Survival Analysis, Lecture 6 Competing risks and multi-state models

Aurélien Belot

Inequalities in Cancer Outcomes Network (ICON)
London School of Hygiene and Tropical Medicine

Intended Learning Outcomes

At the end of this lecture, you should be able to

- Define what is meant by competing risks
- Explain why different methods are needed to analyse competing risks data
- Define, plot and interpret the cumulative incidence function
- Explain the difference between the subdistribution hazard and the cause-specific hazard approaches
- Explain key terms used in multi-state modelling, including the Markov property

What is a competing risks setting?

Definition

When the occurrence of a competing event modifies the rate of the event of interest, or prevents the event of interest from happening.

- This depletion of the risk set due to the competing risk is informative and needs to be accounted for if the aim is modeling (or making inference) for event-specific quantities

What is a competing risks setting?

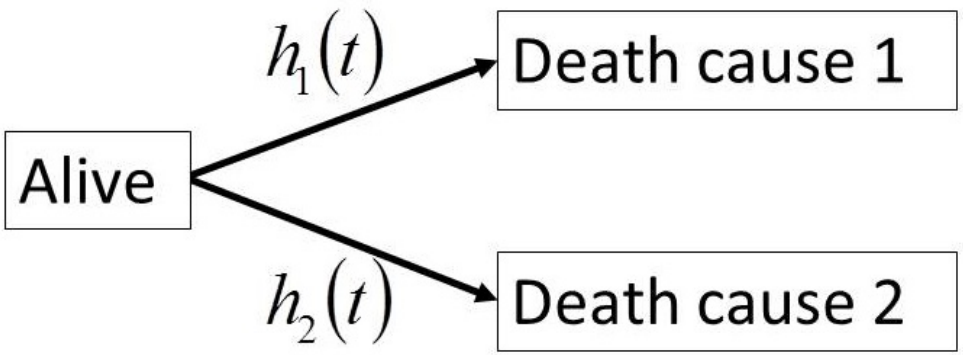
Definition

When the occurrence of a competing event modifies the rate of the event of interest, or prevents the event of interest from happening.

- This depletion of the risk set due to the competing risk is informative and needs to be accounted for if the aim is modeling (or making inference) for event-specific quantities

Examples of competing risks setting

- Time to heart disease death (competing risks: other causes of death)
- Time to cancer relapse or death (competing risks for each other)
- ...



- From overall mortality hazard (marginal hazard) to cause-specific hazards (event-specific hazards)

Additional information available

Notations

For a patient i, in addition to the the time of event T_i, the vital status indicator δ_i, and a vector of explanatory variables $\mathbf{x_i}$, we now have the cause of failure $D_i \in (1, ..., K)$:

$$\{T_i, \delta_i, \mathbf{x}_i, D_i\}, i = 1, .., N$$

Central quantities in competing risks setting

Cause-specific hazard

It represents the instantaneous probability of dying from cause k, given that at risk of dying, and divided by the infinitesimal interval length (i.e. it's a rate...)

$$h_k(t) = \lim_{dt \to 0} \frac{P(t \leq T < t + dt, D = k \mid T \geq t)}{dt}$$

Cumulative incidence function (a.k.a absolute cause-specific risk, crude probability of event k)

It defines the probability of failing from cause k before time t

$$I_k(t) = P(T \leq t, D = k) = \int_0^t S(u-) h_k(u) du$$

where $S(u)$ is the overall survival

Central quantities in competing risks setting

Cause-specific hazard

It represents the instantaneous probability of dying from cause k, given that at risk of dying, and divided by the infinitesimal interval length (i.e. it's a rate...)

$$h_k(t) = \lim_{dt \to 0} \frac{P(t \leq T < t + dt, D = k \mid T \geq t)}{dt}$$

Cumulative incidence function (a.k.a absolute cause-specific risk, crude probability of event k)

It defines the probability of failing from cause k before time t

$$I_k(t) = P(T \leq t, D = k) = \int_0^t S(u-) h_k(u) du$$

where $S(u)$ is the overall survival

Overall survival in competing risks setting

- Overall mortality hazard is the sum of all the cause-specific mortality hazards

$$h(t) = \sum_{k=1}^{K} h_k(t)$$

All events are exclusive (disjoints), so we have

$$P(t \leq T < t+dt) = P(\{t \leq T < t+dt, D=1\} \cup \cdots \cup \{t \leq T < t+dt, D=K\})$$

$$P(t \leq T < t+dt) = \sum_{k=1}^{K} P(t \leq T < t+dt, D=k)$$

Overall survival in competing risks setting

Overall survival

- Overall mortality hazard is the sum of all the cause-specific mortality hazards

$$h(t) = \sum_{k=1}^{K} h_k(t)$$

- So the overall survival is then defined as

$$S(t) = \exp(-\int_0^t \sum_{k=1}^{K} h_k(u)du) = \exp(-\sum_{k=1}^{K} H_k(t))$$

Why do we need new quantities?

A single event type

$P(T \leq t) = 1 - S(t) = \int_0^t h(u)S(u)du = \int_0^t h(u)exp(-\int_0^u h(v)dv)du$

Two event types

$P(T \leq t) = \int_0^t h(u)S(u)du = \int_0^t \{h_1(u) + h_2(u)\}S(u)du$
where $S(u) = \exp(-\int_0^u h_1(v) + h_2(v)dv$

In competing risks setting, the one-to-one relationship between hazard and risk (probability) is lost

Why do we need new quantities? an illustrative example

- We observed N patients with a given cancer
- When the patient died, we knew the cause of death (cancer or other causes)
- What's the probability of dying from cancer?

Why do we need new quantities? an illustrative example

Naive analysis

- Censor patients dying from other cause
- Calculate the corresponding survival (e.g. with Kaplan-Meier, KM_c).
- Use $1 - KM_c$ as an estimation of the probability of dying from cancer

It is a biased estimation (upward) of the probability of dying from cancer.

Why?

$$1 - KM_c(t) = \int_0^t h_c(u) KM_c(u) du \geq \int_0^t h_c(u) S(u) du$$

Why do we need new quantities? an illustrative example

Naive analysis

- Censor patients dying from other cause
- Calculate the corresponding survival (e.g. with Kaplan-Meier, KM_c).
- Use $1 - KM_c$ as an estimation of the probability of dying from cancer

It is a biased estimation (upward) of the probability of dying from cancer.
Why?

$$1 - KM_c(t) = \int_0^t h_c(u) KM_c(u) du \geq \int_0^t h_c(u) S(u) du$$

Why do we need new quantities? an illustrative example

Naive analysis

- Censor patients dying from other cause
- Calculate the corresponding survival (e.g. with Kaplan-Meier, KM_c).
- Use $1 - KM_c$ as an estimation of the probability of dying from cancer

It is a biased estimation (upward) of the probability of dying from cancer.
Why?

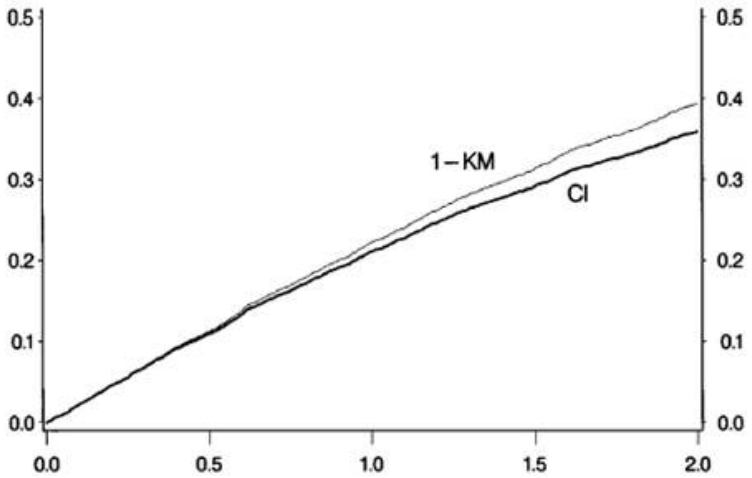
$$1 - KM_c(t) = \int_0^t h_c(u) KM_c(u) du \geq \int_0^t h_c(u) S(u) du$$

Illustration of the difference between the CIF and 1-KM

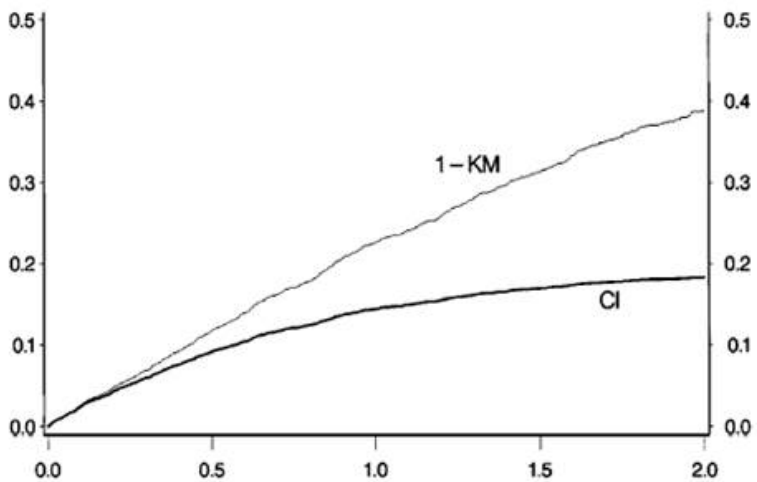
Example with 2 different levels of the competing-event rates

5000 simulated data

Low competing-event mortality hazard: 0.10



High competing-event mortality hazard: 0.90



From Gooley T.A. *et al.*, Stat med 1999.

Exercice 6.1: calculating the CIF

Cancer death = d_{1j}; Other-cause death = d_{2j} ; all death =d_j

Time	d_{1j}	d_{2j}	Cens.	d_j	n_j	$h_1(t_j)$	$h_2(t_j)$	$S(t_j)$	$I_1(t_j)$	$I_2(t_j)$
1	2	8	0		490					
2	4	7	0							
3	0	3	0							
4	3	2	4							
5	6	5	0							

Exercice 6.1: calculating the CIF

Cancer death = d_{1j}; Other-cause death = d_{2j}

Time	d_{1j}	d_{2j}	Cens.	d_j	n_j	$h_1(t_j)$	$h_2(t_j)$	$S(t_j)$	$I_1(t_j)$	$I_2(t_j)$
1	2	8	0	10	490	2/490	8/490	0.9796	0.0041	0.0163
2	4	7	0	11	480	4/480	7/480	0.9571	0.0123	0.0306
3	0	3	0	3	469	0/469	3/469	0.9510	0.0123	0.0367
4	3	2	4	5	466	3/466	2/466	0.9408	0.0184	0.0408
5	6	5	0	11	457	6/457	5/457	0.9182	0.0308	0.0511

$$S(1) = [1 - (10/490)] * S(0) = 0.9796$$
$$S(2) = [1 - (11/480)] * S(1) = 0.9571$$

$$I_1(1) = h_1(1) * S(0) = 0.0041$$
$$I_1(2) = h_1(2) * S(1) + I_1(1) = 0.0123$$

$$I_2(1) = h_2(1) * S(0) = 0.0163$$
$$I_2(2) = h_2(2) * S(1) + I_2(1) = 0.0306$$

Exercice 6.1: calculating the CIF

Cancer death = d_{1j}; Other-cause death = d_{2j}

Time	d_{1j}	d_{2j}	Cens.	d_j	n_j	$h_1(t_j)$	$h_2(t_j)$	$S(t_j)$	$I_1(t_j)$	$I_2(t_j)$
1	2	8	0	10	490	2/490	8/490	0.9796	0.0041	0.0163
2	4	7	0	11	480	4/480	7/480	0.9571	0.0123	0.0306
3	0	3	0	3	469	0/469	3/469	0.9510	0.0123	0.0367
4	3	2	4	5	466	3/466	2/466	0.9408	0.0184	0.0408
5	6	5	0	11	457	6/457	5/457	0.9182	0.0308	0.0511

$$S(1) = [1 - (10/490)] * S(0) = 0.9796$$
$$S(2) = [1 - (11/480)] * S(1) = 0.9571$$

$$I_1(1) = h_1(1) * S(0) = 0.0041$$
$$I_1(2) = h_1(2) * S(1) + I_1(1) = 0.0123$$

$$I_2(1) = h_2(1) * S(0) = 0.0163$$
$$I_2(2) = h_2(2) * S(1) + I_2(1) = 0.0306$$

Exercice 6.1: calculating the CIF

Cancer death = d_{1j}; Other-cause death = d_{2j}

Time	d_{1j}	d_{2j}	Cens.	d_j	n_j	$h_1(t_j)$	$h_2(t_j)$	$S(t_j)$	$I_1(t_j)$	$I_2(t_j)$
1	2	8	0	10	490	2/490	8/490	0.9796	0.0041	0.0163
2	4	7	0	11	480	4/480	7/480	0.9571	0.0123	0.0306
3	0	3	0	3	469	0/469	3/469	0.9510	0.0123	0.0367
4	3	2	4	5	466	3/466	2/466	0.9408	0.0184	0.0408
5	6	5	0	11	457	6/457	5/457	0.9182	0.0308	0.0511

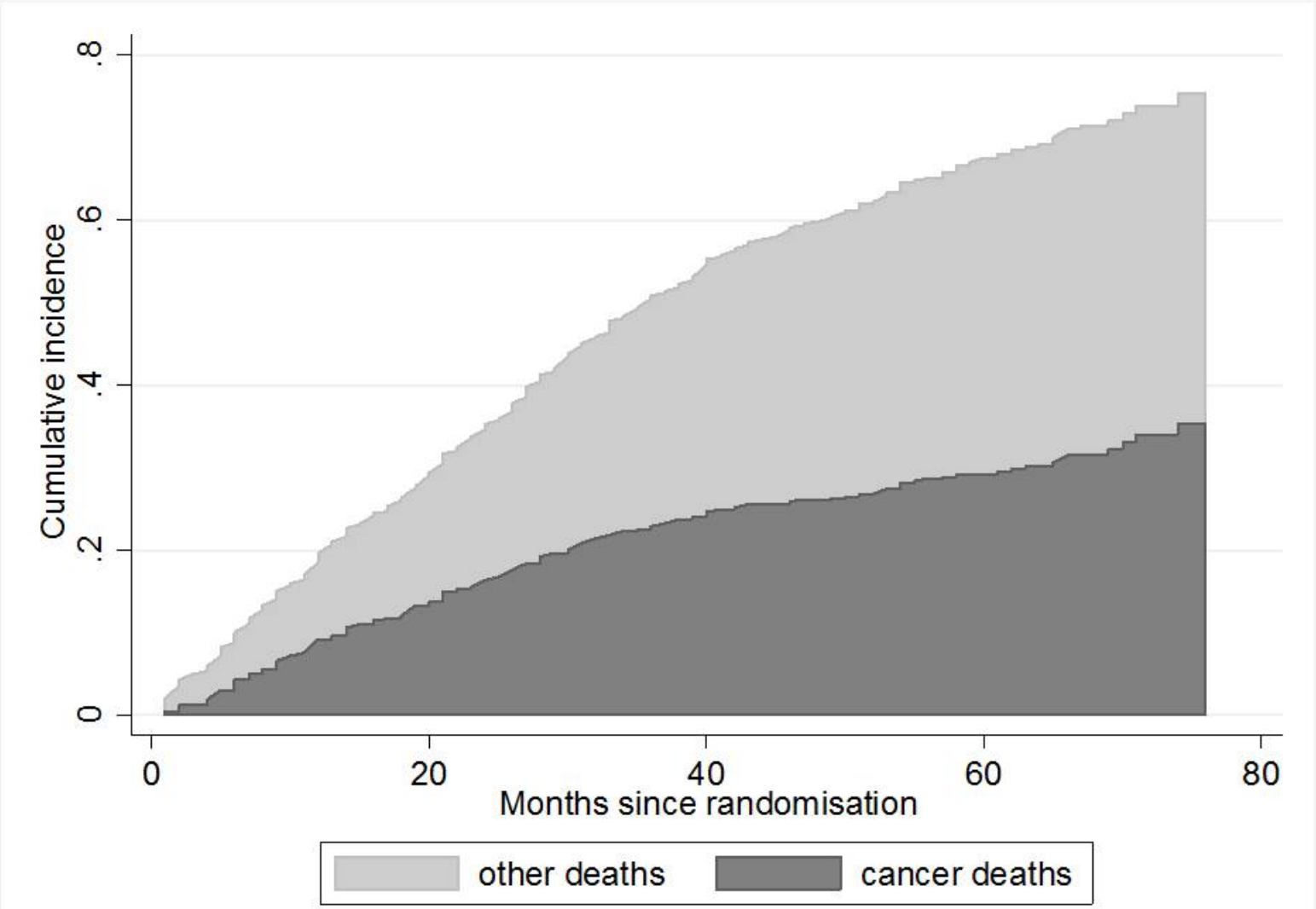
$$S(1) = [1 - (10/490)] * S(0) = 0.9796$$
$$S(2) = [1 - (11/480)] * S(1) = 0.9571$$

$$I_1(1) = h_1(1) * S(0) = 0.0041$$
$$I_1(2) = h_1(2) * S(1) + I_1(1) = 0.0123$$

$$I_2(1) = h_2(1) * S(0) = 0.0163$$
$$I_2(2) = h_2(2) * S(1) + I_2(1) = 0.0306$$

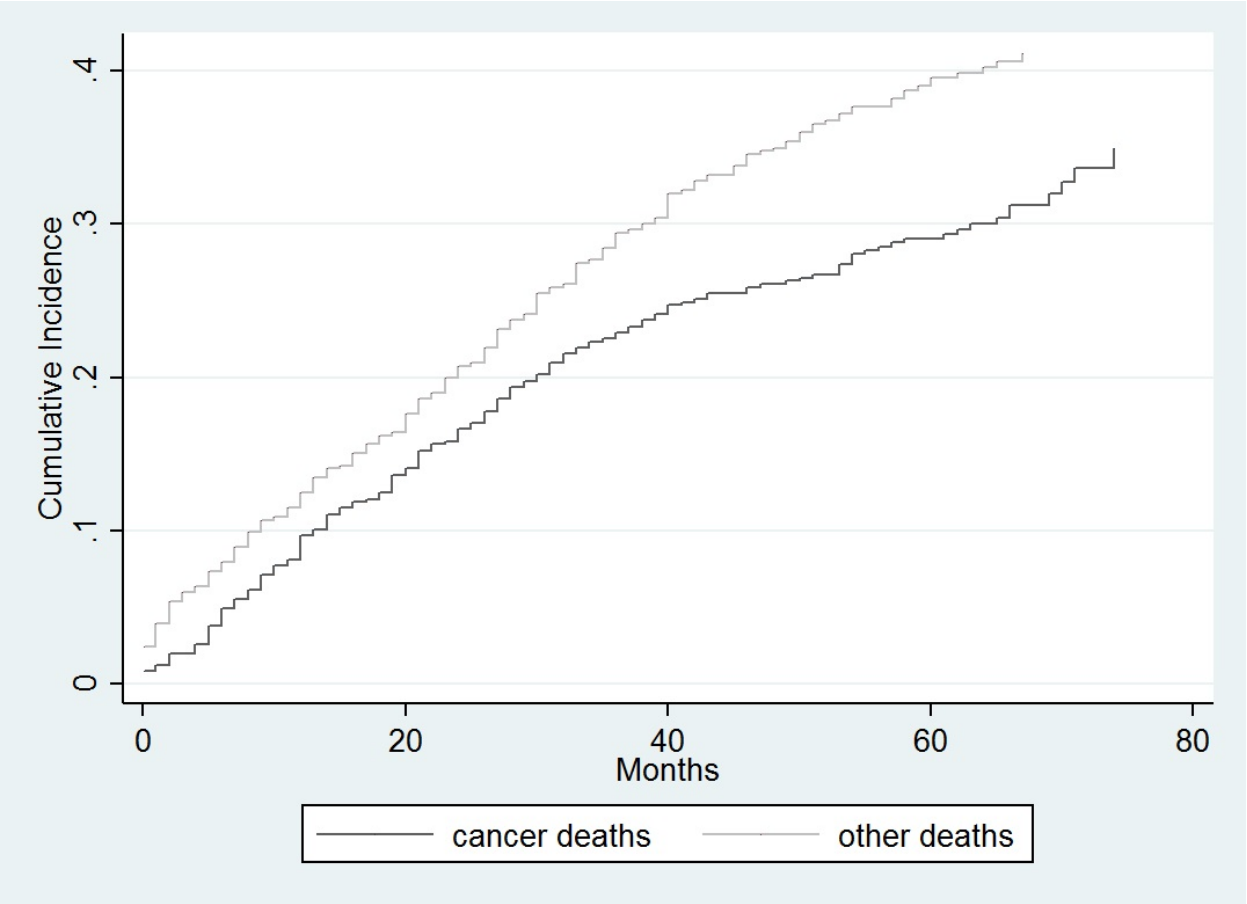
Graphical representation of CIF

Stacked format: informative if we want to emphasize how all event types add up to the overall risk



Graphical representation of CIF

Overlaid format: informative when comparing the progression to the different competing risks



Comparison between groups

Log-rank type test

- Comparison of cause-specific hazards: the same as the log-rank test
- Testing difference between CIFs
  - Test proposed by R.Gray
  - Test proposed by M. Pepe and M. Mori (Implemented in Stata with the command stpepemori)

For more details, please see chapter 2.6 from R.Geskus book.

Estimating effect of explanatory variables

Through cause-specific hazards modelling

- A regression model can be used for each cause-specific hazard (censoring other event types that the one of interest, see why in the appendix)
- This approach is perfectly fine
- But the effect of a variable on the cause-specific hazard translates hardly to the CIF scale (as it involves all the other cause-specific hazards)

Effect of variables using cause-specific hazards: an example

- Consider a binary variable x (0/1)
- We fit 2 cause-specific hazards with x as a covariate: $h_1(t;x) = h_{1,0}(t)e^{\beta_1 x}$ and $h_2(t;x) = h_{2,0}(t)e^{\beta_2 x}$
- Overall Survival $S(t;x) = \exp(-\int_0^t \{h_1(u;x) + h_2(u;x)\}du)$

Cumulative incidence of cause 1 for group with $x = 0$:
$I_1(t;x=0) = \int_0^t h_1(u;x=0)S(u;x=0)du = \int_0^t h_{1,0}(t)S(u;x=0)du$

Cumulative incidence of cause 1 for group with $x = 1$:
$I_1(t;x=1) = \int_0^t h_1(u;x=1)S(u;x=1)du = \int_0^t h_{1,0}(t)e^{\beta_1}S(u;x=1)du$

Contrasting $I_1(t;x=0)$ and $I_1(t;x=1)$ involves both β_1 and β_2
$\Rightarrow$ The subdistribution hazard

Effect of variables using cause-specific hazards: an example

- Consider a binary variable x (0/1)
- We fit 2 cause-specific hazards with x as a covariate: $h_1(t;x) = h_{1,0}(t)e^{\beta_1 x}$ and $h_2(t;x) = h_{2,0}(t)e^{\beta_2 x}$
- Overall Survival $S(t;x) = \exp(-\int_0^t \{h_1(u;x) + h_2(u;x)\}du)$

Cumulative incidence of cause 1 for group with $x = 0$:
$I_1(t;x=0) = \int_0^t h_1(u;x=0)S(u;x=0)du = \int_0^t h_{1,0}(t)S(u;x=0)du$

Cumulative incidence of cause 1 for group with $x = 1$:
$I_1(t;x=1) = \int_0^t h_1(u;x=1)S(u;x=1)du = \int_0^t h_{1,0}(t)e^{\beta_1}S(u;x=1)du$

Contrasting $I_1(t;x=0)$ and $I_1(t;x=1)$ involves both β_1 and β_2
$\Rightarrow$ The subdistribution hazard

Effect of variables using cause-specific hazards: an example

- Consider a binary variable x (0/1)
- We fit 2 cause-specific hazards with x as a covariate: $h_1(t;x) = h_{1,0}(t)e^{\beta_1 x}$ and $h_2(t;x) = h_{2,0}(t)e^{\beta_2 x}$
- Overall Survival $S(t;x) = \exp(-\int_0^t \{h_1(u;x) + h_2(u;x)\}du)$

Cumulative incidence of cause 1 for group with $x = 0$:
$I_1(t;x=0) = \int_0^t h_1(u;x=0)S(u;x=0)du = \int_0^t h_{1,0}(t)S(u;x=0)du$

Cumulative incidence of cause 1 for group with $x = 1$:
$I_1(t;x=1) = \int_0^t h_1(u;x=1)S(u;x=1)du = \int_0^t h_{1,0}(t)e^{\beta_1}S(u;x=1)du$

Contrasting $I_1(t;x=0)$ and $I_1(t;x=1)$ involves both β_1 and β_2
$\Rightarrow$ The subdistribution hazard

Effect of variables using cause-specific hazards: an example

- Consider a binary variable x (0/1)
- We fit 2 cause-specific hazards with x as a covariate: $h_1(t;x) = h_{1,0}(t)e^{\beta_1 x}$ and $h_2(t;x) = h_{2,0}(t)e^{\beta_2 x}$
- Overall Survival $S(t;x) = \exp(-\int_0^t \{h_1(u;x) + h_2(u;x)\}du)$

Cumulative incidence of cause 1 for group with $x = 0$:
$I_1(t;x=0) = \int_0^t h_1(u;x=0)S(u;x=0)du = \int_0^t h_{1,0}(t)S(u;x=0)du$

Cumulative incidence of cause 1 for group with $x = 1$:
$I_1(t;x=1) = \int_0^t h_1(u;x=1)S(u;x=1)du = \int_0^t h_{1,0}(t)e^{\beta_1} S(u;x=1)du$

Contrasting $I_1(t;x=0)$ and $I_1(t;x=1)$ involves both β_1 and β_2
$\Rightarrow$ The subdistribution hazard

The subdistribution hazard

- Motivation: to define a regression model allowing for estimated parameter directly associated to the CIF
- The most used model is the one proposed by Fine & Gray
- Its parametrisation follows the same principle as the semi-parametric Cox model

The subdistribution hazard

$$h_k^s(t) = \lim_{dt \to 0} \frac{P(t \leq T < t + dt, D = k \mid T \geq t \cup (T \leq t, D \neq k))}{dt}$$

$$h_k^s(t; x) = h_{0,k}^s(t) \exp(\gamma_k x)$$

Main difference with the cause-specific hazard: the risk-set

Subdistribution hazard: how does it link to the CIF?

The relationship

$h_k^s(t;x) = -\frac{d}{dt} log\big(1 - I_k(t;x)\big)$ gives

$$I_k(t;x) = 1 - \exp(-\int_0^t h_k^s(u;x)du)$$

where $h_k^s(t;x) = h_{0,k}^s(t)\exp(\gamma_k x)$

Main aim of the subdistribution hazard: make inference of the CIF

For a binary variable x we have

$$1 - I_k(t;x=1) = \big(1 - I_k(t;x=0)\big)^{\exp(\gamma_k)}$$

Contrasting $I_k(t;x=1)$ and $I_k(t;x=0)$ involves only e^{γ_k}

Subdistribution hazard: how does it link to the CIF?

The relationship

$h_k^s(t;x) = -\frac{d}{dt} log\big(1 - I_k(t;x)\big)$ gives

$$I_k(t;x) = 1 - \exp(-\int_0^t h_k^s(u;x)du)$$

where $h_k^s(t;x) = h_{0,k}^s(t)\exp(\gamma_k x)$

Main aim of the subdistribution hazard: make inference of the CIF

For a binary variable x we have

$$1 - I_k(t;x=1) = \big(1 - I_k(t;x=0)\big)^{\exp(\gamma_k)}$$

Contrasting $I_k(t;x=1)$ and $I_k(t;x=0)$ involves only e^{γ_k}

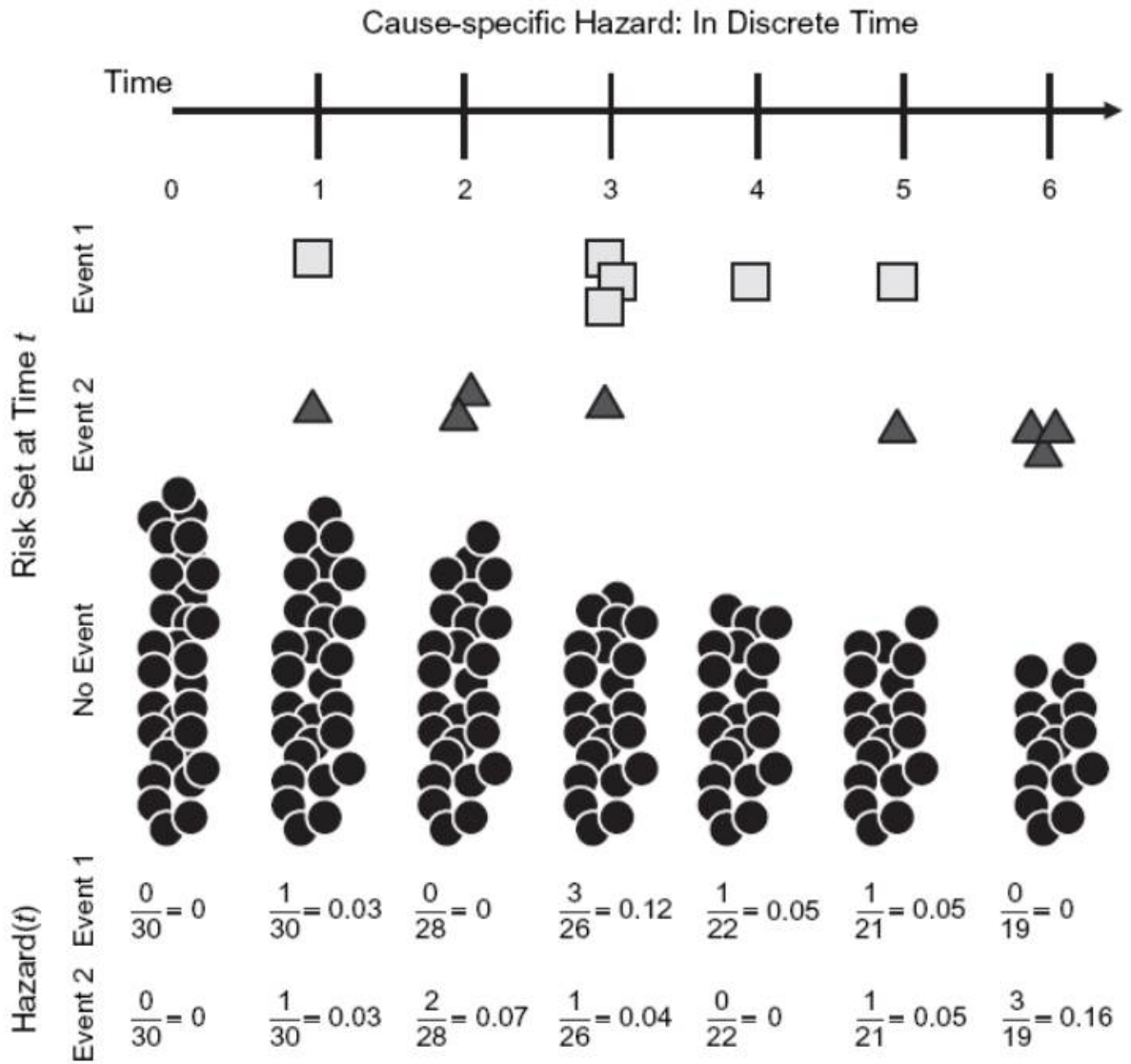
Subdistribution hazard: Interpretation

The numerical (quantitative) interpretation of sHR is not straightforward (due to the risk-set). However, a qualitative interpretation is useful:

- a $sHR = 1$ implies no association between the covariate and the corresponding CIF
- a $sHR > 1$ implies that an increase of the covariate value is associated with an increased risk,
- a $sHR < 1$ implies that an increase of the covariate value is associated with an decreased risk

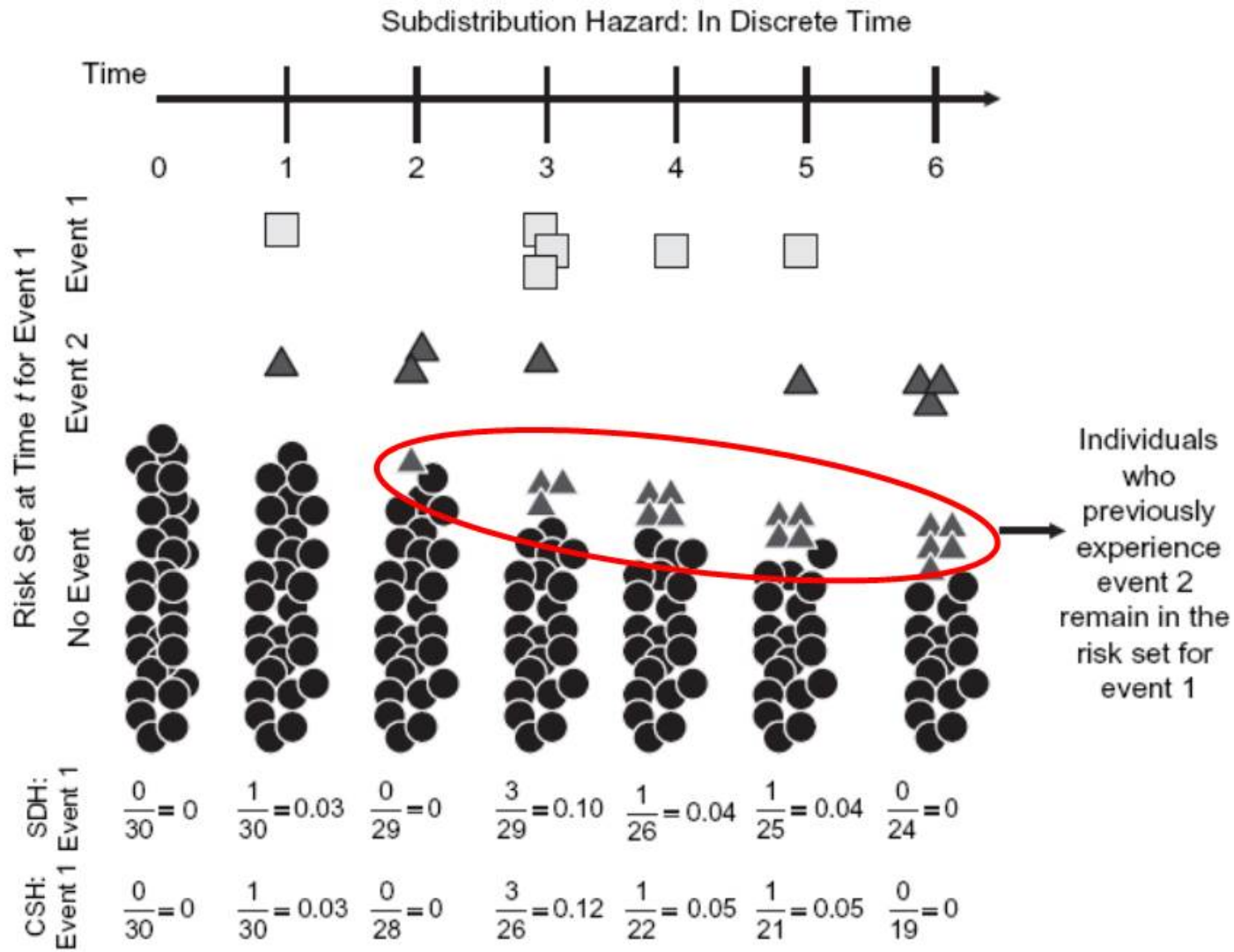
Moreover, equation linking the subdistribution hazard and the Cumulative Incidence Function is useful for prediction

Risk set for the Cause-specific Hazard



From Lau et al., American Journal of Epidemiology 2009

Risk set for the Subdistribution Hazard

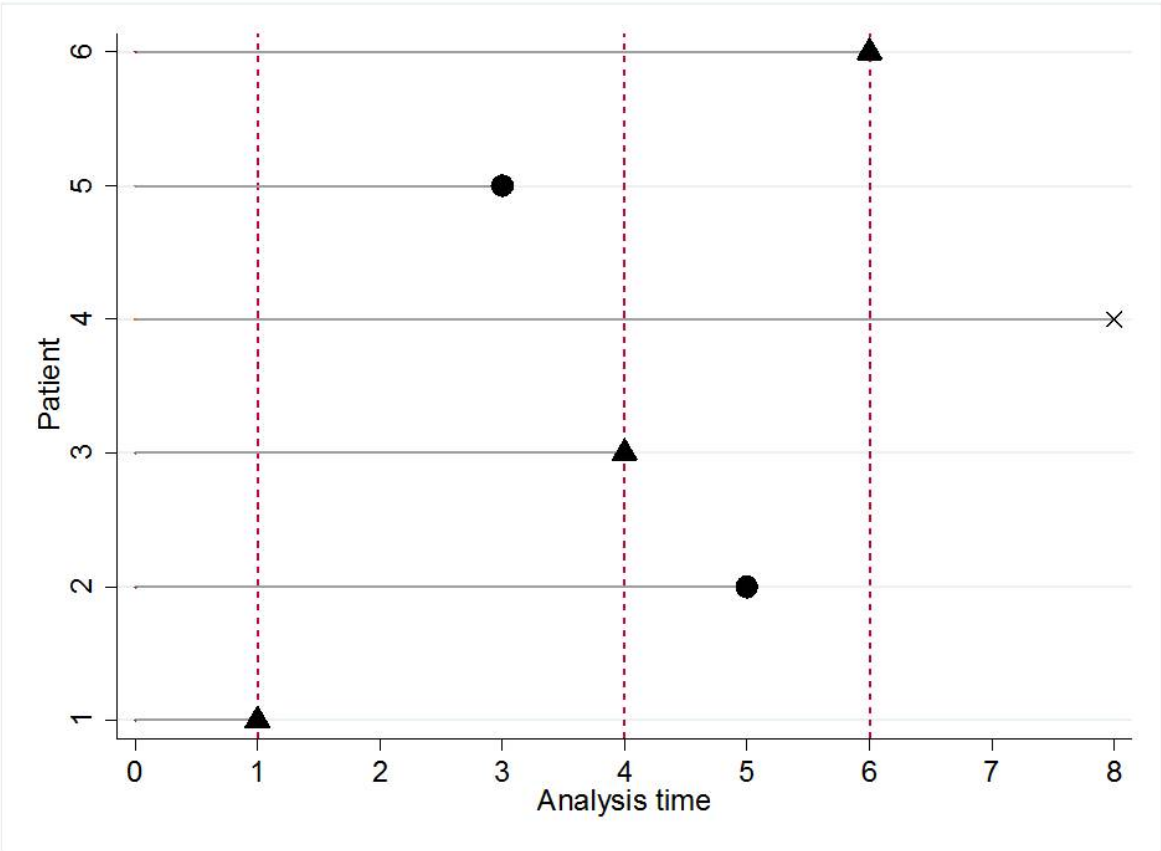


From Lau et al., American Journal of Epidemiology 2009

Risk set: Exercice 6.2

Cause-specific Hazard

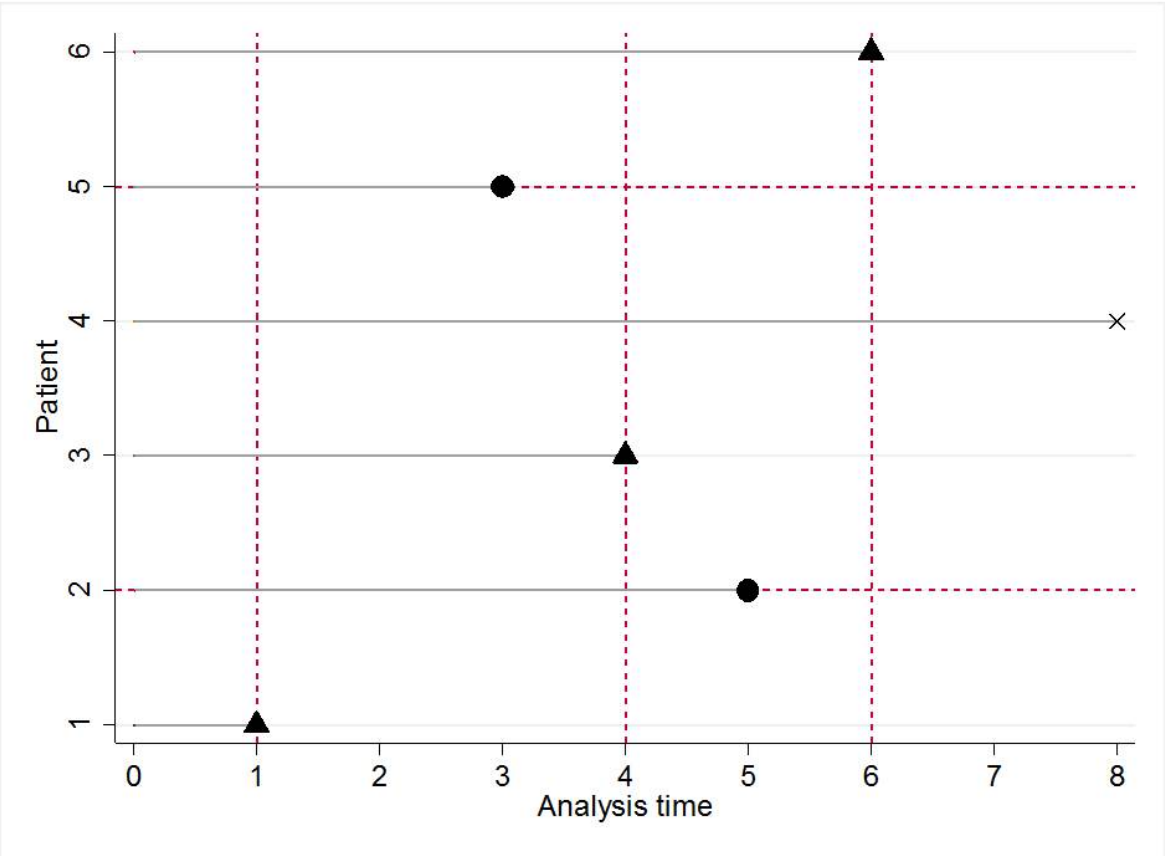
(triangles represent the event of interest (cancer death), circles represent competing risk events (other deaths) and crosses represent end-of-follow-up censoring)



Risk set: Exercice 6.2

Subdistribution Hazard

(triangles represent the event of interest (cancer death), circles represent competing risk events (other deaths) and crosses represent end-of-follow-up censoring)



Risk set: Exercice 6.2

Complete the table, calculating the cause-specific hazard (CSH) and the subdistribution hazard (SH) for cancer death at each time-point shown in the previous figure

Time	1	2	3	4	5	6	7	8
CSH								
SH								

Risk set: Exercice 6.2

Complete the table, calculating the cause-specific hazard (CSH) and the subdistribution hazard (SH) for cancer death at each time-point shown in the previous figure

Time	1	2	3	4	5	6	7	8
CSH	1/6 =0.17	0/5	0/5	1/4=0.25	0/3	1/2=0.5	0/1	0/1
SH	1/6 =0.17	0/5	0/5	1/5=0.20	0/4	1/4=0.25	0/3	0/3

Summary and recommendations

- Two ways of analysing competing risks data: cause-specific hazards analysis or subdistribution hazards analysis
- Cause-specific hazard useful to assess aetiological effects of a variable
- Subdistribution hazard useful to make inference and prediction of the CIF
- Competing risks analysis should report the results of both analysis (cause-specific and subdistribution) to give a full picture of the story

Appendix: likelihood for competing-risks data

Defined accordingly to all the cause-specific hazards

Observed data: $\{t_i, D_i, \delta_i\}$

$$\prod_{i=1}^{N} (h_{D_i}(t_i))^{\delta_i} . exp(- \int_0^{t_i} h(u) du)$$

where $h(t) = \sum_{k=1}^{K} h_k(t)$

The likelihood factors into a separate component for each failure type k (the "cause k-specific likelihood")

$$\prod_{k=1}^{K} \prod_{i=1}^{N} (h_k(t_i))^{\mathbb{I}(D_i=k)} . exp(- \int_0^{t_i} h_k(u) du)$$

Where $\mathbb{I}()$ defines the indicator function

So it shows why, to estimate the cause-k specific hazard, censoring all other causes works.

Appendix: likelihood for competing-risks data

Defined accordingly to all the cause-specific hazards

Observed data: $\{t_i, D_i, \delta_i\}$

$$\prod_{i=1}^{N} (h_{D_i}(t_i))^{\delta_i} . exp(- \int_0^{t_i} h(u) du)$$

where $h(t) = \sum_{k=1}^{K} h_k(t)$

The likelihood factors into a separate component for each failure type k (the "cause k-specific likelihood")

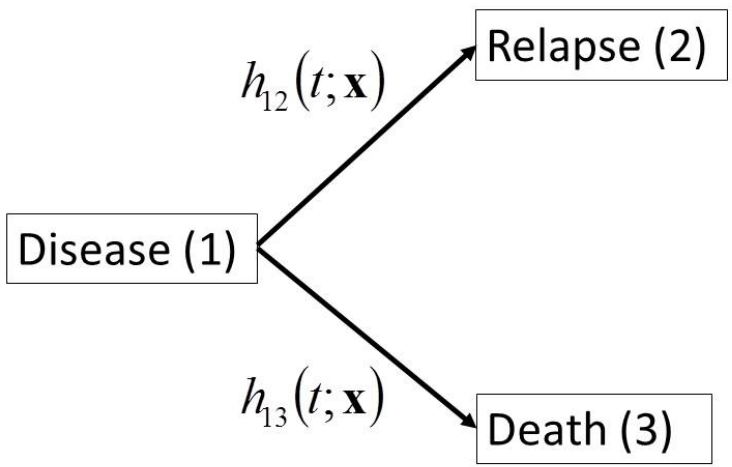
$$\prod_{k=1}^{K} \prod_{i=1}^{N} (h_k(t_i))^{\mathbb{I}(D_i=k)} . exp(- \int_0^{t_i} h_k(u) du)$$

Where $\mathbb{I}()$ defines the indicator function

So it shows why, to estimate the cause-k specific hazard, censoring all other causes works.

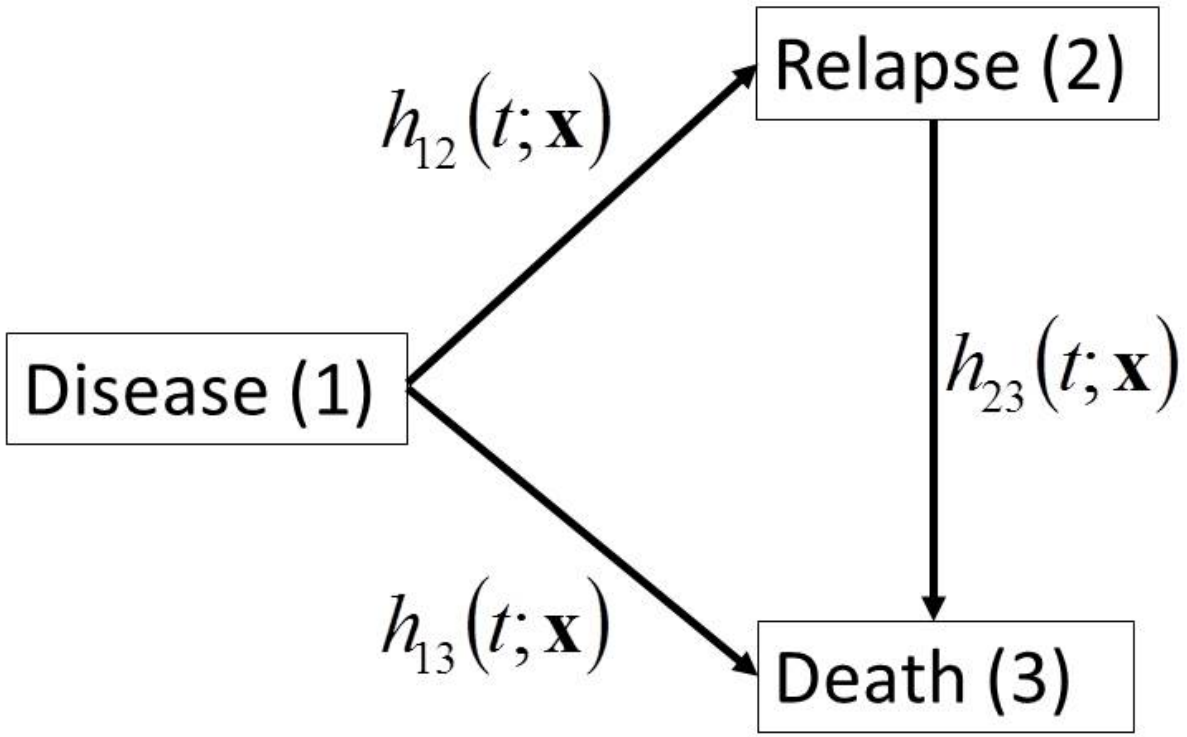
Multistate models: Motivations

Competing-risks models are specific multistate models, where we have only (pseudo) absorbing events



Studying intermediate events (like relapse) may give important information $\Rightarrow$ multistate models

Multistate models: Motivations



Multistate models: Terminology and notations

Terminology

- A box in previous graphs is called a “state”
  - Absorbing state: from which no transition is possible
  - Transient state: not absorbing state
- A direct arrow defines a transition from one state to another

Notations

- State space : $\mathbf{E} : \{1, ..., E\}$
- Random process $\{X(t); t \geq 0\}$ taking values in $\mathbf{E}$
- History of the process until time s: $\mathbb{H}_s = \{X(u); 0 \leq u \leq s\}$

Statistical quantities

Transition intensity

We can describe the process via the transition intensity/hazard

$$h_{ij}(t) = \lim_{dt \to 0} \frac{P(X(t+dt) = j \mid X(t) = i, \mathbb{H}_t)}{dt}$$

The Markov Assumption

$$P(X(t+dt) = j \mid X(t) = i, \mathbb{H}_t) = P(X(t+dt) = j \mid X(t) = i)$$

The transition intensity (and so the transition probability) depends on the past only through the current state

Extended markov model (or semi-Markov model): the transition probability depend on the current state and on how long ago it was reached

Estimation

Data management

- From a wide format (one line per patient) to a long format (multiple lines per patients, one for each transition for which the patient is at risk)
- With starting and stopping time

Estimation

- Apply either non-parametric estimate or regression model (Cox) for each transition separately (i.e. from the long format dataset, by selecting only the rows of the transition of interest).
- Patients moving to another state that the one of interest are censored in the estimation process for this transition.

Example: Exercice 6.3 (to be dicussed in the practical)

- 2204 patients receiving a bone marrow transplant.
- Three-state model:
  - State 1: transplanted
  - State 2: platelet recovery
  - State 3: relapse/death (from either state 1 or 2)
- Two multistate models were fitted

Example: Exercice 6.3

Stratified Analysis

For $i \in \{1,2\}$ and $j \in \{2,3\}$

$$h_{ij}(t;x) = h_{0,ij}(t)\exp(-0.05 age_{20-40} + 0.145 age_{40+} + 0.041 sex_{mismatch})$$

Interpretation:

- Age 40+ have increased transition hazard compared to age < 20
- From the 95% Conf. Interval: no evidence of differences in transition hazard between age < 20 and age 20-40, or by donor match status
- Assumptions: Markov (independent of history) and common effect of each variable across all transitions

Example: Exercice 6.3

Separated Analysis

For state 1 to 2 transition

$$h_{12}(t;x) = h_{0,12}(t)\exp(-0.183 age_{20-40} - 0.139 age_{40+} + 0.039 sexmismatch)$$

Interpretation:

- Older ages associated with reduced rate of platelet recovery – good evidence of a difference between < 20 and 20-40 age group
- From the 95% Conf. Interval: no evidence of difference in hazard of platelet recovery by donor match status
- Assumptions: Markov (independent of history)