Causal inference with time-to-event outcomes under competing risk

Jon Michael Gran

j.m.gran@medisin.uio.no

Oslo Centre for Biostatistics and epidemiology,
Department of biostatistics, University of Oslo

September 10th, 2024

Causal inference with time-to-event outcomes under competing risk

Jon Michael Gran

j.m.gran@medisin.uio.no

Oslo Centre for Biostatistics and epidemiology,
Department of biostatistics, University of Oslo

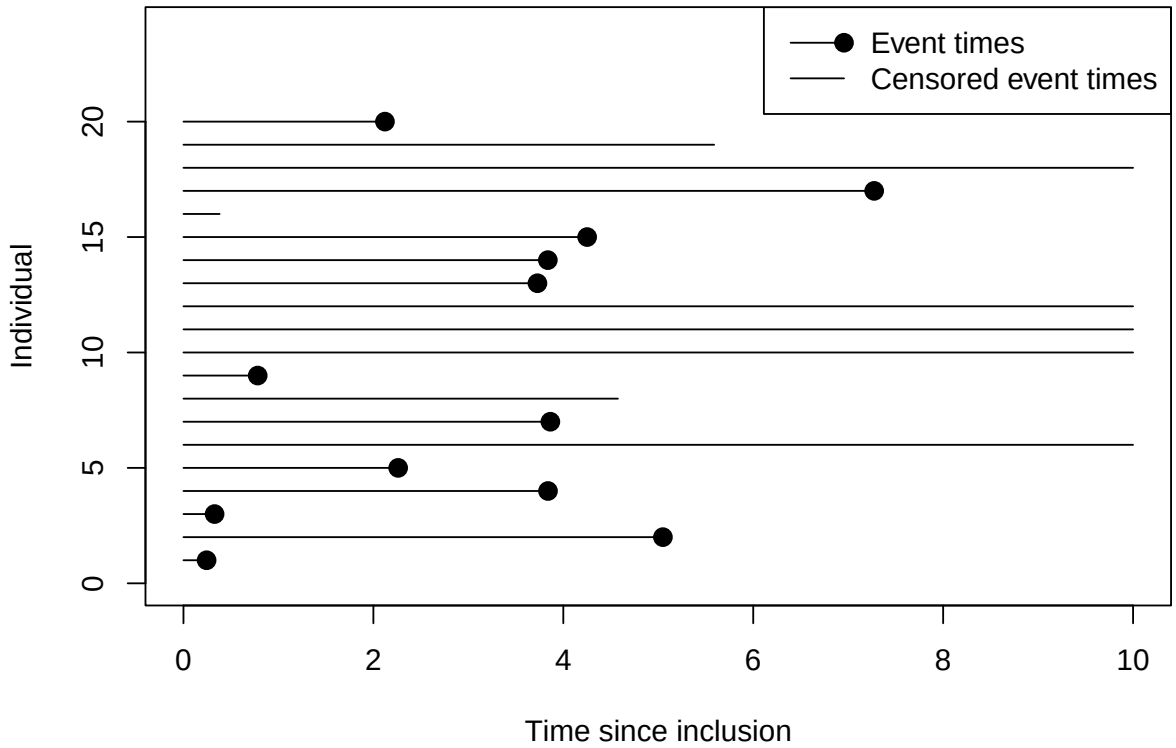
September 10th, 2024

Thanks to Bjarte Aagnes, Tor Åge Myklebust and Paul Lambert for Stata input

Outline

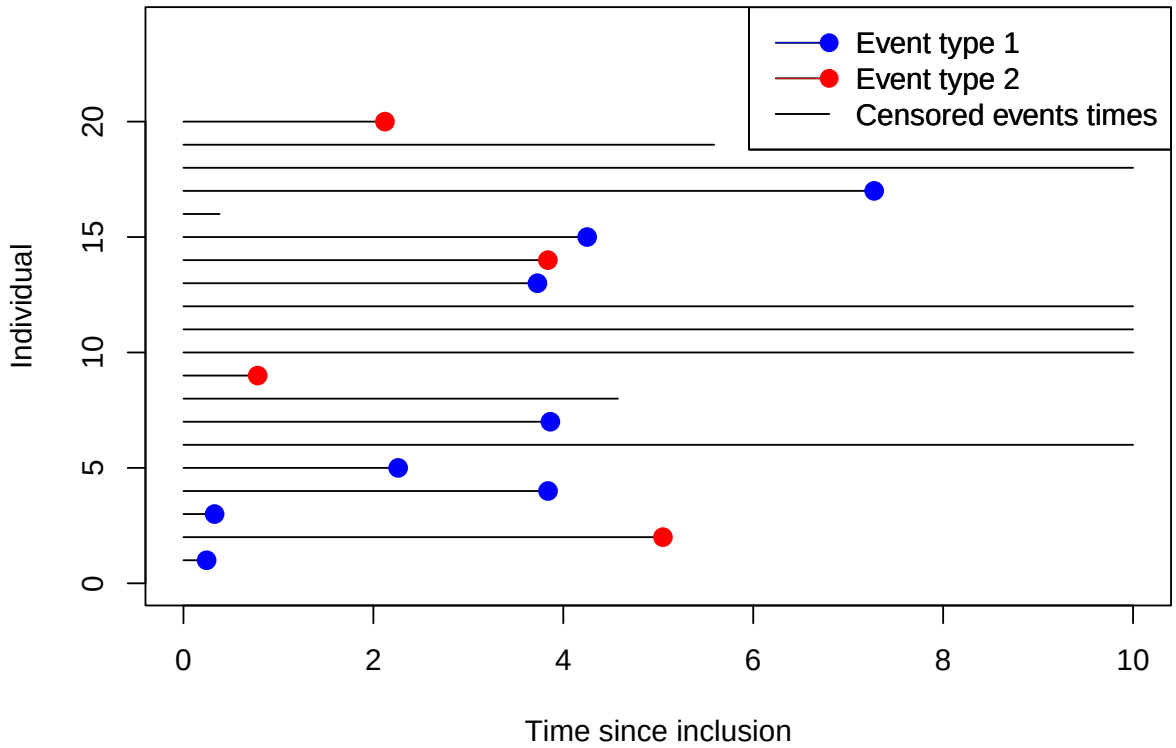
1 The problem of competing risks

- Let us consider possibly right censored event times:



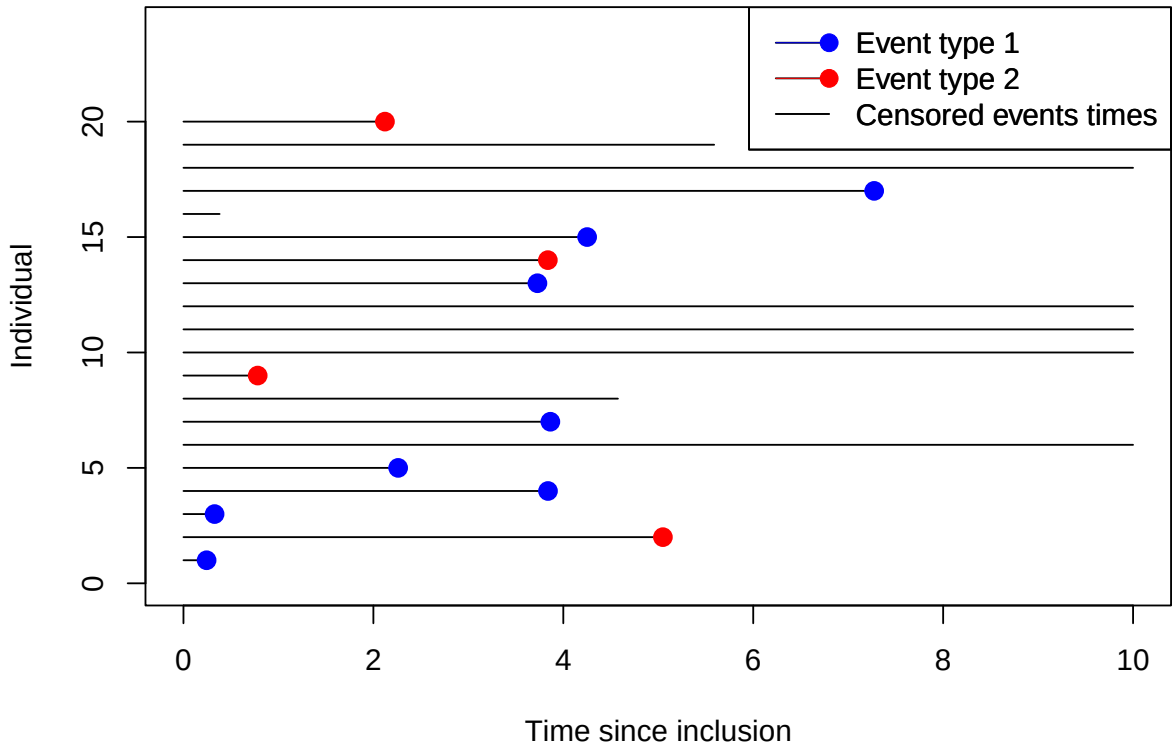
1 The problem of competing risks

- Let us consider possibly right censored event times:



1 The problem of competing risks

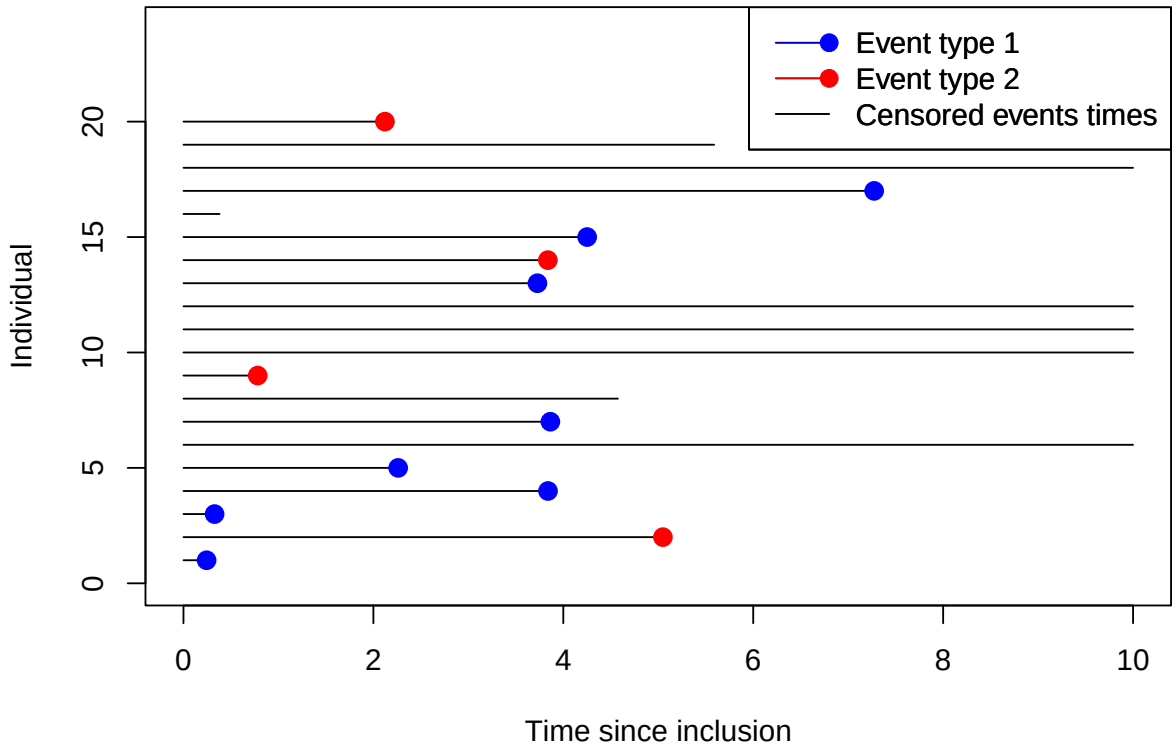
- Let us consider possibly right censored event times:



Interest might be in **events of only one type** (e.g. red)

1 The problem of competing risks

- Let us consider possibly right censored event times:



Interest might be in **events of only one type** (e.g. red)

For example; cancer specific death (with the competing event of death from other causes) or the positive event of return-to-work after traumatic injury (with the negative competing event of death)

- **By definition**: competing events make events of interest *impossible*

- **By definition**: competing events make events of interest *impossible*

  Different from censoring, which make events *unobservable*

- **By definition**: competing events make events of interest *impossible*

  Different from censoring, which make events *unobservable*
- If a competing risk problem is present **depends on**

- **By definition**: competing events make events of interest *impossible*

  Different from censoring, which make events *unobservable*
- If a competing risk problem is present **depends on**:
  1. choice of event of interest

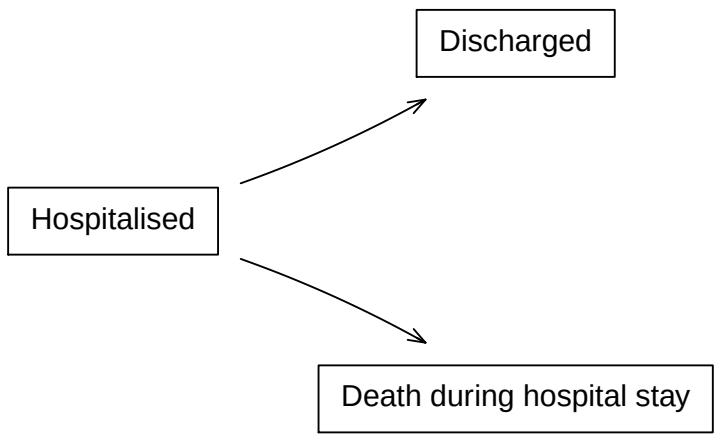
- **By definition**: competing events make events of interest *impossible*

  Different from censoring, which make events *unobservable*
- If a competing risk problem is present **depends on**:
  1. choice of event of interest
  2. the data being collected

- **By definition**: competing events make events of interest *impossible*

  Different from censoring, which make events *unobservable*
- If a competing risk problem is present **depends on**:
  1. choice of event of interest
  2. the data being collected

  Take for example:

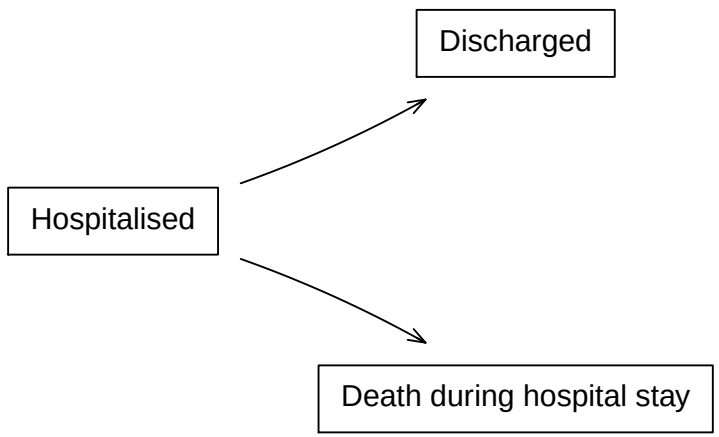


Data from hospital only

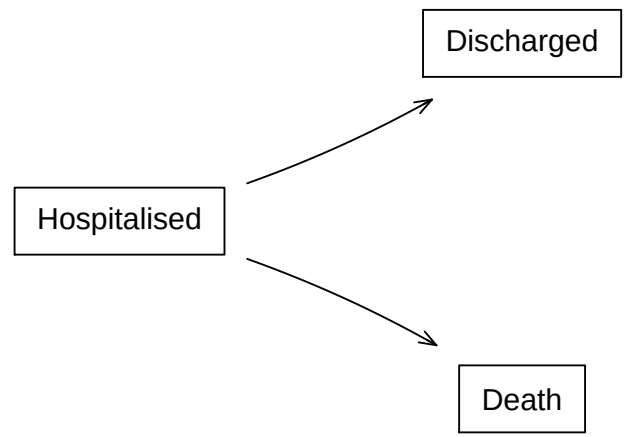
- **By definition**: competing events make events of interest *impossible*

  Different from censoring, which make events *unobservable*
- If a competing risk problem is present **depends on**:
  1. choice of event of interest
  2. the data being collected

  Take for example:



Data from hospital only

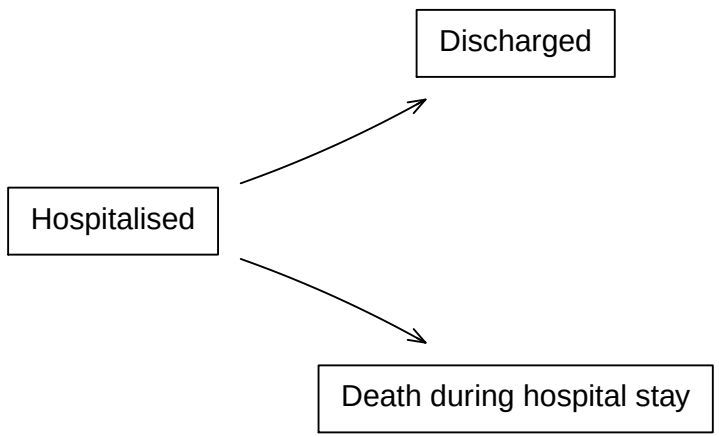


Data from hospital + national registries

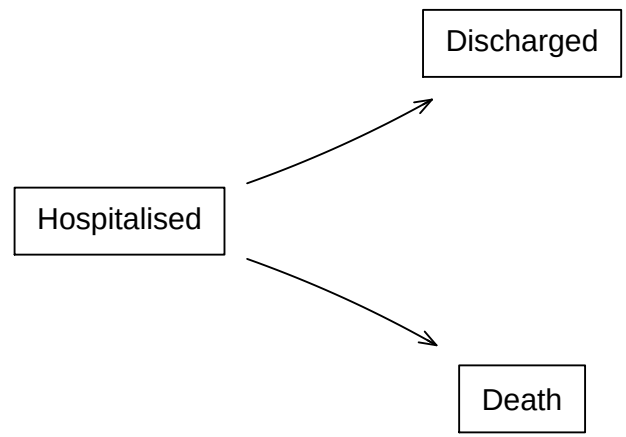
- **By definition**: competing events make events of interest *impossible*

  Different from censoring, which make events *unobservable*
- If a competing risk problem is present **depends on**:
  1. choice of event of interest
  2. the data being collected

  Take for example:



Data from hospital only



Data from hospital + national registries

- Treated by **Bernoulli (1766)**, Florence Nightingale (1860), Neyman (1950), Andersen et al. (2012) and many others

To censor or not censor

- From a **modelling perspective** it may appear that the question boils down to censoring or not censoring competing events

To censor or not censor

- From a **modelling perspective** it may appear that the question boils down to censoring or not censoring competing events
- Remember, the way censoring is dealt with in time-to-event analysis is that the **uncensored are made to represent the censored**

To censor or not censor

- From a **modelling perspective** it may appear that the question boils down to censoring or not censoring competing events
- Remember, the way censoring is dealt with in time-to-event analysis is that the **uncensored are made to represent the censored**

  This is possible with the independent censoring assumption (or conditional independent censoring and e.g. censor weighting)

To censor or not censor

- From a **modelling perspective** it may appear that the question boils down to censoring or not censoring competing events
- Remember, the way censoring is dealt with in time-to-event analysis is that the **uncensored are made to represent the censored**

  This is possible with the independent censoring assumption (or conditional independent censoring and e.g. censor weighting)

  Allow us to make valid inference for a world without censoring

To censor or not censor

- From a **modelling perspective** it may appear that the question boils down to censoring or not censoring competing events
- Remember, the way censoring is dealt with in time-to-event analysis is that the **uncensored are made to represent the censored**

  This is possible with the independent censoring assumption (or conditional independent censoring and e.g. censor weighting)

  Allow us to make valid inference for a world without censoring
- Censoring competing events indicate that we are interested in **a world without competing events**

To censor or not censor

- From a **modelling perspective** it may appear that the question boils down to censoring or not censoring competing events
- Remember, the way censoring is dealt with in time-to-event analysis is that the **uncensored are made to represent the censored**

  This is possible with the independent censoring assumption (or conditional independent censoring and e.g. censor weighting)

  Allow us to make valid inference for a world without censoring
- Censoring competing events indicate that we are interested in **a world without competing events**

  When is this world relevant?

To censor or not censor

- From a **modelling perspective** it may appear that the question boils down to censoring or not censoring competing events
- Remember, the way censoring is dealt with in time-to-event analysis is that the **uncensored are made to represent the censored**

  This is possible with the independent censoring assumption (or conditional independent censoring and e.g. censor weighting)

  Allow us to make valid inference for a world without censoring
- Censoring competing events indicate that we are interested in **a world without competing events**

  When is this world relevant?

  When are the independent censoring assumptions reasonable?

2 Classical survival methods for competing risks

Fundamental quantities

2 Classical survival methods for competing risks

Fundamental quantities

- The cause specific **cumulative incidence function** at time t,

$$F_j(t) = \mathbb{P}(T \leq t, Y = j),$$

for a given cause $j \in \{1, ..., J\}$

2 Classical survival methods for competing risks

Fundamental quantities

- The cause specific **cumulative incidence function** at time t,

$$F_j(t) = \mathbb{P}(T \leq t, Y = j),$$

for a given cause $j \in \{1, ..., J\}$

- Only meaningful to consider the **overall survival function**,

$$S(t) = \mathbb{P}(T > t) = 1 - \sum_{j=1}^{J} (F_j(t)),$$

which only can be constructed from *all* the cause specific cumulative incidence functions

- Can also consider the **cause specific hazard function** at time t,

$$h_j(t) = \lim_{\Delta t \to 0} \frac{1}{\Delta t} \mathbb{P}(t \leq T < t + \Delta t, Y = j \mid T \geq t),$$

which is the rate of (only) events by cause j, in small time intervals $t + \Delta t$, among those who have not yet died by any cause

- Can also consider the **cause specific hazard function** at time t,

$$h_j(t) = \lim_{\Delta t \to 0} \frac{1}{\Delta t} \mathbb{P}(t \leq T < t + \Delta t, Y = j \mid T \geq t),$$

which is the rate of (only) events by cause j, in small time intervals $t + \Delta t$, among those who have not yet died by any cause

Now have $S(t) = \exp\{-\left(\sum_{j=1}^{J} \int_0^t h_j(s)ds\right)\}$ and $F_j(t) = \int_0^t S(s)h_j(s)ds$

- Can also consider the **cause specific hazard function** at time t,

$$h_j(t) = \lim_{\Delta t \to 0} \frac{1}{\Delta t} \mathbb{P}(t \leq T < t + \Delta t, Y = j \mid T \geq t),$$

which is the rate of (only) events by cause j, in small time intervals $t + \Delta t$, among those who have not yet died by any cause

Now have $S(t) = \exp\{-\left(\sum_{j=1}^{J} \int_0^t h_j(s)ds\right)\}$ and $F_j(t) = \int_0^t S(s)h_j(s)ds$

- The **subdistribution hazard function** has also been suggested;

$$\tilde{h}_j(t) = \lim_{\Delta t \to 0} \frac{1}{\Delta t} \mathbb{P}(t \leq T_j < t + \Delta t, Y = j \mid T_j \geq t),$$

where $T_j = T \times I(Y = j) + \infty \times I(Y \neq j)$

Competing risk data

- Let us now consider the following **data structure**:

```
. list in 1/6

     +---------------------------------+
     | id   A          L          T   D |
     |---------------------------------|
  1. |  1   0   4.579293   12.72125   0 |
  2. |  2   0   3.884528   14.63008   2 |
  3. |  3   1   67.89856   12.08012   1 |
  4. |  4   0   27.68397   20.87179   2 |
  5. |  5   0   5.103034   15.05113   2 |
     |---------------------------------|
  6. |  6   0    46.6393   37.35997   1 |
     +---------------------------------+
```

Competing risk data

- Let us now consider the following **data structure**:

```
. list in 1/6

     +---------------------------------+
     | id   A          L          T   D |
     |---------------------------------|
  1. |  1   0   4.579293   12.72125   0 |
  2. |  2   0   3.884528   14.63008   2 |
  3. |  3   1   67.89856   12.08012   1 |
  4. |  4   0   27.68397   20.87179   2 |
  5. |  5   0   5.103034   15.05113   2 |
     |---------------------------------|
  6. |  6   0    46.6393   37.35997   1 |
     +---------------------------------+
```

where T are event times, D an event indicator (1 if event of type 1, and 2 if event of type 2) and $\boldsymbol{X} = \{A, \boldsymbol{L}\}$ baseline covariates

Competing risk data

- Let us now consider the following **data structure**:

```
. list in 1/6

     +---------------------------------+
     | id   A          L          T   D |
     |---------------------------------|
  1. |  1   0   4.579293   12.72125   0 |
  2. |  2   0   3.884528   14.63008   2 |
  3. |  3   1   67.89856   12.08012   1 |
  4. |  4   0   27.68397   20.87179   2 |
  5. |  5   0   5.103034   15.05113   2 |
     |---------------------------------|
  6. |  6   0    46.6393   37.35997   1 |
     +---------------------------------+
```

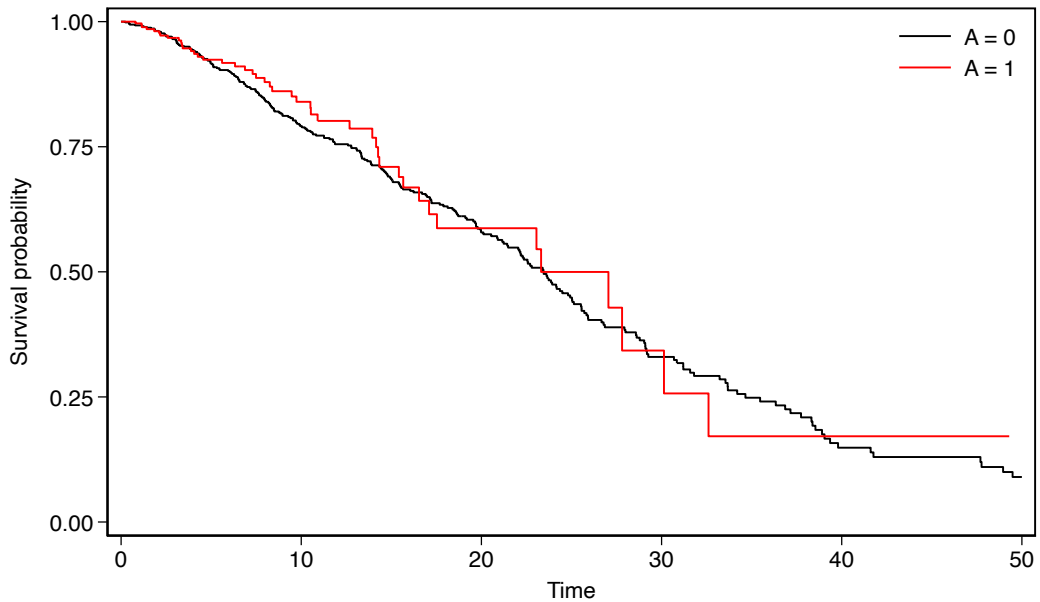
where T are event times, D an event indicator (1 if event of type 1, and 2 if event of type 2) and $\boldsymbol{X} = \{A, \boldsymbol{L}\}$ baseline covariates

Simulated data, where we can imagine that A denote two treatments given at time zero, and L a variable that affects both treatment choice and time to event of interest

The Kaplan-Meier estimator

- Easy to produce a **"cause specific" survival** curve:

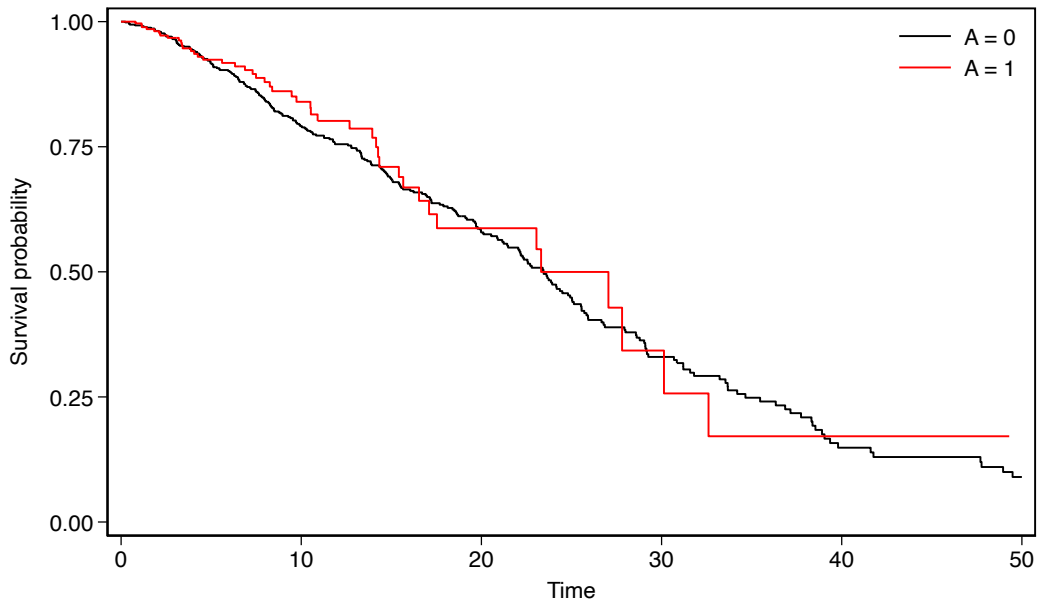
```
stset T, fail(D == 2) id(id)
sts graph, by(A) plot1opts(lcolor(black)) plot2opts(lcolor(red)) ytitle("Survival
probability") xtitle("Time") legend(ring(0) pos(1)) title("") ylab(, nogrid) xlab(, nogrid)
plotregion(lstyle(refline))
```



The Kaplan-Meier estimator

- Easy to produce a **"cause specific" survival** curve:

```
stset T, fail(D == 2) id(id)
sts graph, by(A) plot1opts(lcolor(black)) plot2opts(lcolor(red)) ytitle("Survival
probability") xtitle("Time") legend(ring(0) pos(1)) title("") ylab(, nogrid) xlab(, nogrid)
plotregion(lstyle(refline))
```

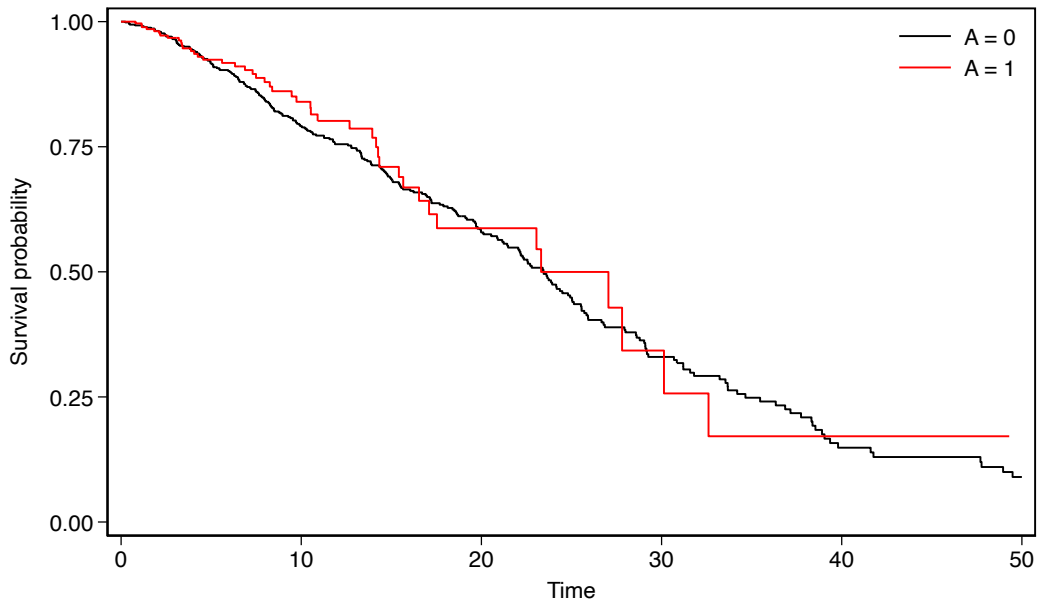


but interpretation is unclear, and this should generally be avoided

The Kaplan-Meier estimator

- Easy to produce a **"cause specific" survival** curve:

```
stset T, fail(D == 2) id(id)
sts graph, by(A) plot1opts(lcolor(black)) plot2opts(lcolor(red)) ytitle("Survival
probability") xtitle("Time") legend(ring(0) pos(1)) title("") ylab(, nogrid) xlab(, nogrid)
plotregion(lstyle(refline))
```



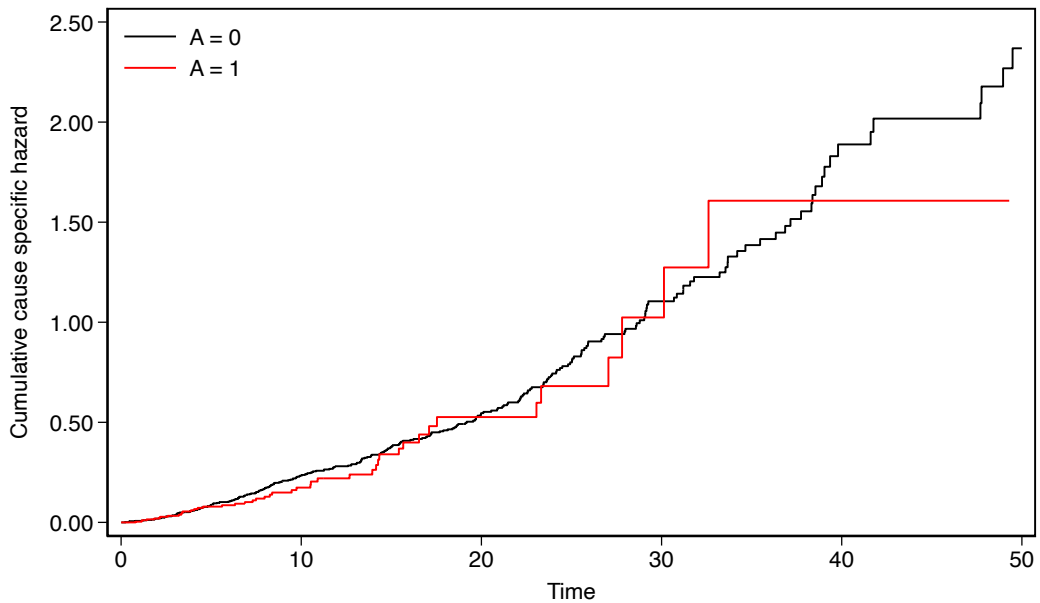
but interpretation is unclear, and this should generally be avoided

(data has 429 events (type 2), 500 competing events and 71 censored)

The Nelson-Aalen estimator

- Similarly, easy to estimate **cause specific cumulative hazard**:

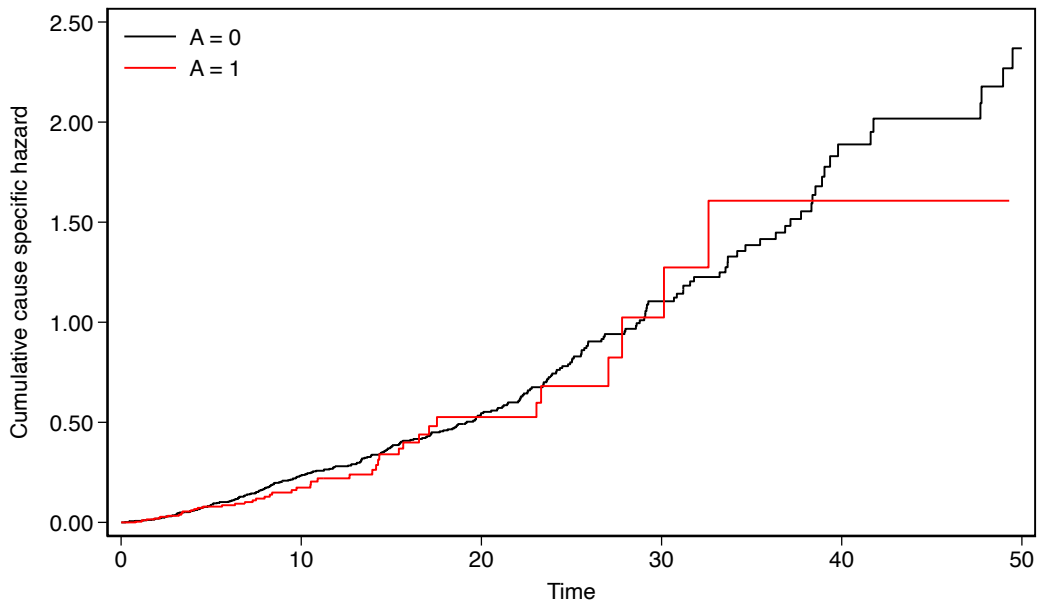
```
sts graph, by(A) na plot1opts(lcolor(black)) plot2opts(lcolor(red)) ytitle("Cumulative cause
specific hazard") xtitle("Time") legend(ring(0) pos(11)) title("") ylab(, nogrid)
xlab(, nogrid) plotregion(lstyle(refline))
```



The Nelson-Aalen estimator

- Similarly, easy to estimate **cause specific cumulative hazard**:

```
sts graph, by(A) na plot1opts(lcolor(black)) plot2opts(lcolor(red)) ytitle("Cumulative cause
specific hazard") xtitle("Time") legend(ring(0) pos(11)) title("") ylab(, nogrid)
xlab(, nogrid) plotregion(lstyle(refline))
```

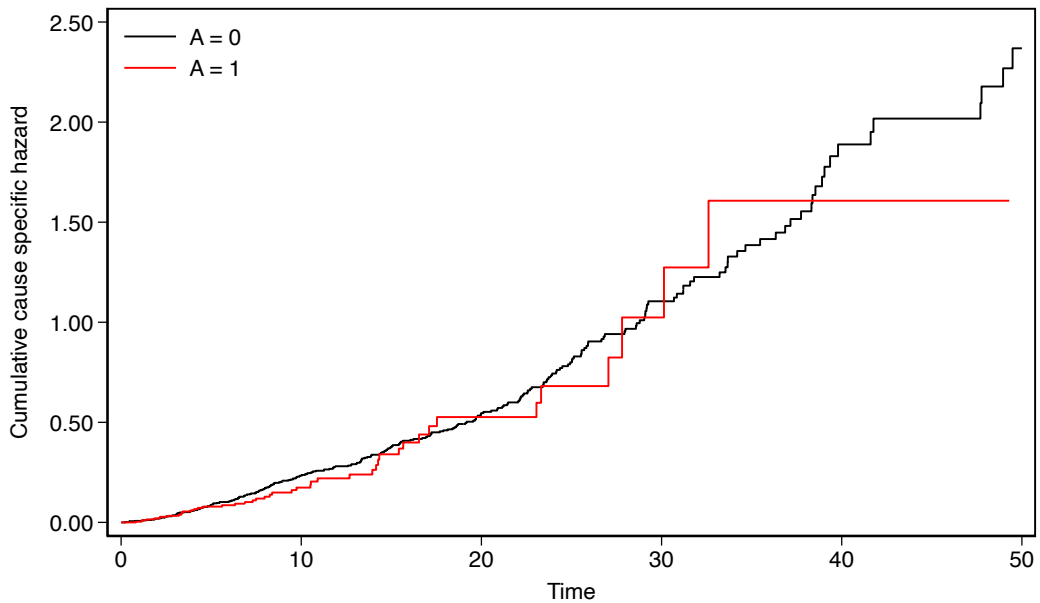


These curves *can* be interpreted as describing the movement from the "alive" state to the "event 2" state

The Nelson-Aalen estimator

- Similarly, easy to estimate **cause specific cumulative hazard**:

```
sts graph, by(A) na plot1opts(lcolor(black)) plot2opts(lcolor(red)) ytitle("Cumulative cause
specific hazard") xtitle("Time") legend(ring(0) pos(11)) title("") ylab(, nogrid)
xlab(, nogrid) plotregion(lstyle(refline))
```



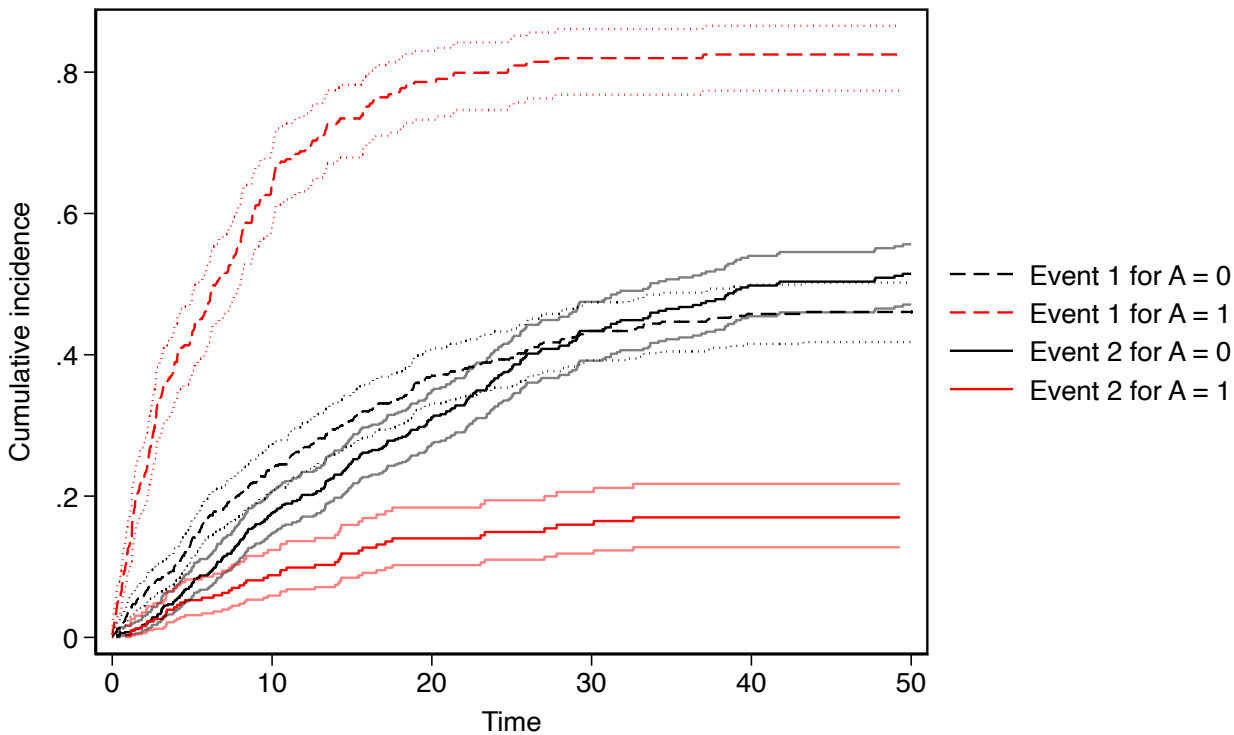
These curves *can* be interpreted as describing the movement from the "alive" state to the "event 2" state

But, the shape of these (cumulative) hazards is now a result of i) individual risk, ii) selection, and iii) rate of competing events

The Aalen-Johansen plug-in estimator

- Can also calculate **cause specific cumulative incidence**, where competing events are allowed:

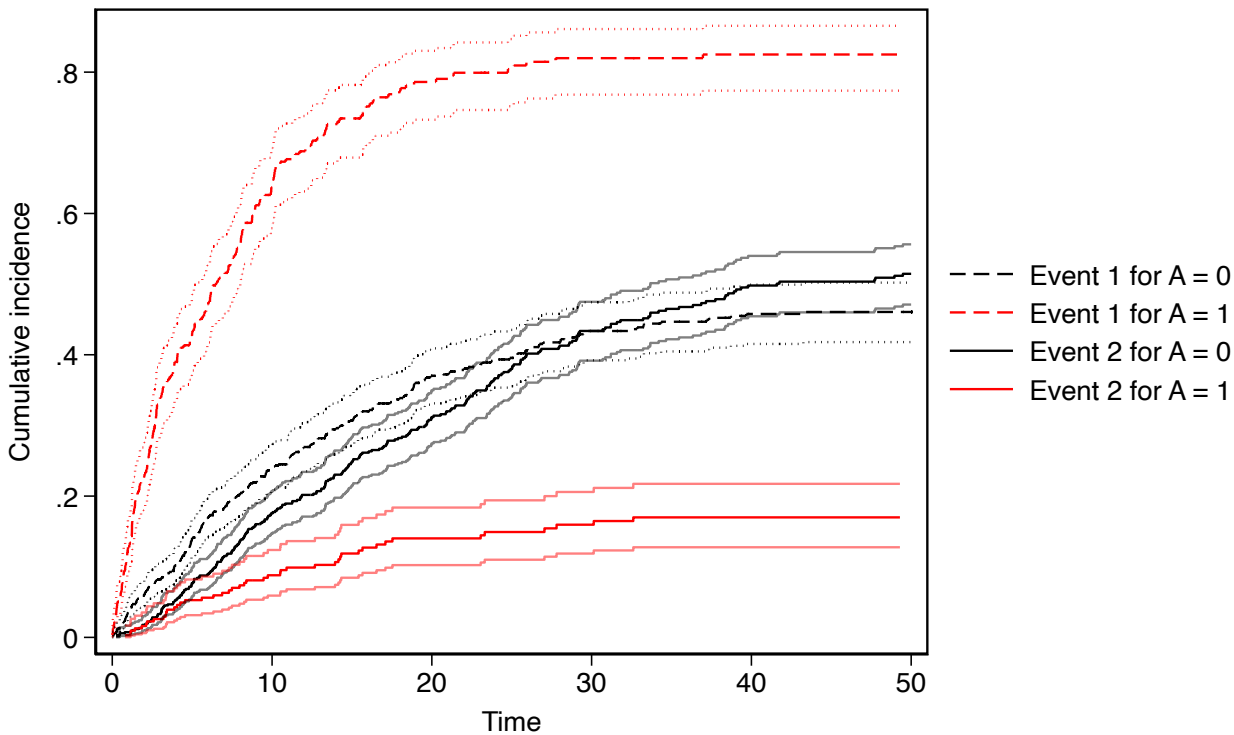
```
stcompet cuminc = ci ub = hi lb = lo, compet1(1) by(A)
```



The Aalen-Johansen plug-in estimator

- Can also calculate **cause specific cumulative incidence**, where competing events are allowed:

```
stcompet cuminc = ci ub = hi lb = lo, compet1(1) by(A)
```



This give a very different picture than cumulative incidence from Kaplan-Meier (1-KM two slides back), which always will overestimate the incidence in the presence of competing events

- Note that, under competing risk, the cumulative incidence is

$$F_j(t) = \int_0^t S(t) h_j(t),$$

where $S(t)$ is the overall survival $S(t) = \exp\left(-\sum_k H_k(t)\right)$ for all k events and $h_j(t)$ is the cause specific hazard for event j

- Note that, under competing risk, the cumulative incidence is

$$F_j(t) = \int_0^t S(t)h_j(t),$$

where $S(t)$ is the overall survival $S(t) = \exp\left(-\sum_k H_k(t)\right)$ for all k events and $h_j(t)$ is the cause specific hazard for event j

So, there is **no longer a one-to-one correspondence** between (cause specific) hazard and cumulative incidence

- Note that, under competing risk, the cumulative incidence is

$$F_j(t) = \int_0^t S(t)h_j(t),$$

where $S(t)$ is the overall survival $S(t) = \exp\left(-\sum_k H_k(t)\right)$ for all k events and $h_j(t)$ is the cause specific hazard for event j

So, there is **no longer a one-to-one correspondence** between (cause specific) hazard and cumulative incidence

- A general **"empirical transition matrix"** estimator given by Aalen and Johansen (1978):

$$\hat{\mathbf{P}}(0,t) = \pi_0 \prod_{u \in (0,t]} (\mathbf{I} + d\widehat{\mathbf{A}}(u)),$$

where π_0 is the initial state distribution vector (for us $\{1,0,0\}$) and $\widehat{\mathbf{A}}(u))$ is the cumulative transition hazard matrix

Cause specific hazard models

- Common to also fit **Cox models for cause specific hazards**, e.g. using `stcox A`:

```
...
Cox regression with no ties

No. of subjects =      1,000                          Number of obs =  1,000
No. of failures =        315
Time at risk    = 11,185.729
                                                      LR chi2(1)    =   0.49
Log likelihood = -1759.2309                           Prob > chi2   = 0.4848

------------------------------------------------------------------------------
          _t | Haz. ratio   Std. err.      z    P>|z|     [95% conf. interval]
-------------+----------------------------------------------------------------
           A |   .8928204   .1466702    -0.69   0.490     .6470397    1.231962
------------------------------------------------------------------------------
...
```

Cause specific hazard models

- Common to also fit **Cox models for cause specific hazards**, e.g. using `stcox A`:

```
...
Cox regression with no ties

No. of subjects =      1,000                          Number of obs =  1,000
No. of failures =        315
Time at risk    = 11,185.729
                                                      LR chi2(1)    =   0.49
Log likelihood = -1759.2309                           Prob > chi2   = 0.4848

------------------------------------------------------------------------------
          _t | Haz. ratio   Std. err.      z    P>|z|     [95% conf. interval]
-------------+----------------------------------------------------------------
           A |   .8928204   .1466702    -0.69   0.490     .6470397    1.231962
------------------------------------------------------------------------------
...
```

Tempting because it is easy, but this HR can be hard to motivate

Cause specific hazard models

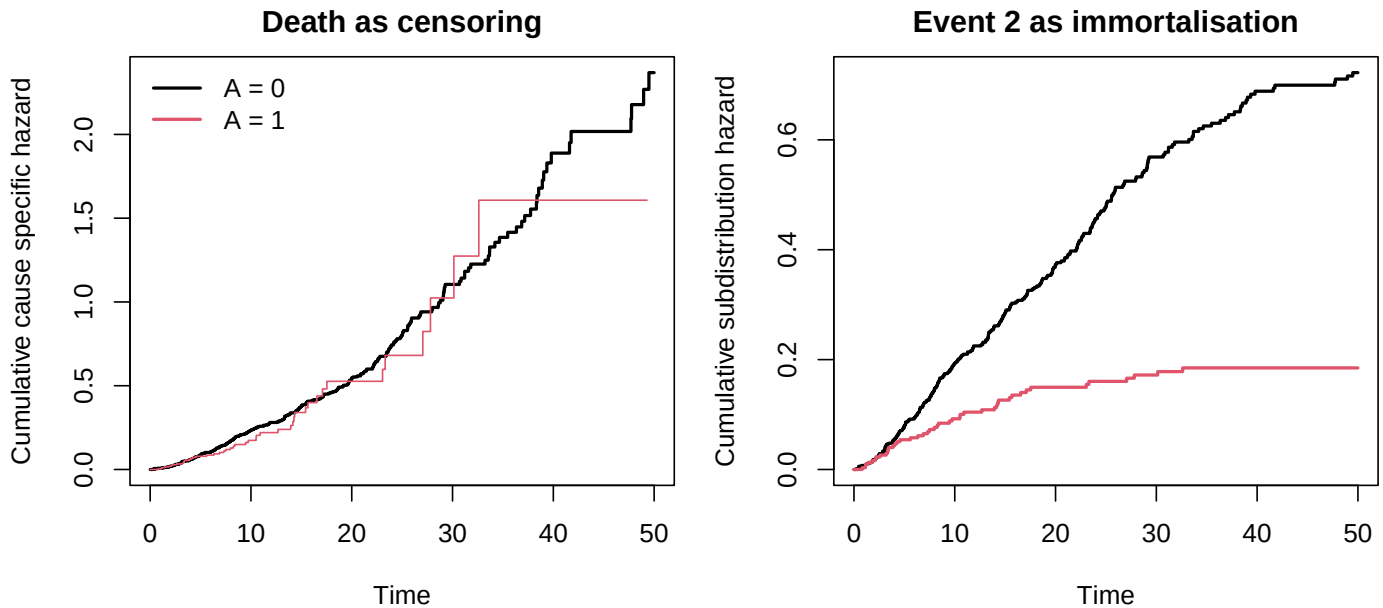
- Common to also fit **Cox models for cause specific hazards**, e.g. using `stcox A`:

```
...
Cox regression with no ties

No. of subjects =      1,000                          Number of obs =  1,000
No. of failures =        315
Time at risk    = 11,185.729
                                                      LR chi2(1)    =   0.49
Log likelihood = -1759.2309                           Prob > chi2   = 0.4848

------------------------------------------------------------------------------
          _t | Haz. ratio   Std. err.      z    P>|z|     [95% conf. interval]
-------------+----------------------------------------------------------------
           A |   .8928204   .1466702    -0.69   0.490     .6470397    1.231962
------------------------------------------------------------------------------
...
```

Tempting because it is easy, but this HR can be hard to motivate (useful component for calculating cumulative incidence though)

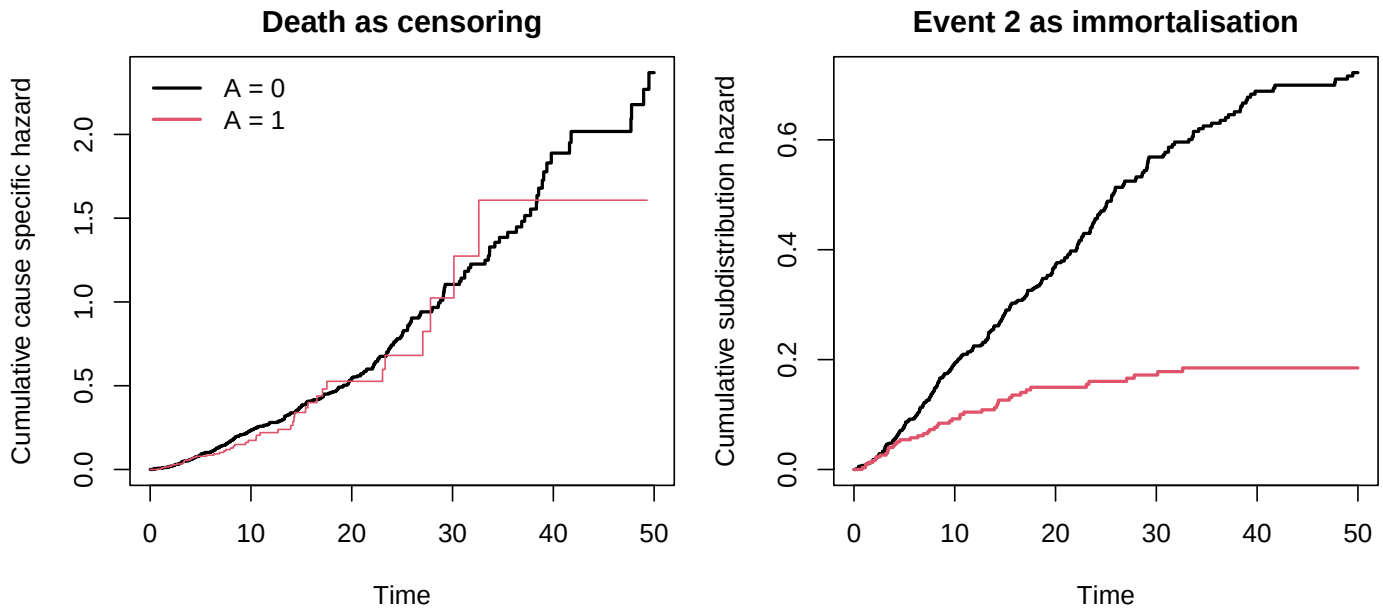
The Fine and Gray model

- Fine et al. (1999) showed that the one-to-one correspondence can be restored by estimating **the subdistribution hazard**, identified (in a censor free world) by setting all competing event times to ∞



The Fine and Gray model

- Fine et al. (1999) showed that the one-to-one correspondence can be restored by estimating **the subdistribution hazard**, identified (in a censor free world) by setting all competing event times to ∞



Interpretation is awkward, but this makes 1-KM in the subdistribution data equal to the real world $F_j(t)$

- The **subdistribution dataset** (and censoring weights) can be created manually

- The **subdistribution dataset** (and censoring weights) can be created manually
- See also `stcrreg A compete(D == 2)`, fitting **the Fine and Gray model**, corresponding to a (weighted) Cox model on the subdistribution dataset:

```
...

Competing-risks regression                          No. of obs      =      1,000
                                                    No. of subjects =      1,000
Failure event:   D == 2                             No. failed      =        315
Competing event: D == 1                             No. competing   =        500
                                                    No. censored    =        185


                                                    Wald chi2(1)    =      51.90
Log pseudolikelihood = -1999.7238                   Prob > chi2     =     0.0000

                                (Std. err. adjusted for 1,000 clusters in id)
------------------------------------------------------------------------------
             |               Robust
          _t |        SHR   std. err.      z    P>|z|     [95% conf. interval]
-------------+----------------------------------------------------------------
           A |   .3040094   .0502443    -7.20   0.000     .2198908    .4203072
------------------------------------------------------------------------------
...
```

- The **subdistribution dataset** (and censoring weights) can be created manually
- See also `stcrreg A compete(D == 2)`, fitting **the Fine and Gray model**, corresponding to a (weighted) Cox model on the subdistribution dataset:

```
...

Competing-risks regression                          No. of obs      =      1,000
                                                    No. of subjects =      1,000
Failure event:   D == 2                             No. failed      =        315
Competing event: D == 1                             No. competing   =        500
                                                    No. censored    =        185


                                                    Wald chi2(1)    =      51.90
Log pseudolikelihood = -1999.7238                   Prob > chi2     =     0.0000

                                (Std. err. adjusted for 1,000 clusters in id)
------------------------------------------------------------------------------
             |               Robust
          _t |        SHR   std. err.      z    P>|z|     [95% conf. interval]
-------------+----------------------------------------------------------------
           A |   .3040094   .0502443    -7.20   0.000     .2198908    .4203072
------------------------------------------------------------------------------
...
```

The coefficient is not easy to interpret, but the model *can* be used as a test for difference in cumulative incidence, equivalent to the log-rank test in settings without competing events

- The **subdistribution dataset** (and censoring weights) can be created manually
- See also `stcrreg A compete(D == 2)`, fitting **the Fine and Gray model**, corresponding to a (weighted) Cox model on the subdistribution dataset:

```
...

Competing-risks regression                          No. of obs      =      1,000
                                                    No. of subjects =      1,000
Failure event:   D == 2                             No. failed      =        315
Competing event: D == 1                             No. competing   =        500
                                                    No. censored    =        185


                                                    Wald chi2(1)    =      51.90
Log pseudolikelihood = -1999.7238                   Prob > chi2     =     0.0000

                                (Std. err. adjusted for 1,000 clusters in id)
------------------------------------------------------------------------------
             |               Robust
          _t |        SHR   std. err.      z    P>|z|     [95% conf. interval]
-------------+----------------------------------------------------------------
           A |   .3040094   .0502443    -7.20   0.000     .2198908    .4203072
------------------------------------------------------------------------------
...
```

The coefficient is not easy to interpret, but the model *can* be used as a test for difference in cumulative incidence, equivalent to the log-rank test in settings without competing events[1]

[1] For regression of $F_j(t)$, see also pseudo values (Klein et al. 2005)

An example (out of many)

thebmj covid-19 Research Education News & Views Campaigns Jobs

Research

Covid-19 vaccine effectiveness against post-covid-19 condition among 589 722 individuals in Sweden: population based cohort study

BMJ 2023 ; 383 doi: https://doi.org/10.1136/bmj-2023-076990 (Published 22 November 2023)
Cite this as: *BMJ* 2023;383:e076990

Linked Editorial
Does timely vaccination help prevent post-viral conditions?

An example (out of many)

- Consider **time to post covid condition** (PCC) after covid-19

An example (out of many)

- Consider **time to post covid condition** (PCC) after covid-19

  Compare vaccinated and unvaccinated (from time of infection)

An example (out of many)

thebmj covid-19 Research Education News & Views Campaigns Jobs

Research

Covid-19 vaccine effectiveness against post-covid-19 condition among 589 722 individuals in Sweden: population based cohort study

BMJ 2023 ; 383 doi: https://doi.org/10.1136/bmj-2023-076990 (Published 22 November 2023)
Cite this as: *BMJ* 2023;383:e076990

Linked Editorial
Does timely vaccination help prevent post-viral conditions?

- Consider **time to post covid condition** (PCC) after covid-19

  Compare vaccinated and unvaccinated (from time of infection)

  After covid infection people can get PCC, but are censored if they get vaccinated, reinfected, emigrate or die

- **Does it matter?**

- **Does it matter?**

	Not vaccinated before covid-19 n = 290,030	Vaccinated before covid-19 n = 299,692
Vaccination, n (%)	200,965 (69)	167,000 (56)
Reached end of follow-up[1], n (%)	73,872 (26)	126,835 (42)
Reinfection[2], n (%)	9,613 (3.3)	3,275 (1.1)
PCC, n (%)	4,118 (1.4)	1,201 (0.4)
Death, n (%)	821 (0.3)	1,076 (0.4)
Emigration, n (%)	641 (0.2)	305 (0.1)

[1]30 November 2022.
[2]New covid-19 infection at least 90 days after covid-19 index date.

PCC=post-covid-19 condition

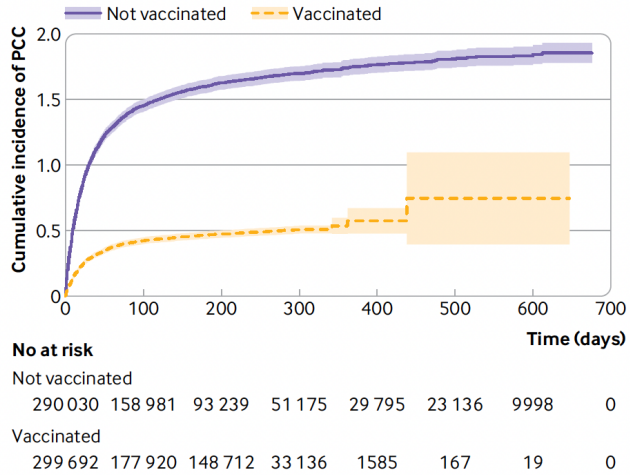


Fig 2 | Cumulative incidence of PCC, using Kaplan-Meier failure function, for individuals vaccinated or not vaccinated against covid-19. Study population included all adult (≥18 years) residents in the two largest regions of Sweden with covid-19 first registered during the study inclusion period, 27 December 2020 to 9 February 2022. PCC=post-covid-19 condition

- **Does it matter?**

	Not vaccinated before covid-19 n = 290,030	Vaccinated before covid-19 n = 299,692
Vaccination, n (%)	200,965 (69)	167,000 (56)
Reached end of follow-up[1], n (%)	73,872 (26)	126,835 (42)
Reinfection[2], n (%)	9,613 (3.3)	3,275 (1.1)
PCC, n (%)	4,118 (1.4)	1,201 (0.4)
Death, n (%)	821 (0.3)	1,076 (0.4)
Emigration, n (%)	641 (0.2)	305 (0.1)

[1]30 November 2022.
[2]New covid-19 infection at least 90 days after covid-19 index date.

PCC=post-covid-19 condition

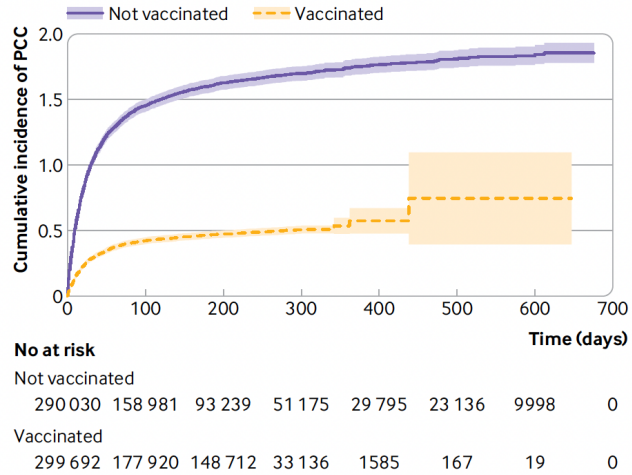


Fig 2 | Cumulative incidence of PCC, using Kaplan-Meier failure function, for individuals vaccinated or not vaccinated against covid-19. Study population included all adult (≥18 years) residents in the two largest regions of Sweden with covid-19 first registered during the study inclusion period, 27 December 2020 to 9 February 2022. PCC=post-covid-19 condition

Estimated cumulative incidence not representative of real world incidence

- **Does it matter?**

	Not vaccinated before covid-19 n = 290,030	Vaccinated before covid-19 n = 299,692
Vaccination, n (%)	200,965 (69)	167,000 (56)
Reached end of follow-up[1], n (%)	73,872 (26)	126,835 (42)
Reinfection[2], n (%)	9,613 (3.3)	3,275 (1.1)
PCC, n (%)	4,118 (1.4)	1,201 (0.4)
Death, n (%)	821 (0.3)	1,076 (0.4)
Emigration, n (%)	641 (0.2)	305 (0.1)

[1]30 November 2022.
[2]New covid-19 infection at least 90 days after covid-19 index date.

PCC=post-covid-19 condition

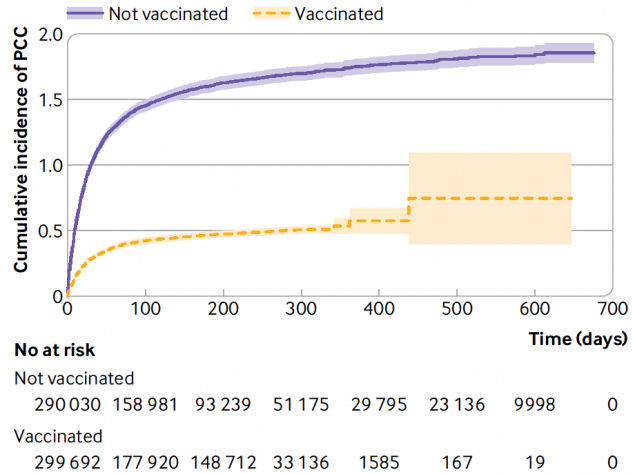


Fig 2 | Cumulative incidence of PCC, using Kaplan-Meier failure function, for individuals vaccinated or not vaccinated against covid-19. Study population included all adult (≥18 years) residents in the two largest regions of Sweden with covid-19 first registered during the study inclusion period, 27 December 2020 to 9 February 2022. PCC=post-covid-19 condition

Estimated cumulative incidence not representative of real world incidence

Hard to tell how it affect contrasts between groups (curves) and for vaccine effectiveness calculated using 1-HR

- **Does it matter?**

	Not vaccinated before covid-19 n = 290,030	Vaccinated before covid-19 n = 299,692
Vaccination, n (%)	200,965 (69)	167,000 (56)
Reached end of follow-up[1], n (%)	73,872 (26)	126,835 (42)
Reinfection[2], n (%)	9,613 (3.3)	3,275 (1.1)
PCC, n (%)	4,118 (1.4)	1,201 (0.4)
Death, n (%)	821 (0.3)	1,076 (0.4)
Emigration, n (%)	641 (0.2)	305 (0.1)

[1]30 November 2022.
[2]New covid-19 infection at least 90 days after covid-19 index date.

PCC=post-covid-19 condition

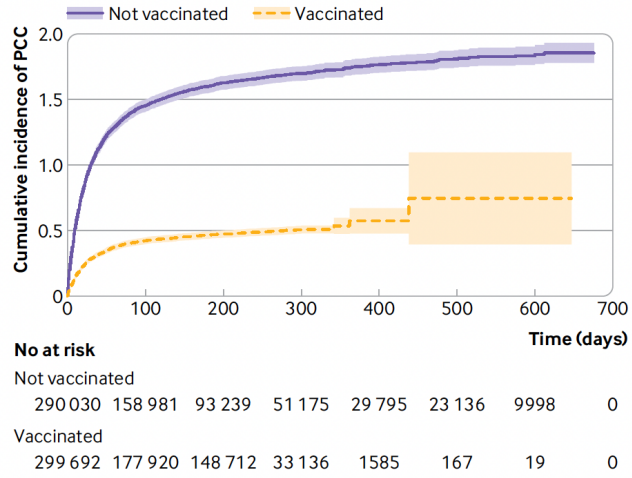


Fig 2 | Cumulative incidence of PCC, using Kaplan-Meier failure function, for individuals vaccinated or not vaccinated against covid-19. Study population included all adult (≥18 years) residents in the two largest regions of Sweden with covid-19 first registered during the study inclusion period, 27 December 2020 to 9 February 2022. PCC=post-covid-19 condition

Estimated cumulative incidence not representative of real world incidence

Hard to tell how it affect contrasts between groups (curves) and for vaccine effectiveness calculated using 1-HR

(paper generated a lot of discussion, a BMJ editorial, 8 online comments and a correction, but nothing on competing risks)

3 Causal estimands under competing risk

Recent developments in the analysis of competing risks

- **Young et al. (2020)** put classical statistical estimands for competing risk in a causal frame, drawing parallels to mediation analysis and defining total and direct effects

3 Causal estimands under competing risk

Recent developments in the analysis of competing risks

- **Young et al. (2020)** put classical statistical estimands for competing risk in a causal frame, drawing parallels to mediation analysis and defining total and direct effects

  Followed up by Stensrud et al. (2022) on separable effects and Janvin et al. (2024) on recurrent and competing events

3 Causal estimands under competing risk

Recent developments in the analysis of competing risks

- **Young et al. (2020)** put classical statistical estimands for competing risk in a causal frame, drawing parallels to mediation analysis and defining total and direct effects

  Followed up by Stensrud et al. (2022) on separable effects and Janvin et al. (2024) on recurrent and competing events

- **The ICH E9 R1 addendum (2019)** on estimands and sensitivity analysis in clinical trials open for more causal inference formalism

3 Causal estimands under competing risk

Recent developments in the analysis of competing risks

- **Young et al. (2020)** put classical statistical estimands for competing risk in a causal frame, drawing parallels to mediation analysis and defining total and direct effects

  Followed up by Stensrud et al. (2022) on separable effects and Janvin et al. (2024) on recurrent and competing events

- **The ICH E9 R1 addendum (2019)** on estimands and sensitivity analysis in clinical trials open for more causal inference formalism

  Give strategies on how to handle "intercurrent events": *events occurring after treatment initiation affecting interpretation or existence of the measurements associated with the clinical question of interest*

Start by justifying an estimand

Start by justifying an estimand

- Causal inference formalism boils down to **three distinct steps**

Start by justifying an estimand

- Causal inference formalism boils down to **three distinct steps**:
  1. Estimand: Motivate and describe a well-defined causal contrast

Start by justifying an estimand

- Causal inference formalism boils down to **three distinct steps**:
  1. Estimand: Motivate and describe a well-defined causal contrast
  2. Identification: Lay out assumptions needed to identify it

Start by justifying an estimand

- Causal inference formalism boils down to **three distinct steps**:
  1. Estimand: Motivate and describe a well-defined causal contrast
  2. Identification: Lay out assumptions needed to identify it
  3. Estimation: Chose a statistical estimator

Start by justifying an estimand

- Causal inference formalism boils down to **three distinct steps**:
  1. Estimand: Motivate and describe a well-defined causal contrast
  2. Identification: Lay out assumptions needed to identify it
  3. Estimation: Chose a statistical estimator
- Encouraged through **the causal roadmap** (Petersen et al. 2014) and target trial emulation (Hernán and Robins 2016)

Start by justifying an estimand

- Causal inference formalism boils down to **three distinct steps**:
  1. Estimand: Motivate and describe a well-defined causal contrast
  2. Identification: Lay out assumptions needed to identify it
  3. Estimation: Chose a statistical estimator
- Encouraged through **the causal roadmap** (Petersen et al. 2014) and target trial emulation (Hernán and Robins 2016)

  ... no reason to start by asking whether to censor or not

Average causal effects

- Let us consider **general average total effects** (ATEs) on form

$$\text{ATE} = \mathbb{E}(Y^1) \text{ vs. } \mathbb{E}(Y^0),$$

for outcome Y under interventions 1 and 0

Average causal effects

- Let us consider **general average total effects** (ATEs) on form

$$\text{ATE} = \mathbb{E}(Y^1) \text{ vs. } \mathbb{E}(Y^0),$$

for outcome Y under interventions 1 and 0

Note: this is a marginal effect (as opposed to conditional), and that interest may be in both adjusted or unadjusted marginal effects

Average causal effects

- Let us consider **general average total effects** (ATEs) on form

$$\text{ATE} = \mathbb{E}(Y^1) \text{ vs. } \mathbb{E}(Y^0),$$

  for outcome Y under interventions 1 and 0

  Note: this is a marginal effect (as opposed to conditional), and that interest may be in both adjusted or unadjusted marginal effects

- In a time-to-event setting we have **various options**, where hazard ratios (or other contrasts of hazards) are not the best ones (Hernán 2010; Aalen, Cook et al. 2015)

Average causal effects

- Let us consider **general average total effects** (ATEs) on form

$$\text{ATE} = \mathbb{E}(Y^1) \text{ vs. } \mathbb{E}(Y^0),$$

  for outcome Y under interventions 1 and 0

  Note: this is a marginal effect (as opposed to conditional), and that interest may be in both adjusted or unadjusted marginal effects

- In a time-to-event setting we have **various options**, where hazard ratios (or other contrasts of hazards) are not the best ones (Hernán 2010; Aalen, Cook et al. 2015)

- Under competing risks **contrasts of cumulative incidence** are a natural choice, e.g. ATEs as a contrast of

$$\mathbb{P}(T^1 \leq t, J = j) \text{ vs. } \mathbb{P}(T^0 \leq t, J = j),$$

  where J denote the event type, taking values 1, 2, ...

Average causal effects

- Let us consider **general average total effects** (ATEs) on form

$$\text{ATE} = \mathbb{E}(Y^1) \text{ vs. } \mathbb{E}(Y^0),$$

  for outcome Y under interventions 1 and 0

  Note: this is a marginal effect (as opposed to conditional), and that interest may be in both adjusted or unadjusted marginal effects

- In a time-to-event setting we have **various options**, where hazard ratios (or other contrasts of hazards) are not the best ones (Hernán 2010; Aalen, Cook et al. 2015)

- Under competing risks **contrasts of cumulative incidence** are a natural choice, e.g. ATEs as a contrast of

$$\mathbb{P}(T^1 \leq t, J = j) \text{ vs. } \mathbb{P}(T^0 \leq t, J = j),$$

  where J denote the event type, taking values 1, 2, ...

  (contrasts of restricted mean time lost is a related alternative)

Inverse probability weighting

- For baseline differences between groups, adjusted marginal effects can be estimated with **propensity score weights**

Inverse probability weighting

- For baseline differences between groups, adjusted marginal effects can be estimated with **propensity score weights**:

$$W_{PS} = \frac{1}{\mathbb{P}(A \mid L)}$$

Inverse probability weighting

- For baseline differences between groups, adjusted marginal effects can be estimated with **propensity score weights**:

$$W_{PS} = \frac{1}{\mathbb{P}(A \mid L)}$$

In Stata:

```
mkspline ans=L, cubic nknots(5)
logistic A ans1 ans2 ans3 ans4
predict double phat, pr
gen psw=1/phat if A==1
replace psw=1/(1-phat) if A==0
stset T [pweight=psw], failure(D==2)
```

Inverse probability weighting

- For baseline differences between groups, adjusted marginal effects can be estimated with **propensity score weights**:

$$W_{PS} = \frac{1}{\mathbb{P}(A \mid L)}$$

In Stata:

```
mkspline ans=L, cubic nknots(5)
logistic A ans1 ans2 ans3 ans4
predict double phat, pr
gen psw=1/phat if A==1
replace psw=1/(1-phat) if A==0
stset T [pweight=psw], failure(D==2)
```

Standard methods apply to weighted data, but SE must take into account the uncertainty in weights (e.g. using bootstrap)

Inverse probability weighting

- For baseline differences between groups, adjusted marginal effects can be estimated with **propensity score weights**:

$$W_{PS} = \frac{1}{\mathbb{P}(A \mid L)}$$

In Stata:

```
mkspline ans=L, cubic nknots(5)
logistic A ans1 ans2 ans3 ans4
predict double phat, pr
gen psw=1/phat if A==1
replace psw=1/(1-phat) if A==0
stset T [pweight=psw], failure(D==2)
```

Standard methods apply to weighted data, but SE must take into account the uncertainty in weights (e.g. using bootstrap)

- `stcrreg A, compete(D==1)` give **a weighted test** for difference in cause specific cumulative incidence

- For **time-varying treatment regimes**, more general inverse probability of treatment weighting (IPTW) can be used for

$$\mathbb{P}(T^{\bar{a}} \leq t, J = j) \text{ vs. } \mathbb{P}(T^{\bar{a}'} \leq t, J = j)$$

- For **time-varying treatment regimes**, more general inverse probability of treatment weighting (IPTW) can be used for

$$\mathbb{P}(T^{\bar{a}} \leq t, J = j) \text{ vs. } \mathbb{P}(T^{\bar{a}'} \leq t, J = j)$$

- **Inverse probability of censoring weighting** (IPCW) for to account for (regular or artificial) dependent censoring, using time-dependent weights:

$$W_C(t) = \prod_{k=1}^{t} \frac{\mathbb{P}(\text{not censored at } k | \text{baseline covariates})}{\mathbb{P}(\text{not censored at } k | \text{covariates up to } k)}$$

- For **time-varying treatment regimes**, more general inverse probability of treatment weighting (IPTW) can be used for

$$\mathbb{P}(T^{\bar{a}} \leq t, J = j) \text{ vs. } \mathbb{P}(T^{\bar{a}'} \leq t, J = j)$$

- **Inverse probability of censoring weighting** (IPCW) for to account for (regular or artificial) dependent censoring, using time-dependent weights:

$$W_C(t) = \prod_{k=1}^{t} \frac{\mathbb{P}(\text{not censored at } k | \text{baseline covariates})}{\mathbb{P}(\text{not censored at } k | \text{covariates up to } k)}$$

Censoring can be seen as a time-dependent treatment:

$$\text{ATE} = \mathbb{P}(Y_k^{a=1,\bar{c}_k=\bar{0}} = 1) \text{ vs. } \mathbb{P}(Y_k^{a=0,\bar{c}_k=\bar{0}} = 1)$$

- For **time-varying treatment regimes**, more general inverse probability of treatment weighting (IPTW) can be used for

$$\mathbb{P}(T^{\bar{a}} \leq t, J = j) \text{ vs. } \mathbb{P}(T^{\bar{a}'} \leq t, J = j)$$

- **Inverse probability of censoring weighting** (IPCW) for to account for (regular or artificial) dependent censoring, using time-dependent weights:

$$W_C(t) = \prod_{k=1}^{t} \frac{\mathbb{P}(\text{not censored at } k | \text{baseline covariates})}{\mathbb{P}(\text{not censored at } k | \text{covariates up to } k)}$$

Censoring can be seen as a time-dependent treatment:

$$\text{ATE} = \mathbb{P}(Y_k^{a=1,\bar{c}_k=\bar{0}} = 1) \text{ vs. } \mathbb{P}(Y_k^{a=0,\bar{c}_k=\bar{0}} = 1)$$

See `stteffects ipwra`

- For **time-varying treatment regimes**, more general inverse probability of treatment weighting (IPTW) can be used for

$$\mathbb{P}(T^{\bar{a}} \leq t, J = j) \text{ vs. } \mathbb{P}(T^{\bar{a}'} \leq t, J = j)$$

- **Inverse probability of censoring weighting** (IPCW) for to account for (regular or artificial) dependent censoring, using time-dependent weights:

$$W_C(t) = \prod_{k=1}^{t} \frac{\mathbb{P}(\text{not censored at } k | \text{baseline covariates})}{\mathbb{P}(\text{not censored at } k | \text{covariates up to } k)}$$

  Censoring can be seen as a time-dependent treatment:

$$\text{ATE} = \mathbb{P}(Y_k^{a=1,\bar{c}_k=\bar{0}} = 1) \text{ vs. } \mathbb{P}(Y_k^{a=0,\bar{c}_k=\bar{0}} = 1)$$

  See `stteffects ipwra`

- **Standardisation/g-formula/robust methods** are alternatives

Options 1: Composite endpoints

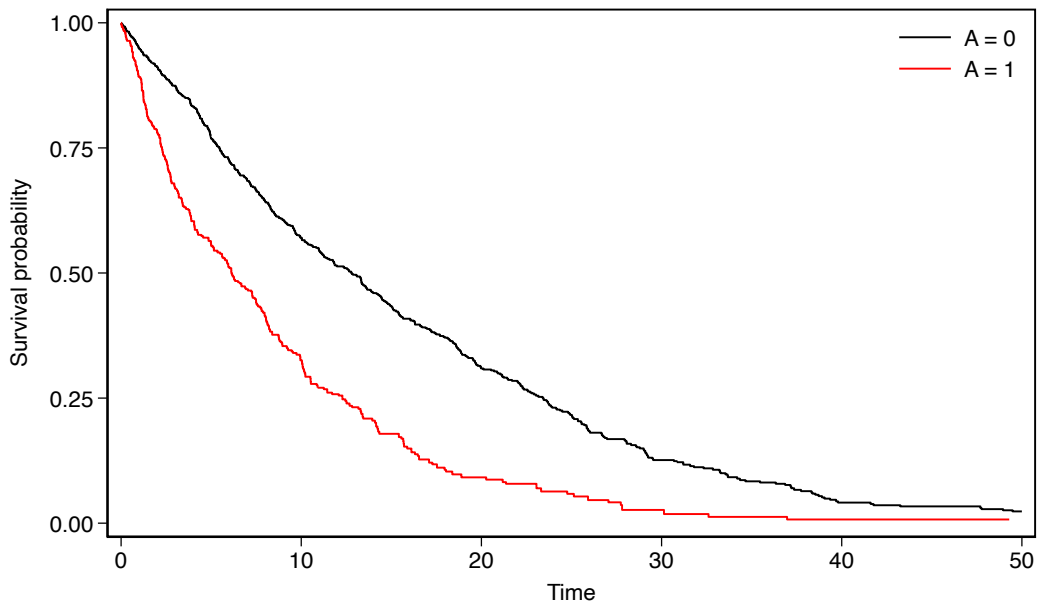
Options 1: Composite endpoints

- **Eliminate the problem** so that traditional time-to-event methods apply (but argue for the relevance of the new endpoint)

Options 1: Composite endpoints

- **Eliminate the problem** so that traditional time-to-event methods apply (but argue for the relevance of the new endpoint):

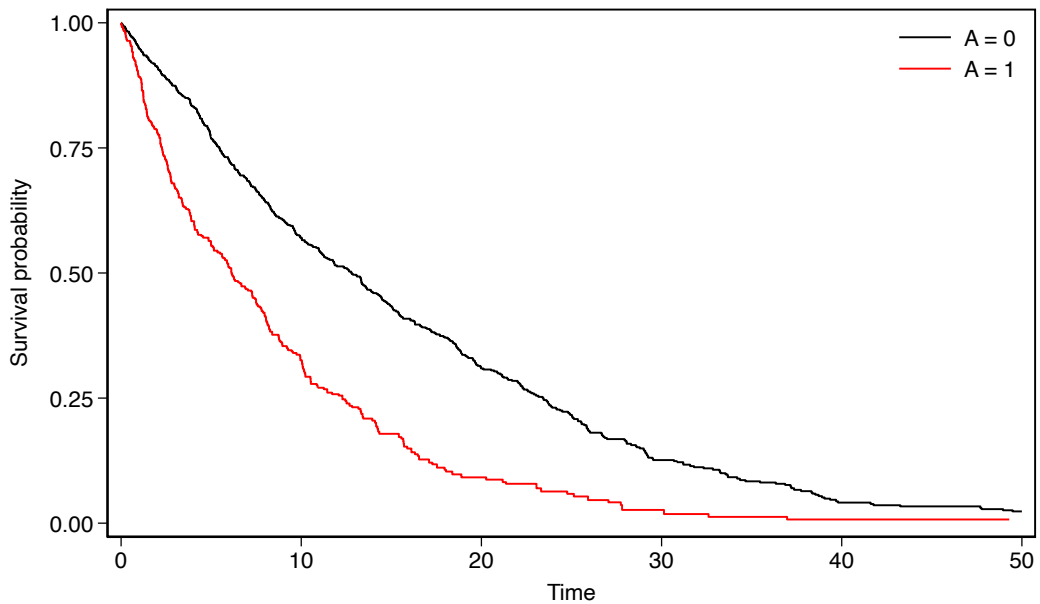
```
stset T [pweight=psw], fail(D == 1,2) id(id)
sts graph, by(A) plot1opts(lcolor(black)) plot2opts(lcolor(red))
ytitle("Survival probability") xtitle("Time") legend(ring(0) pos(1))
title("") ylab(, nogrid) xlab(, nogrid) plotregion(lstyle(refline))
```



Options 1: Composite endpoints

- **Eliminate the problem** so that traditional time-to-event methods apply (but argue for the relevance of the new endpoint):

```
stset T [pweight=psw], fail(D == 1,2) id(id)
sts graph, by(A) plot1opts(lcolor(black)) plot2opts(lcolor(red))
ytitle("Survival probability") xtitle("Time") legend(ring(0) pos(1))
title("") ylab(, nogrid) xlab(, nogrid) plotregion(lstyle(refline))
```



Look at, for example, "death by any cause" or "cancer free survival" and analyse as in simple survival settings

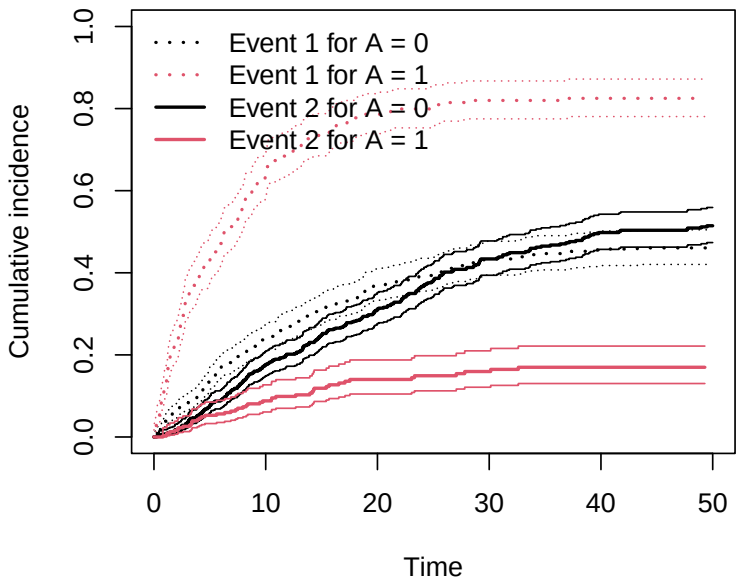
Option 2: Cause specific cumulative incidence

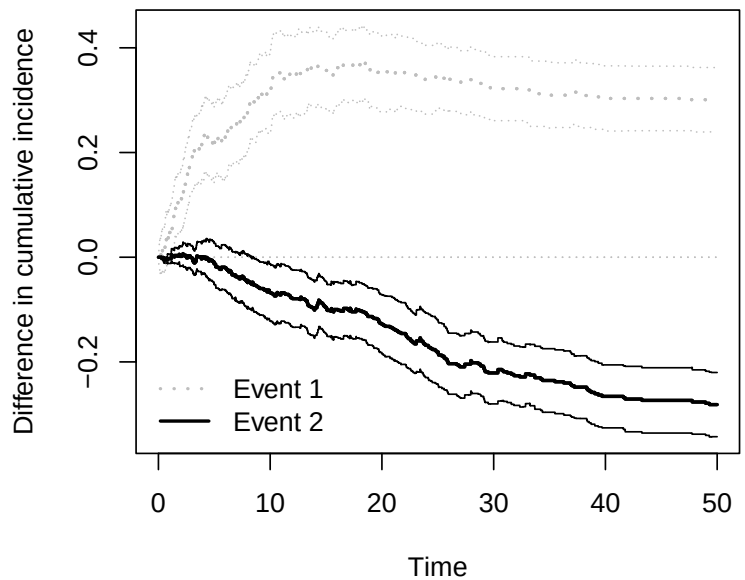
Option 2: Cause specific cumulative incidence

- **Allow for competing events**

Option 2: Cause specific cumulative incidence

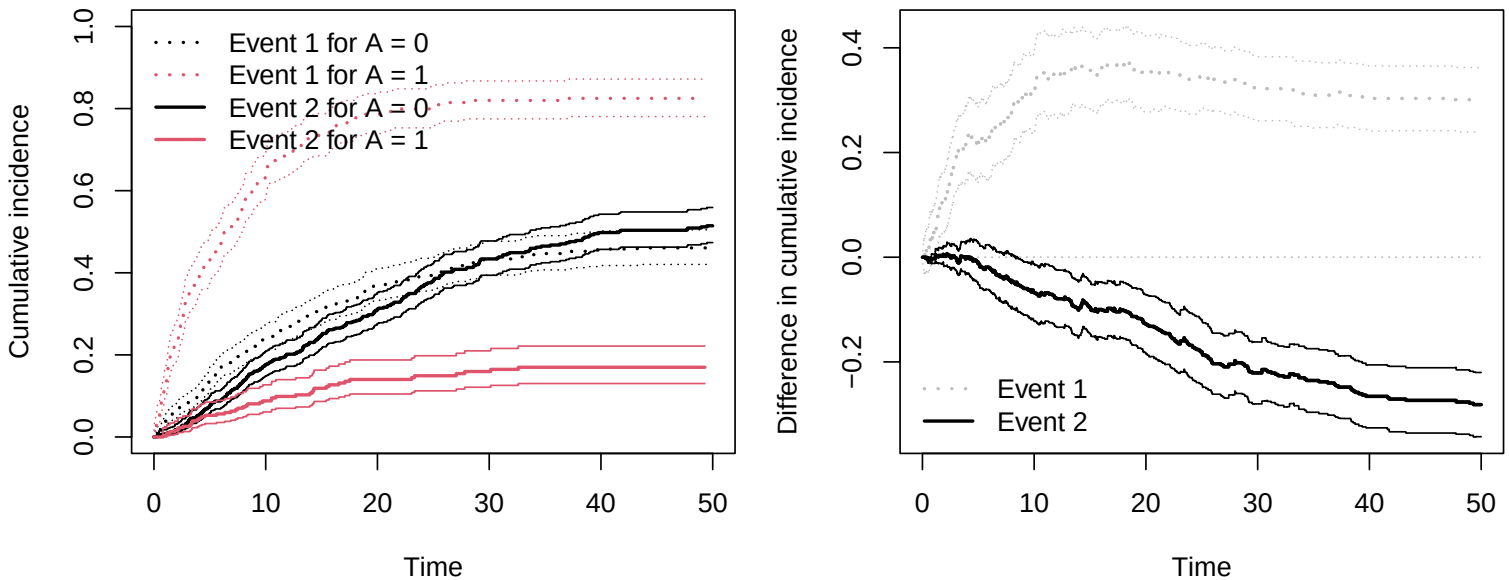
- **Allow for competing events**:





Option 2: Cause specific cumulative incidence

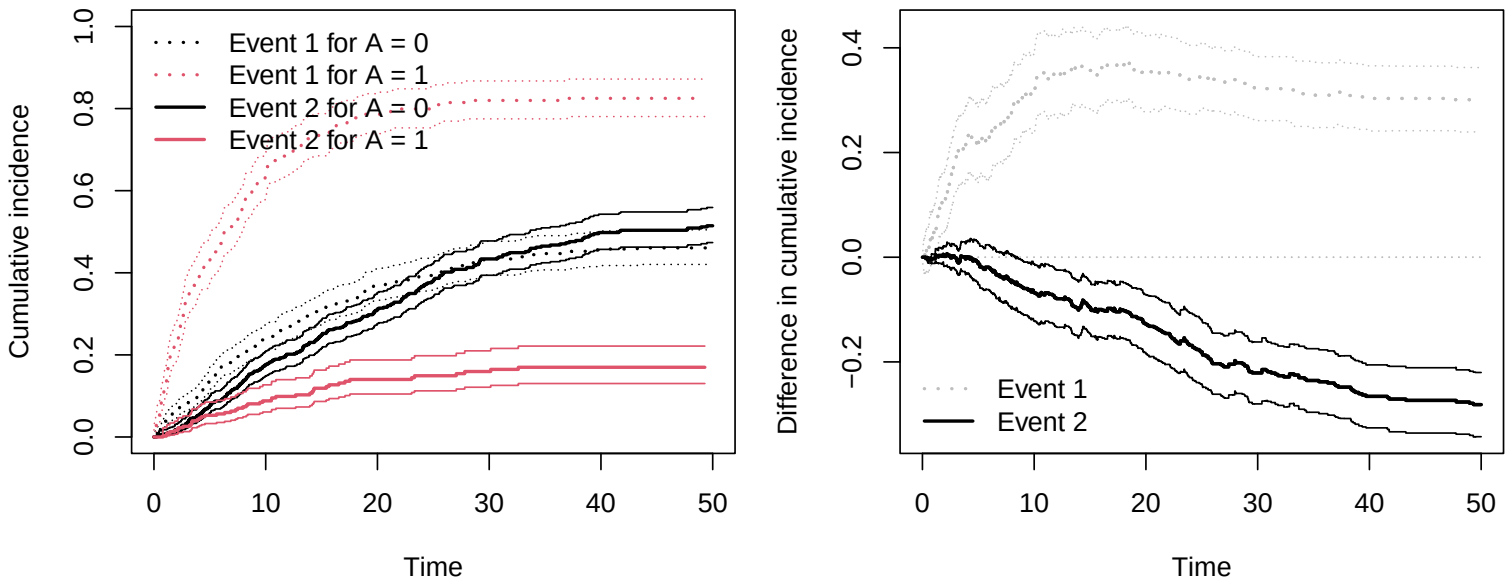
- **Allow for competing events**:



Look at difference between cumulative incidence curves (usually reasonable to show curves for all event types)

Option 2: Cause specific cumulative incidence

- **Allow for competing events**:

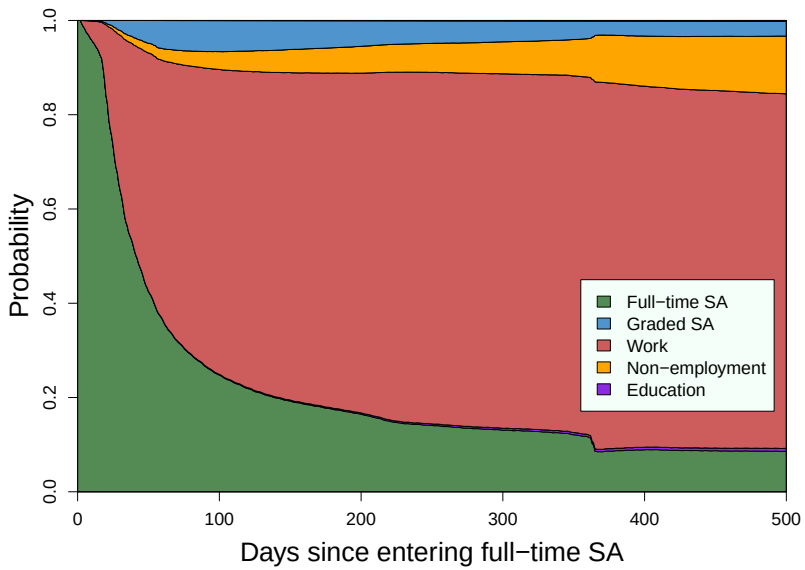


Look at difference between cumulative incidence curves (usually reasonable to show curves for all event types)

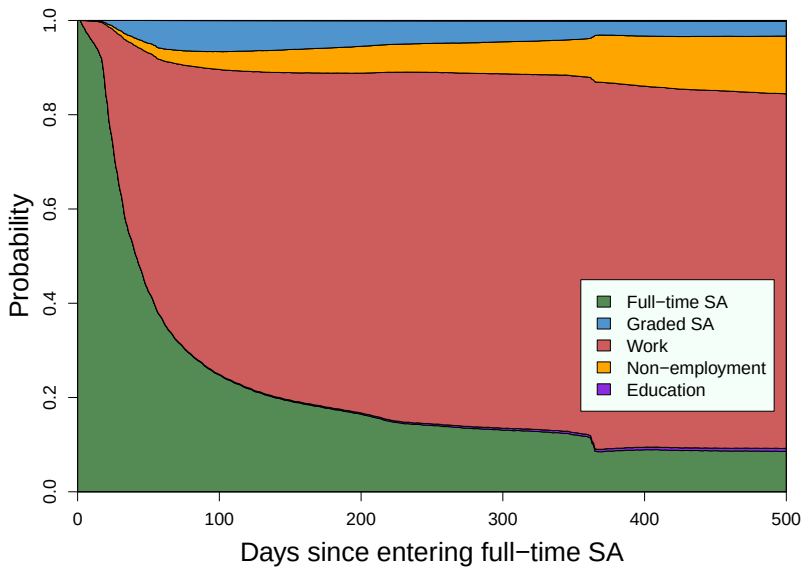
`stcompet` is unfortunately not reacting well to weights, see `stcrprep` and related functions by Lambert

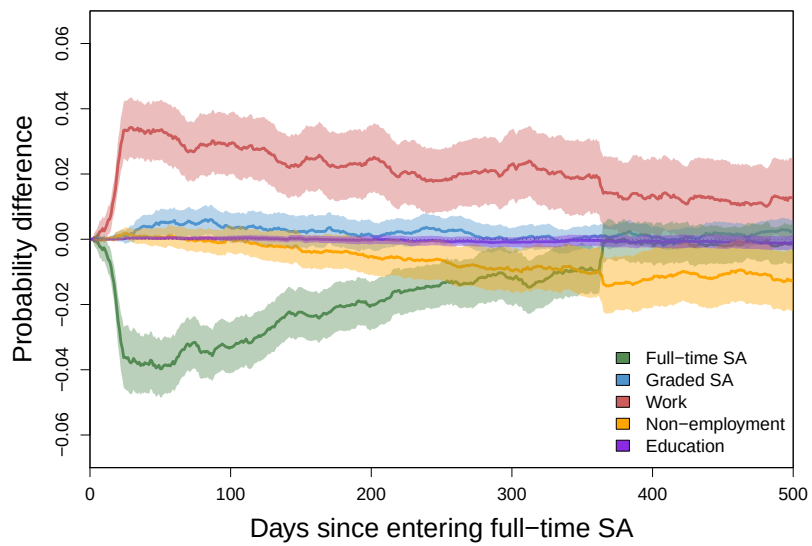
- Note: cumulative incidence easily generalises to more general **multi-state models**, using the Aalen-Johansen estimator

- Note: cumulative incidence easily generalises to more general **multi-state models**, using the Aalen-Johansen estimator (Gran et al. 2015, Hoff et al. 2022) :

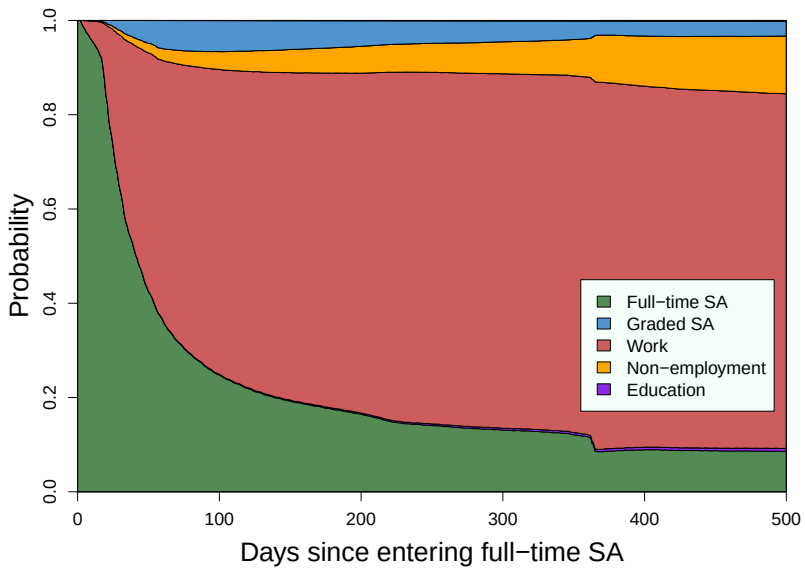


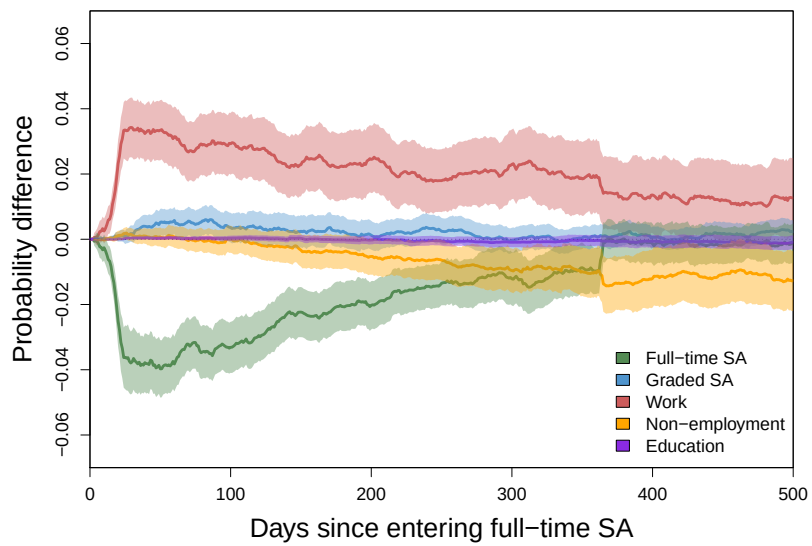
- Note: cumulative incidence easily generalises to more general **multi-state models**, using the Aalen-Johansen estimator (Gran et al. 2015, Hoff et al. 2022) :





- Note: cumulative incidence easily generalises to more general **multi-state models**, using the Aalen-Johansen estimator (Gran et al. 2015, Hoff et al. 2022) :





See `multistate` package by Crowther, Lambert and others for (mostly flexible parametric?) options in Stata

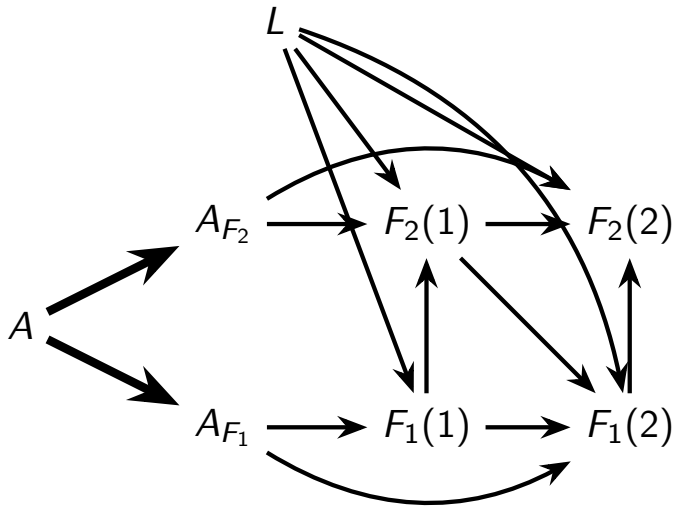
Option 3: More advanced estimands

Option 3: More advanced estimands

- **Direct effects** under elimination of competing events as defined by Young et al. (2020)

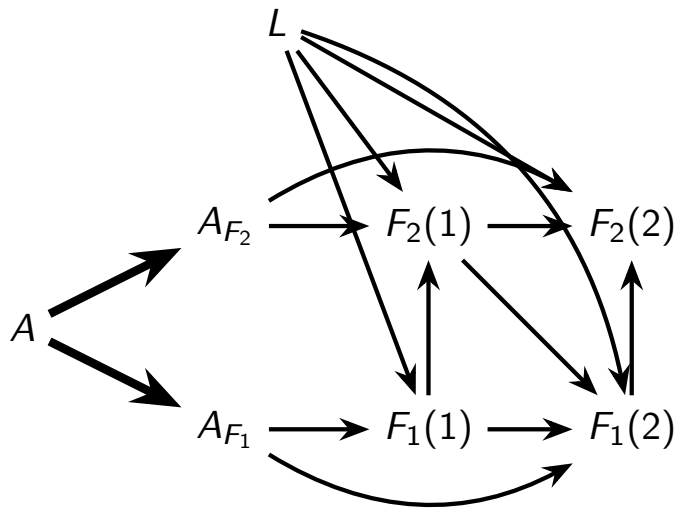
Option 3: More advanced estimands

- **Direct effects** under elimination of competing events as defined by Young et al. (2020)
- **Separable effects** as defined by Stensrud et al. (2022):



Option 3: More advanced estimands

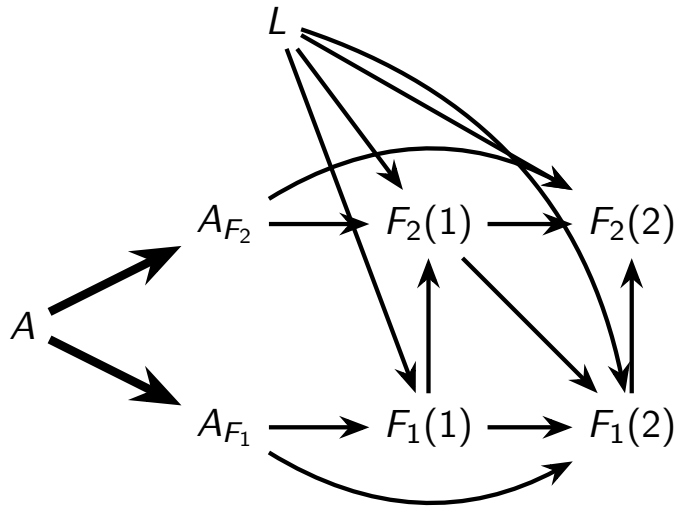
- **Direct effects** under elimination of competing events as defined by Young et al. (2020)
- **Separable effects** as defined by Stensrud et al. (2022):



(but more advanced estimands demand more assumptions)

Option 3: More advanced estimands

- **Direct effects** under elimination of competing events as defined by Young et al. (2020)
- **Separable effects** as defined by Stensrud et al. (2022):

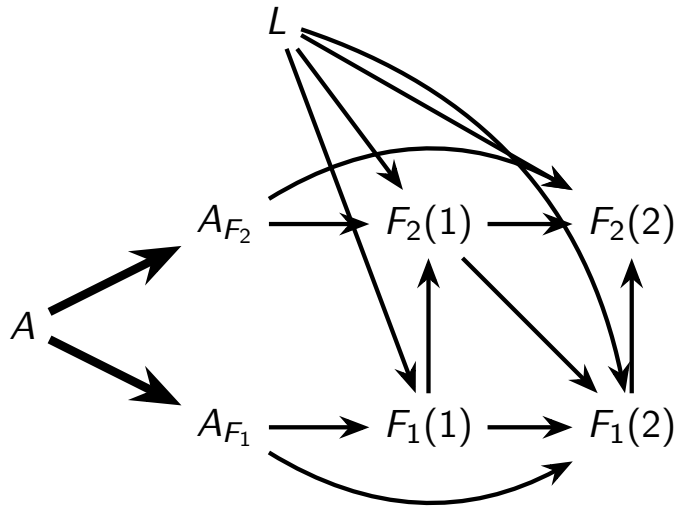


(but more advanced estimands demand more assumptions)

- **Maltzahn et al. (2024)** studies separable and controlled direct effects to estimate component specific effects on sickness absence

Option 3: More advanced estimands

- **Direct effects** under elimination of competing events as defined by Young et al. (2020)
- **Separable effects** as defined by Stensrud et al. (2022):



(but more advanced estimands demand more assumptions)

- **Maltzahn et al. (2024)** studies separable and controlled direct effects to estimate component specific effects on sickness absence
- **Stoltenberg et al.** with ongoing work on dynamic regimes based on opioid saving drug prescriptions

- As for estimating ATEs of **time-varying treatments** on cumulative incidence, this is mostly straight forward

- As for estimating ATEs of **time-varying treatments** on cumulative incidence, this is mostly straight forward

  ... as long as 'artificial censoring' and IPCW can be applied

- As for estimating ATEs of **time-varying treatments** on cumulative incidence, this is mostly straight forward

  ... as long as 'artificial censoring' and IPCW can be applied

- **G-formula or robust methods** get's more involved

- As for estimating ATEs of **time-varying treatments** on cumulative incidence, this is mostly straight forward

  ... as long as 'artificial censoring' and IPCW can be applied

- **G-formula or robust methods** get's more involved

  Some online Stata code accompanying the book by Hernán and Robins (2020), based on pooled logistic regression

- As for estimating ATEs of **time-varying treatments** on cumulative incidence, this is mostly straight forward

  ... as long as 'artificial censoring' and IPCW can be applied

- **G-formula or robust methods** get's more involved

  Some online Stata code accompanying the book by Hernán and Robins (2020), based on pooled logistic regression

  See also sequential trials (Gran et al. 2010, Keogh et al. 2023)

The extent of the problem

The extent of the problem

- In a review of all the 219 NEJM papers published in 2015, Schumacher et al. (2016) found 192 (88%) had a primary time-to-event outcome and 136 used time-to-event methodology (62%)

The extent of the problem

- In a review of all the 219 NEJM papers published in 2015, Schumacher et al. (2016) found 192 (88%) had a primary time-to-event outcome and 136 used time-to-event methodology (62%)

  51 had competing events and **only 26 (51%) dealt with it adequately**

4 References

- Andersen PK, Geskus RB, De Witte T and Putter H (2012). Competing risks in epidemiology: possibilities and pitfalls. *International journal of epidemiology*. DOI: 10.1093/ije/dyr213.
- Bernoulli D (1766). A new analysis of the mortality caused by smallpox. *Mem. Math. Phys. Acad. Roy. Sci. Paris*.
- Beyersmann J and Schrade C (2017). Florence Nightingale, William Farr and competing risks. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*. DOI: 10.1111/rssa.12187.
- Fine JP and Gray RJ (1999). A proportional hazards model for the subdistribution of a competing risk. *Journal of the American statistical association*. DOI: 10.1080/01621459.1999.10474144.
- Gran JM, Lie SA, Øyeflaten I, Borgan Ø and Aalen OO (2015). Causal inference in multi-state models – sickness absence and work for 1145 participants after work rehabilitation. *BMC public health*. DOI: 10.1186/s12889-015-2408-8.
- Gran JM, Røysland K, Wolbers M, Didelez V, Sterne JA, Ledergerber B, Furrer H, Von Wyl V and Aalen OO (2010). A sequential Cox approach for estimating the causal effect of treatment in the presence of time-dependent confounding applied to data from the Swiss HIV Cohort Study. *Statistics in medicine*. DOI: 10.1002/sim.4048.
- Hernán MA (2010). The hazards of hazard ratios. *Epidemiology*. DOI: 10.1097/EDE.0b013e3181c1ea43.

Hernán MA and Robins JM (2016). Using big data to emulate a target trial when a randomized trial is not available. *American journal of epidemiology*. DOI: 10.1093/aje/kwv254.

— (2020). *Causal Inference: What if*. Chapman & Hall/CRC. URL: https://www.hsph.harvard.edu/miguel-hernan/causal-inference-book.

Hoff R, Maltzahn N, Hasting R, Merkus SL, Undem K, Kristensen P, Mehlum IS and Gran JM (2022). An initiative for a more inclusive working life and its effect on return-to-work after sickness absence – a multi-state longitudinal cohort study. *BMJ Open*.

Janvin M, Young JG, Ryalen PC and Stensrud MJ (2024). Causal inference with recurrent and competing events. *Lifetime data analysis*. DOI: 10.1007/s10985-023-09594-8.

Keogh RH, Gran JM, Seaman SR, Davies G and Vansteelandt S (2023). Causal inference in survival analysis using longitudinal observational data: Sequential trials and marginal structural models. *Statistics in Medicine*. DOI: 10.1002/sim.9718.

Klein JP and Andersen PK (2005). Regression modeling of competing risks data based on pseudovalues of the cumulative incidence function. *Biometrics*. DOI: 10.1111/j.0006-341X.2005.031209.x.

Maltzahn NN, Mehlum IS and Gran JM (2024). Separable and controlled direct effects for competing events: Estimation of component specific effects on sickness absence. *Statistics in Medicine*. DOI: 10.1002/sim.10179.

Neyman JS (1950). *First course in probability and statistics*. Henry Holt & Company.

- Petersen ML and van der Laan MJ (2014). Causal models and learning from data: integrating causal modeling and statistical estimation. *Epidemiology*. DOI: 10.1097/EDE.0000000000000078.
- Schumacher M, Ohneberg K and Beyersmann J (2016). Competing risk bias was common in a prominent medical journal. *Journal of Clinical Epidemiology*. DOI: 10.1016/j.jclinepi.2016.07.013.
- Stensrud MJ, Young JG, Didelez V, Robins JM and Hernán MA (2022). Separable effects for causal inference in the presence of competing events. *Journal of the American Statistical Association*. DOI: 10.1080/01621459.2020.1765783.
- Young JG, Stensrud MJ, Tchetgen Tchetgen EJ and Hernán MA (2020). A causal framework for classical statistical estimands in failure-time settings with competing events. *Statistics in Medicine*. DOI: 10.1002/sim.8471.
- Aalen OO, Cook RJ and Røysland K (2015). Does Cox analysis of a randomized survival study yield a causal treatment effect? *Lifetime data analysis*. DOI: 10.1007/s10985-015-9335-y.
- Aalen OO and Johansen S (1978). An empirical transition matrix for non-homogeneous Markov chains based on censored observations. *Scandinavian Journal of Statistics*. URL: https://www.jstor.org/stable/4615704.