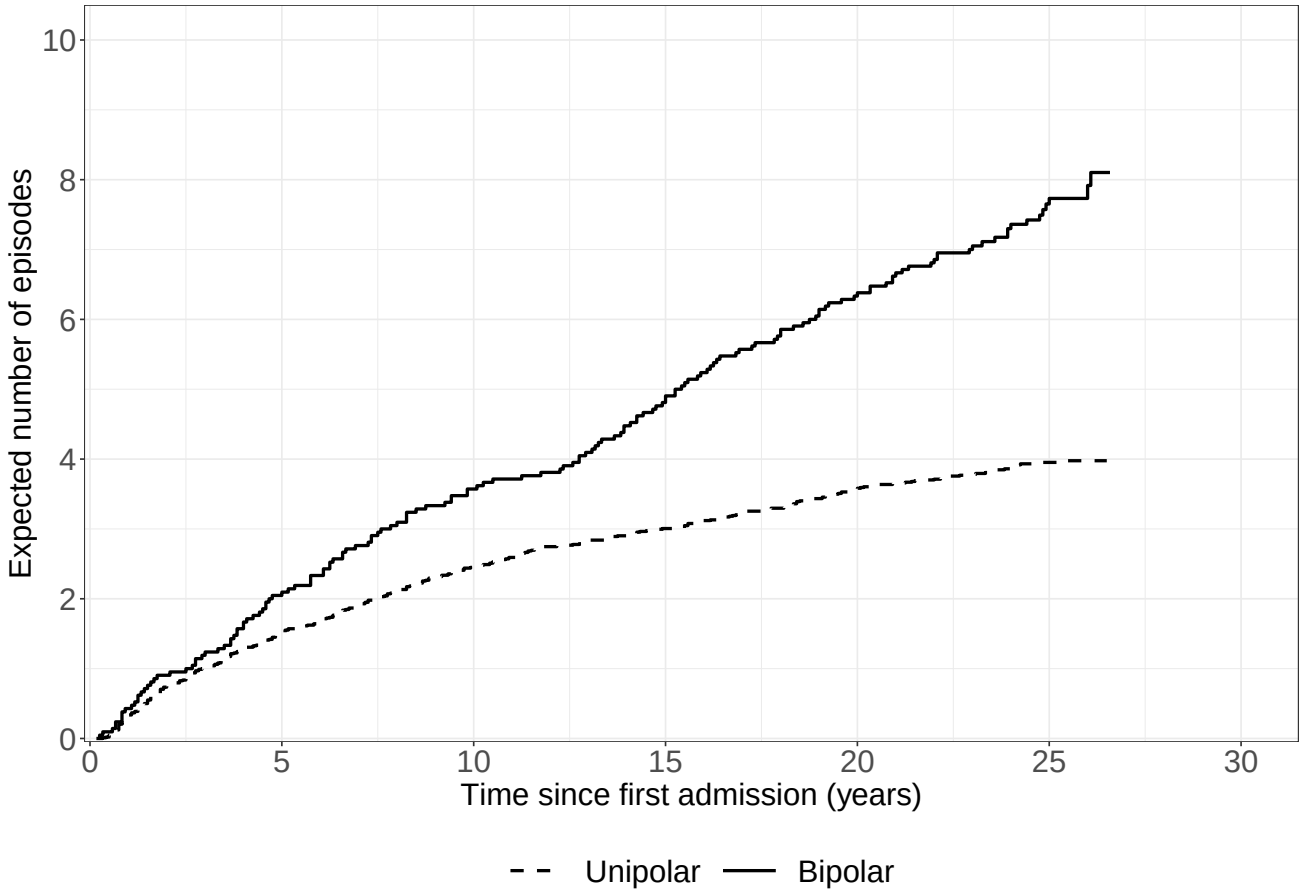
For the data on repeated episodes, there are periods of on-going episodes where patients are not at risk for a new episode ('in hospital'), and these periods were accounted for when estimating the cumulative intensity.

The simplest way of taking these periods into account when estimating the marginal mean is to ignore them – estimate the expected number of 'bullets' in $[0, t]$ using the Nelson-Aalen/Cook-Lawless estimator and ignore the 'circles':



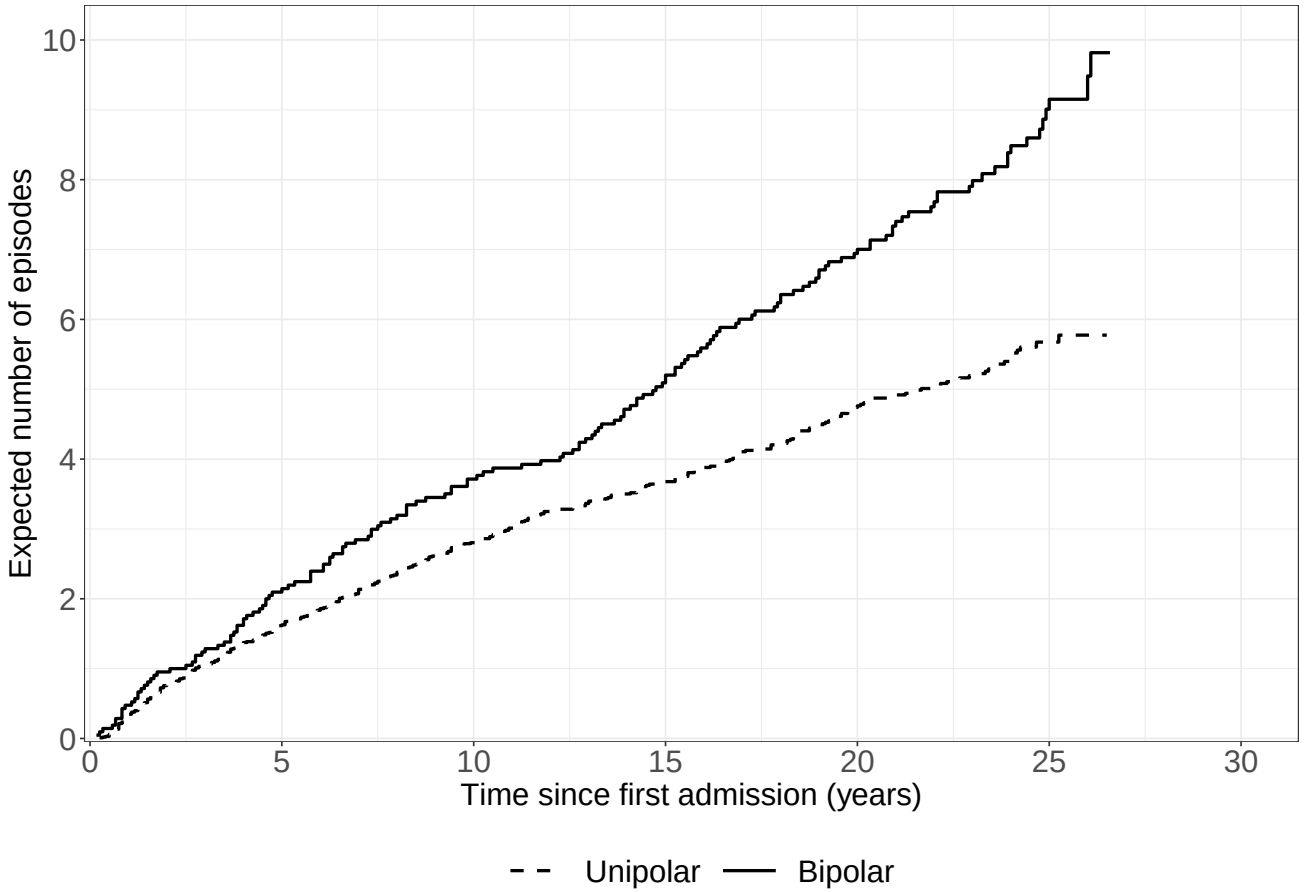
Supplement by analysis of time 'in hospital'/mortality: one way of getting few re-admissions is to keep the patient in hospital/kill him!

Repeated episodes in affective disorder



Cook-Lawless estimator – correct!

Repeated episodes in affective disorder



Nelson-Aalen estimator – biased!

R code (prev is start of 'episode')

```
library(mets)
library(survival)
# Cook-Lawless estimator (bipolar patients)
newstatus<-ifelse(status==1,1,0)
death<-ifelse(status==2,1,0)
xr1 <- phreg(Surv(prev,stop,newstatus)~cluster(id),
      data=subset(affective,bip==1))
dr1 <- phreg(Surv(prev,stop,death)~cluster(id),
      data=subset(affective,bip==1))
out1 <- recurrentMarginal(xr1,dr1)
plot(out1$time,out1$mu)
# Nelson-Aalen
affna <- survfit(Surv(prev, stop, status==1) ~ bip,
           data = subset(affective,state==0))
plot(affna,fun="cumhaz")
```

Non-parametric tests

- The standard non-parametric test for comparison of *intensities* is the *log-rank* test

```
library(survival)
survdiff(Surv(days, status != 0) ~ tment, data = pbc3)
```

- For competing risks data, there is a Gray (1988) test for comparison of cumulative incidences
- For recurrent events with competing risks, there is a Ghosh-Lin (2000) test for comparison of mean value functions.

Exercises

1. Consider the Copenhagen Holter study and estimate non-parametrically the cumulative hazards for stroke-free survival for subjects with and without ESVEA. Compare the two using the logrank test.
2. Repeat the previous exercise, now looking instead at the competing end-points stroke and death without stroke.
3. Consider the Copenhagen Holter study and estimate non-parametrically the probabilities of stroke-free survival for subjects with and without ESVEA.
4. Estimate non-parametrically the cumulative incidences of stroke and death without stroke for subjects with and without ESVEA.

   Estimate also the 10-year restricted mean stroke-free survival times and the expected number of years lost due to stroke or death without stroke for subjects with and without ESVEA.