Sciences Economiques et Sociales de la Santé
& Traitement de l'Information Médicale

sesstim.univ-amu.fr

Ronald GESKUS

Head of biostatistics group - Oxford University Clinical Research Unit Ho Chi Minh City, Vietnam.
Associate professor Nuffield Department of Medicine - University of Oxford, United Kingdom.

Competing risks, analysis and interpretation

juin 2020

Cliquez ici pour voir l'intégralité des ressources associées à ce document

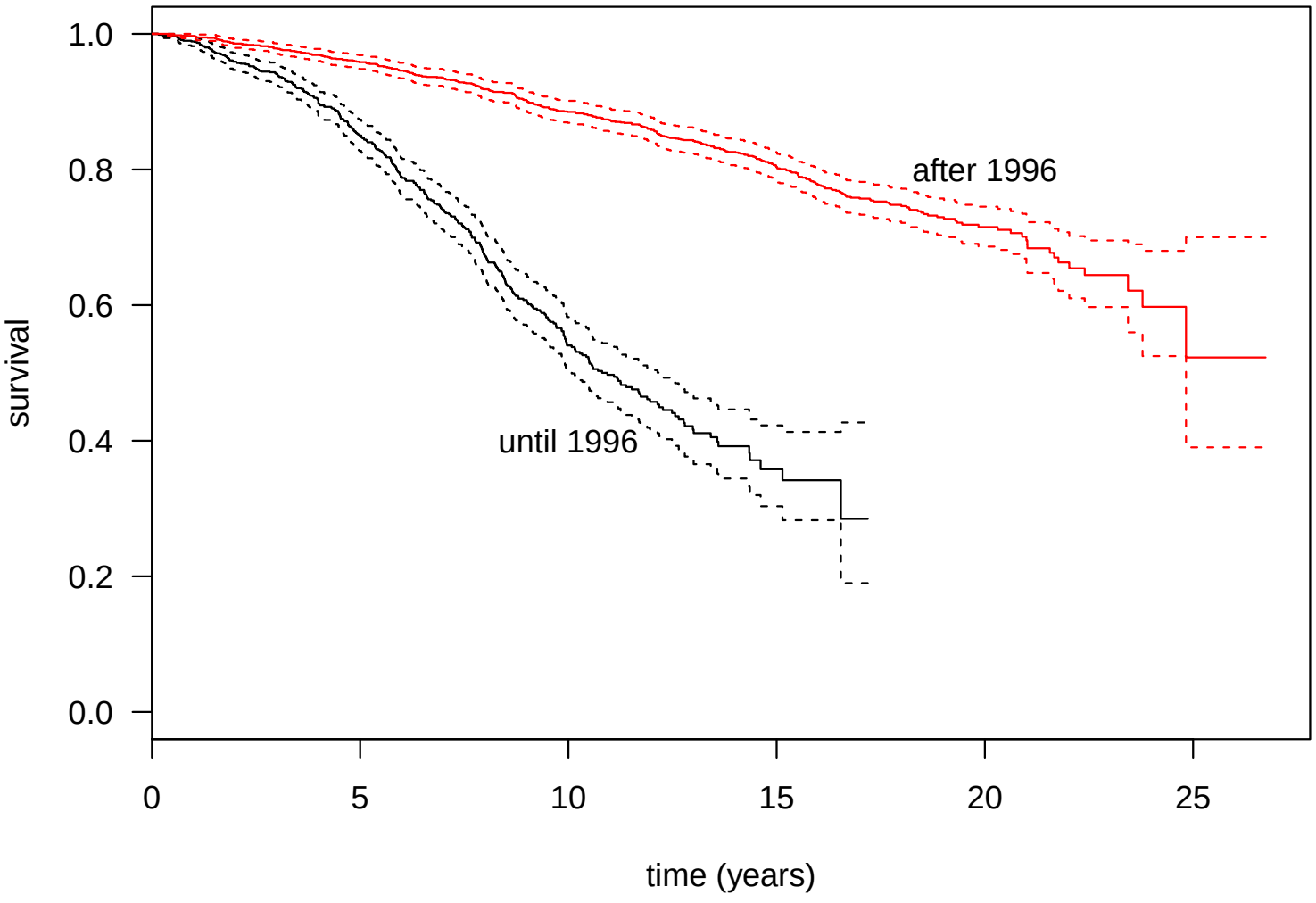
Competing risks, analysis and interpretation

SESSTIM
December 18, 2020

Ronald Geskus
Oxford University Clinical Research Unit (OUCRU)
Hospital for Tropical Diseases, Ho Chi Minh City, Viet Nam

Outline

Example: death after HIV infection



Problem: right censored data $\longrightarrow$ rate/hazard

Rate and Risk

T time to event (e.g. death)

- Risk: $\mathrm{P}(T \leq t)$ (or survival $\mathrm{P}(T > t)$)
- Rate (hazard, incidence):

$$h(t) = \mathrm{P}(T = t | T \geq t) \qquad \textit{discrete}$$

$$\lim_{\Delta t \downarrow 0} \frac{\mathrm{P}(t \leq T < t + \Delta t \,|\, T \geq t)}{\Delta t} \qquad \textit{continuous}$$

Kaplan-Meier

year	0-1	1-2	2-3	3-4	4-5	5-6	6-7	Total
death	1	2	6	11	9	11	2	60
censor	5	9	6	6	9	12	4	86

Kaplan-Meier

year	0-1	1-2	2-3	3-4	4-5	5-6	6-7	Total
death	1	2	6	11	9	11	2	60
censor	5	9	6	6	9	12	4	86

$$\begin{aligned}\mathrm{P}(>\text{6 year alive}) = &\mathrm{P}(\text{ year 0-1 alive})\times \\ &\mathrm{P}(\text{ year 1-2 alive} \mid \text{alive until year 1})\times \\ &\ldots\times \\ &\mathrm{P}(\text{ year 5-6 alive}\mid\text{alive until year 5}) \\ &= (1-h_{0\text{-}1})\times(1-h_{1\text{-}2})\times(1-h_{2\text{-}3})\times\ldots\times(1-h_{5\text{-}6})\end{aligned}$$

Year 0-1: $h_{0\text{-}1} = 1/146 = 0.006849$
Year 1-2: $h_{1\text{-}2} = 2/140 = 0.014286$
Year 2-3: $h_{2\text{-}3} = 6/129 = 0.046512$ etc.

oucru

wellcome

University of Oxford

Kaplan-Meier

year	0-1	1-2	2-3	3-4	4-5	5-6	6-7	Total
death	1	2	6	11	9	11	2	60
censor	5	9	6	6	9	12	4	86

$$\begin{aligned}\mathrm{P}(>\text{6 year alive}) = &\mathrm{P}(\text{ year 0-1 alive})\times \\ &\mathrm{P}(\text{ year 1-2 alive} \mid \text{alive until year 1})\times \\ &\ldots\times \\ &\mathrm{P}(\text{ year 5-6 alive}\mid\text{alive until year 5}) \\ &= (1-h_{0\text{-}1})\times(1-h_{1\text{-}2})\times(1-h_{2\text{-}3})\times\ldots\times(1-h_{5\text{-}6})\end{aligned}$$

Year 0-1: $h_{0\text{-}1} = 1/146 = 0.006849$
Year 1-2: $h_{1\text{-}2} = 2/140 = 0.014286$
Year 2-3: $h_{2\text{-}3} = 6/129 = 0.046512$ etc.

Assumption: censored individuals represented by those at risk

Rate and Risk

T time to event (e.g. death)

- Risk: $\mathrm{P}(T \leq t)$ (or survival $\mathrm{P}(T > t)$)
- Rate (hazard, incidence):

$$h(t) = \mathrm{P}(T = t | T \geq t) \qquad \textit{discrete}$$

$$\lim_{\Delta t \downarrow 0} \frac{\mathrm{P}(t \leq T < t + \Delta t \,|\, T \geq t)}{\Delta t} \qquad \textit{continuous}$$

- One-to-one relation via

$$\mathrm{P}(T > t) = \prod_{t_i \leq t} \{1 - h(t_i)\} \qquad \textit{discrete}$$

$$\mathrm{P}(T > t) = \exp\left\{ - \int_0^t h(u) du \right\} \qquad \textit{continuous}$$

Rate and Risk

T time to event (e.g. death)

- Risk: $\mathrm{P}(T \leq t)$ (or survival $\mathrm{P}(T > t)$)
- Rate (hazard, incidence):

$$h(t) = \mathrm{P}(T = t | T \geq t) \qquad \textit{discrete}$$

$$\lim_{\Delta t \downarrow 0} \frac{\mathrm{P}(t \leq T < t + \Delta t \,|\, T \geq t)}{\Delta t} \qquad \textit{continuous}$$

- One-to-one relation via

$$\mathrm{P}(T > t) = \prod_{t_i \leq t} \{1 - h(t_i)\} \qquad \textit{discrete}$$

$$\mathrm{P}(T > t) = \exp\left\{ - \int_0^t h(u) du \right\} \qquad \textit{continuous}$$

- Hazard basis for many methods
  - Kaplan-Meier

Rate and Risk

T time to event (e.g. death)

- Risk: $\mathrm{P}(T \leq t)$ (or survival $\mathrm{P}(T > t)$)
- Rate (hazard, incidence):

$$h(t) = \mathrm{P}(T = t | T \geq t) \qquad \textit{discrete}$$

$$\lim_{\Delta t \downarrow 0} \frac{\mathrm{P}(t \leq T < t + \Delta t \,|\, T \geq t)}{\Delta t} \qquad \textit{continuous}$$

- One-to-one relation via

$$\mathrm{P}(T > t) = \prod_{t_i \leq t} \{1 - h(t_i)\} \qquad \textit{discrete}$$

$$\mathrm{P}(T > t) = \exp\left\{ - \int_0^t h(u) du \right\} \qquad \textit{continuous}$$

- Hazard basis for many methods
  - Kaplan-Meier
  - Cox model $h(t) = h_0(t) \exp\{\beta_1 X_1 + \ldots + \beta_p X_p\}$

Outline

What and how?

- What is a competing risk?

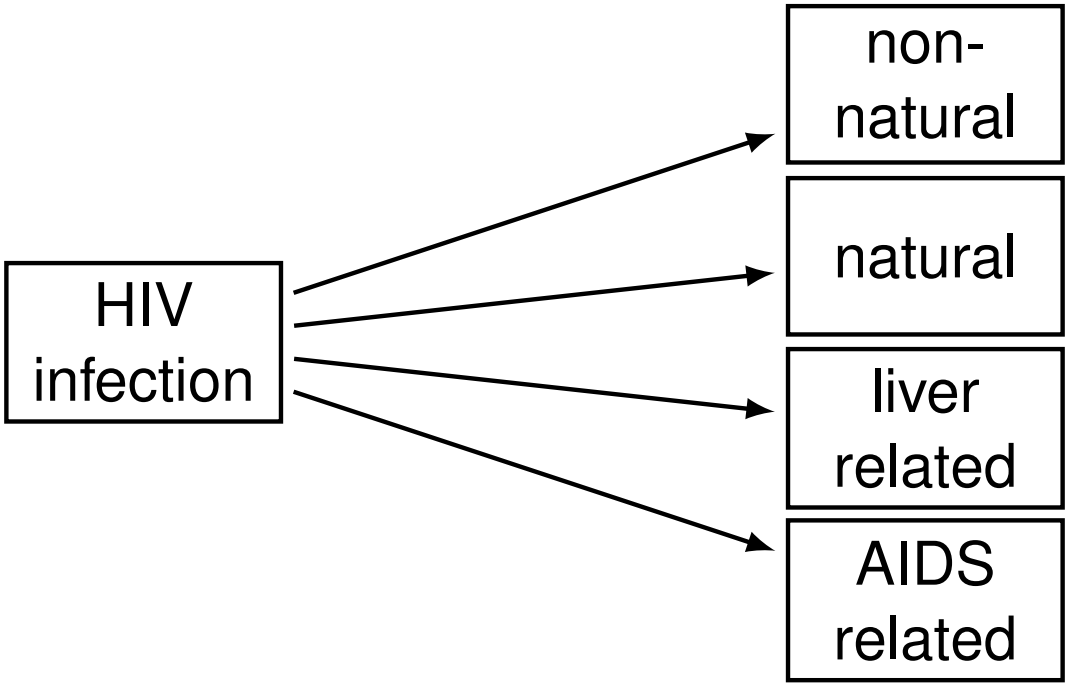
  A competing risk is an event that

  i) prevents the event of interest to occur

  or

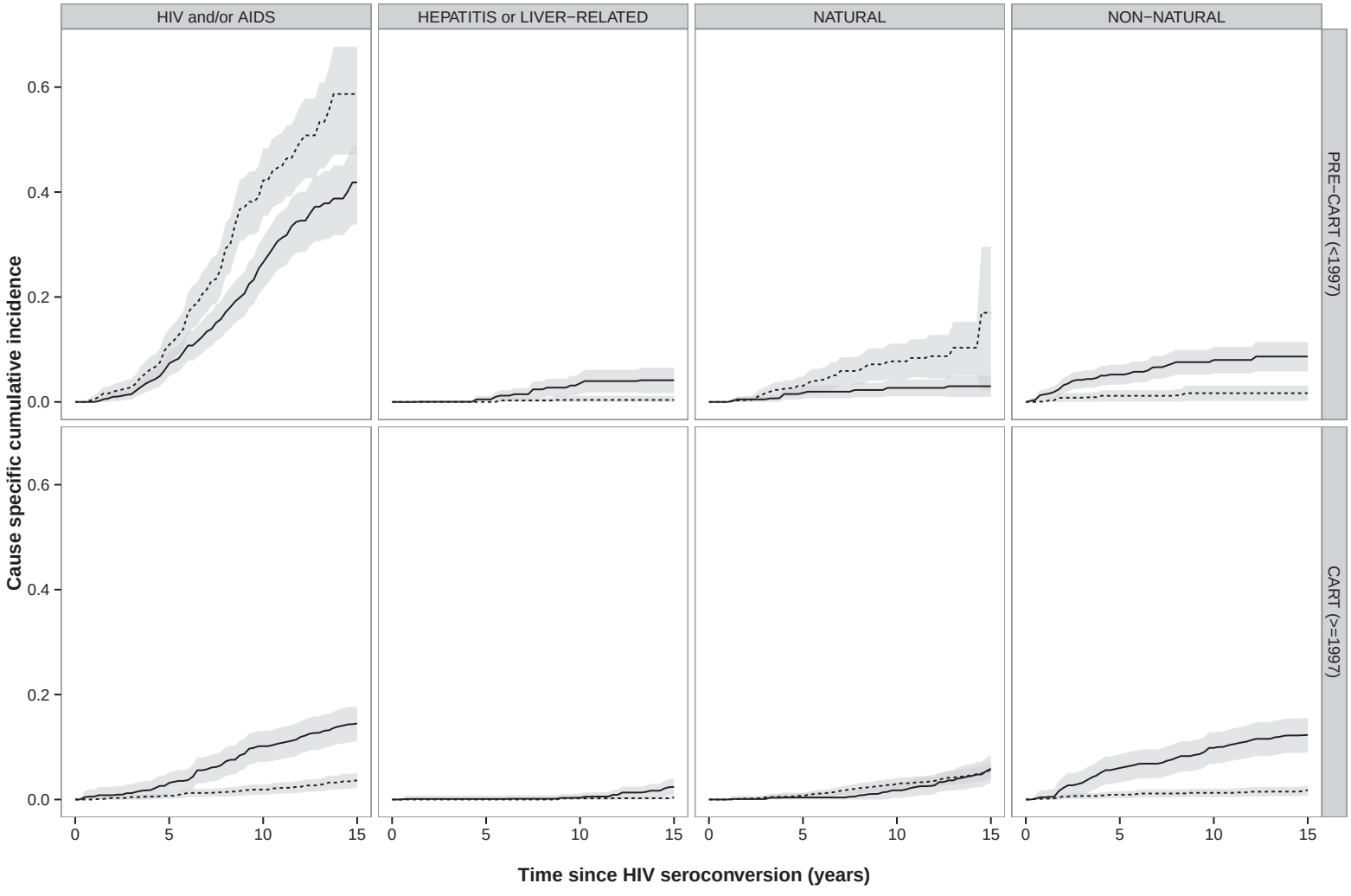
  ii) changes the risk of the event of interest

Example I: Causes of death (COD) after HIV infection



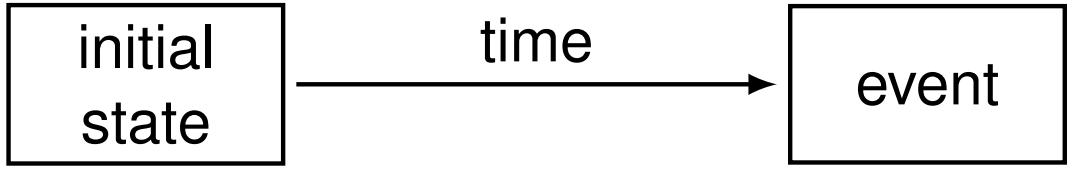
- Has the spectrum in causes of death changed after the introduction of cART (combination anti-retroviral therapy)

Cause-specific mortality by calendar period and hepatitis C status



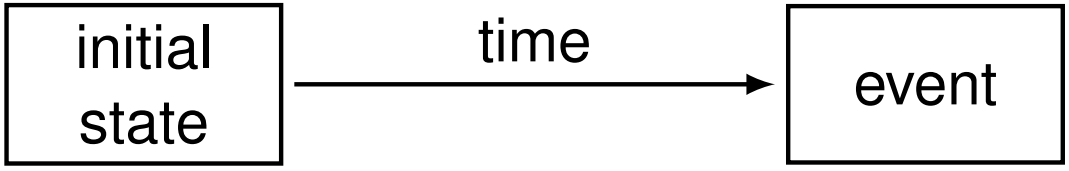
Beyond classical survival analysis

- Classical: transition between two states, one event type.
  "We all die, but not all at the same age"



Beyond classical survival analysis

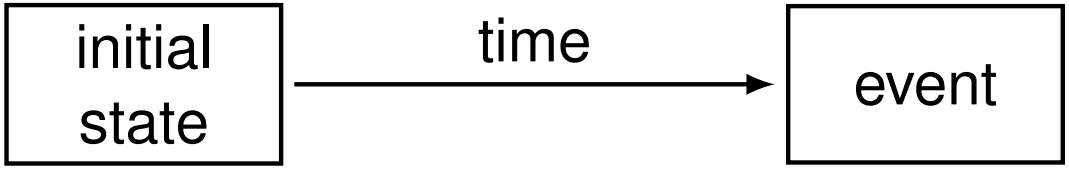
- Classical: transition between two states, one event type.
  "We all die, but not all at the same age"



- Life and death are richer than that
  1. Multiple causes of death. Competing risks:
     "we all die, but not all at the same age and from the same cause"

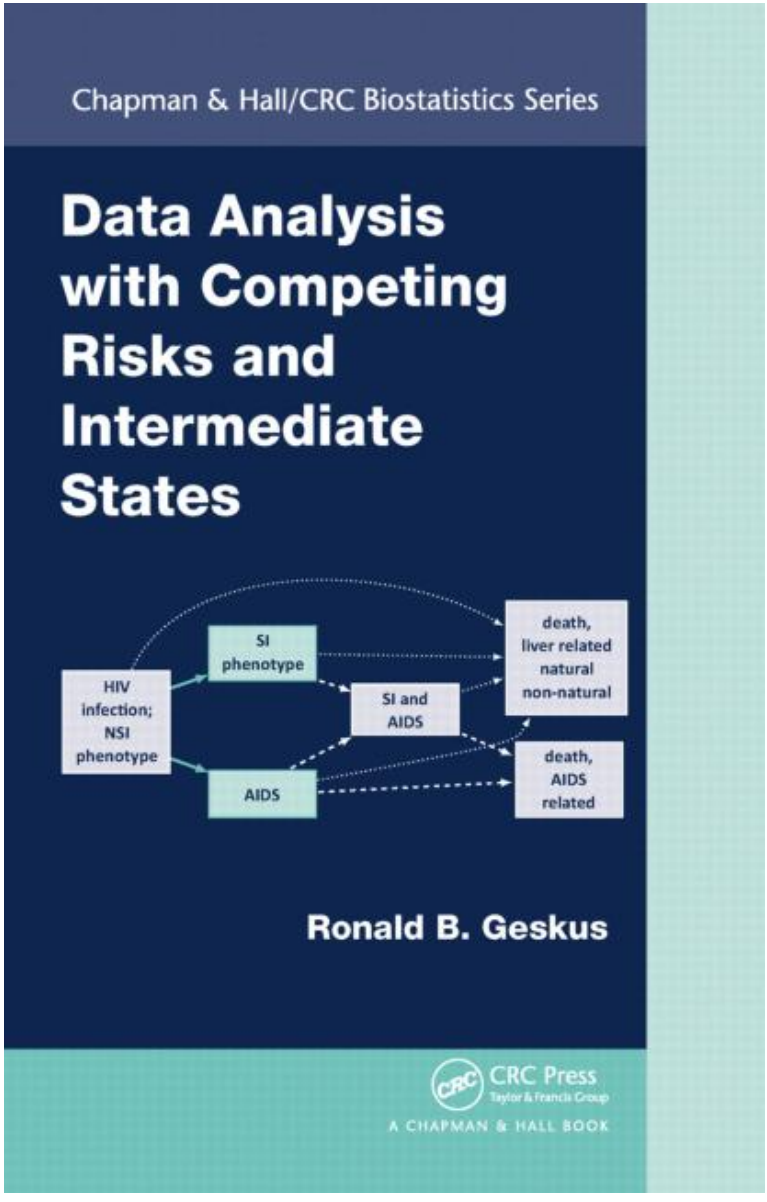
Beyond classical survival analysis

- Classical: transition between two states, one event type.
  "We all die, but not all at the same age"



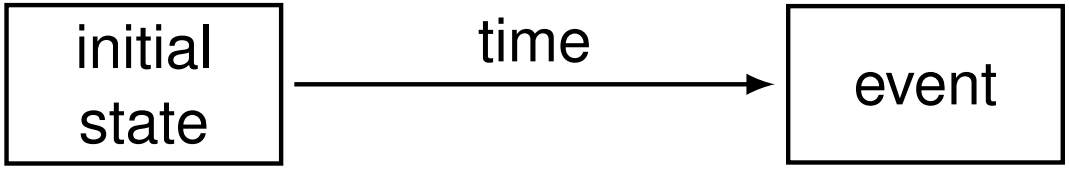
- Life and death are richer than that
  1. Multiple causes of death. Competing risks:
     "we all die, but not all at the same age and from the same cause"
  2. Intermediate events. Multi-state model:
     "we all die, but not all at the same age, not from the same cause and with different life histories"

Published by CRC Press, 2015



Beyond classical survival analysis

- Classical: transition between two states, one event type.
  "We all die, but not all at the same age"



- Life and death are richer than that
  1. Multiple causes of death. Competing risks:
     "we all die, but not all at the same age and from the same cause"
  2. Intermediate events. Multi-state model:
     "we all die, but not all at the same age, not from the same cause and with different life histories"

What and how?

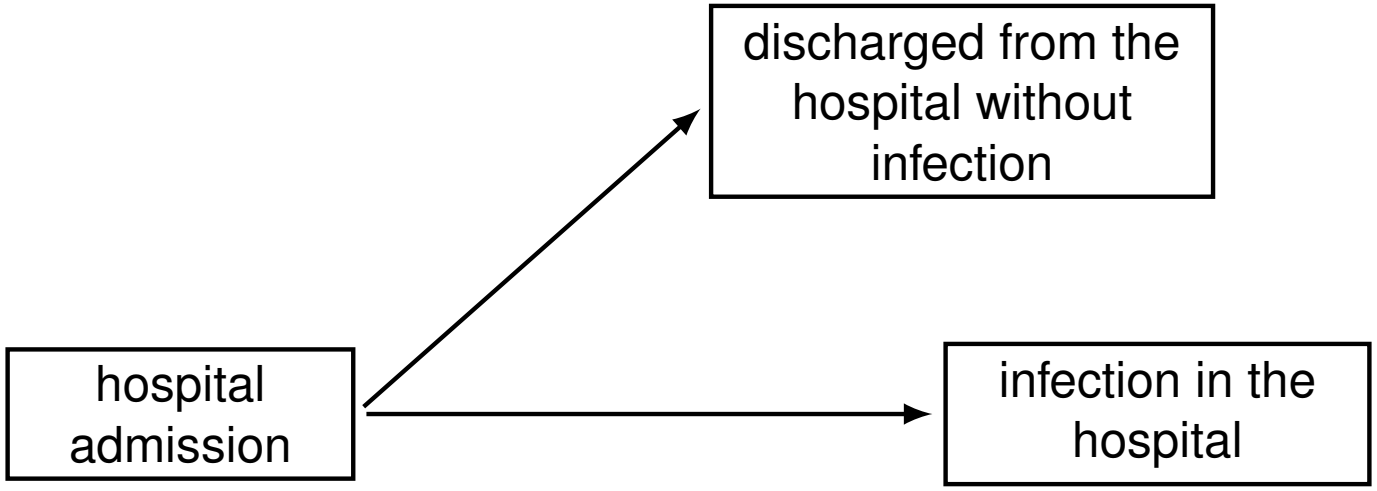
- What is a competing risk?

  A competing risk is an event that
  i) prevents the event of interest to occur
  or
  ii) changes the risk of the event of interest

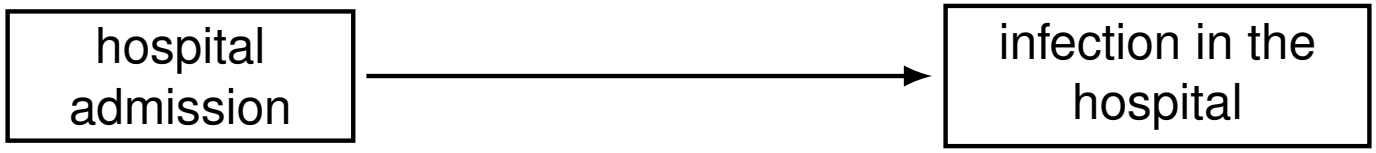
- How do I deal with individuals with competing risk?
  - Marginal analysis: treat as censored and interpret as censored
  - Competing risks, multi-state approach: treat as censored and interpret as competing risk
  - Competing risks, subdistribution approach: ignore occurrence of competing risk in estimation (keep in risk set "forever")

Outline

II: Time to staphylococcus infection during hospital stay

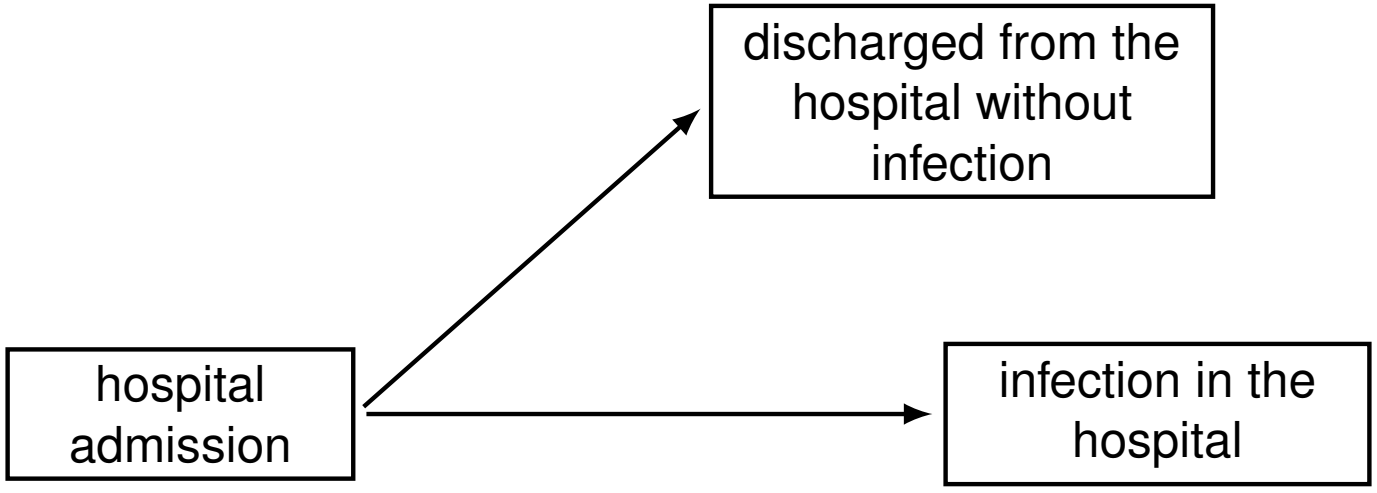


II: Time to staphylococcus infection during hospital stay



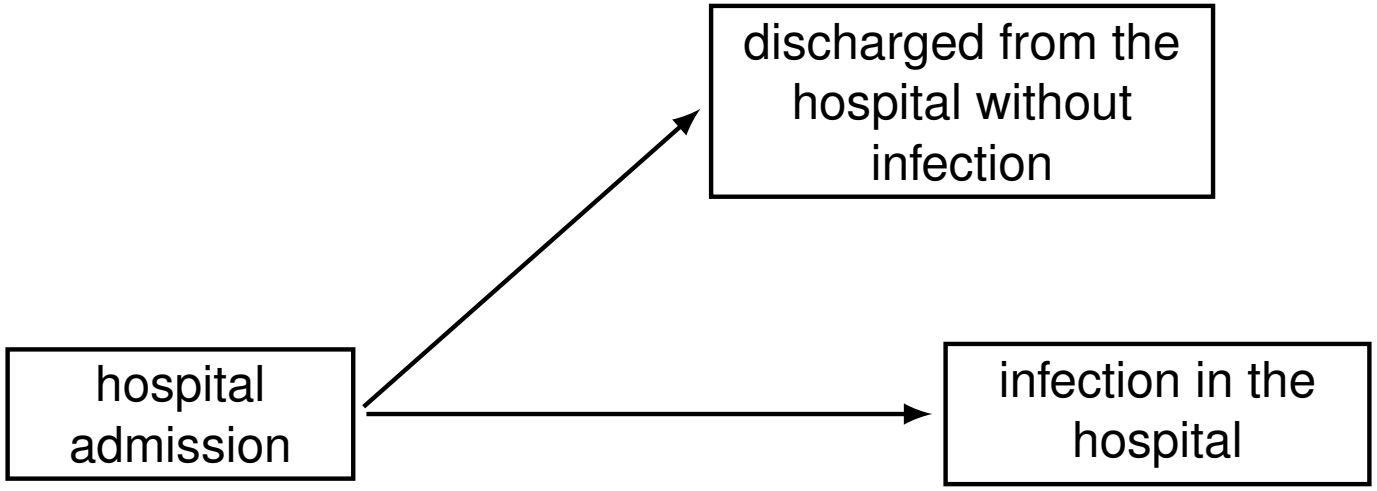
- Etiology (biological question): infection risk in hospital. What would happen if everyone stayed in hospital?
  - Marginal distribution/net risk

II: Time to staphylococcus infection during hospital stay



- Etiology (biological question): infection risk in hospital. What would happen if everyone stayed in hospital?
  - Marginal distribution/net risk
- Predict (clinical question): burden due to infection while in hospital; discharge prevents event to occur
  - Cause-specific cumulative incidence/crude risk

II: Time to staphylococcus infection during hospital stay



- Etiology (biological question): infection risk in hospital. What would happen if everyone stayed in hospital?
  - Marginal distribution/net risk
- Predict (clinical question): burden due to infection while in hospital; discharge prevents event to occur
  - Cause-specific cumulative incidence/crude risk

Estimation with complete follow-up (artificial data)

week	0-1	1-2	2-3	3-4	4-5	5-6	6-7	> 7
infection	1	2	6	11	9	11	2	18
cumulative	1	3	9	20	29	40	42	60
discharge	5	9	6	6	9	12	4	35
cumulative	5	14	20	26	35	47	51	86

Estimation with complete follow-up (artificial data)

week	0-1	1-2	2-3	3-4	4-5	5-6	6-7	> 7
infection	1	2	6	11	9	11	2	18
cumulative	1	3	9	20	29	40	42	60
discharge	5	9	6	6	9	12	4	35
cumulative	5	14	20	26	35	47	51	86

- Crude risk. Estimated as frequency of events:

  $\widehat{P}(\text{infection} \leq 6\,\text{weeks}) = 40/146$

  $\widehat{P}(\text{discharge} \leq 6\,\text{weeks}) = 47/146$

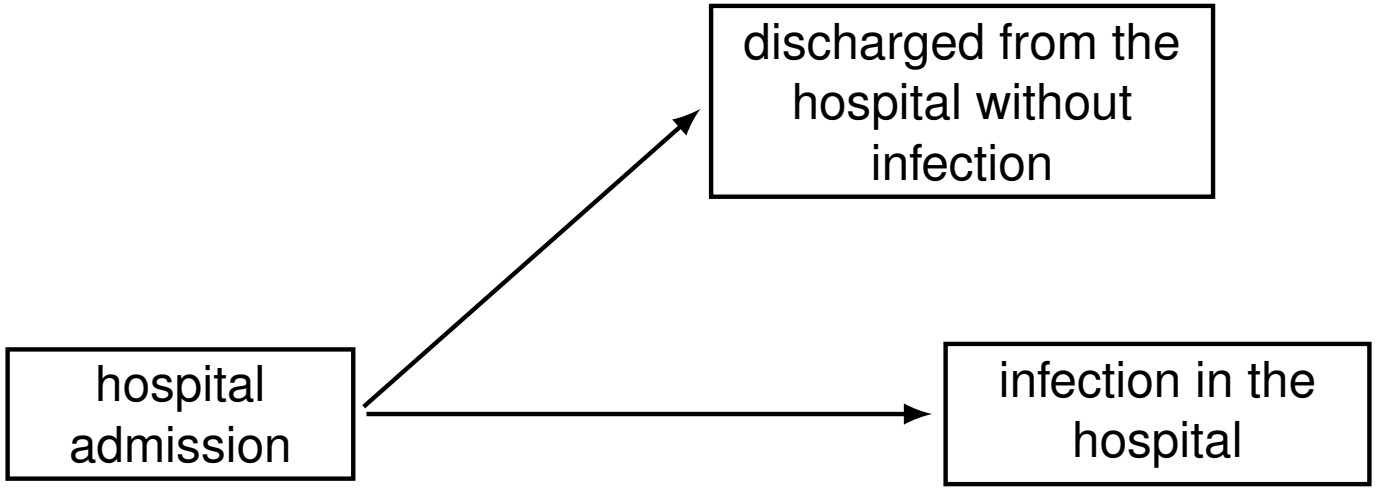
  Individuals with competing event remain in denominator, competing event ignored in estimation

Estimation with complete follow-up (artificial data)

week	0-1	1-2	2-3	3-4	4-5	5-6	6-7	> 7
infection	1	2	6	11	9	11	2	18
cumulative	1	3	9	20	29	40	42	60
discharge	5	9	6	6	9	12	4	35
cumulative	5	14	20	26	35	47	51	86

- Crude risk. Estimated as frequency of events:
  $\widehat{P}(\text{infection} \leq 6\,\text{weeks}) = 40/146$
  $\widehat{P}(\text{discharge} \leq 6\,\text{weeks}) = 47/146$
  Individuals with competing event remain in denominator, competing event ignored in estimation
- Net risk. Discharged individuals treated and interpreted as censored.

II: Time to staphylococcus infection during hospital stay



- Etiology (biological question): infection risk in hospital. What would happen if everyone stayed in hospital?
  - Marginal distribution/net risk
- Predict (clinical question): burden due to infection while in hospital; discharge prevents event to occur
  - Cause-specific cumulative incidence/crude risk
- Which of two hospitals has higher risk may depend on type of question

What and how?

- What is a competing risk?

  A competing risk is an event that
  i) prevents the event of interest to occur
  or
  ii) changes the risk of the event of interest

- How do I deal with individuals with competing risk?
  - Marginal analysis: treat as censored and interpret as censored
  - Competing risks, multi-state approach: treat as censored and interpret as competing risk
  - Competing risks, subdistribution approach: ignore occurrence of competing risk in estimation (keep in risk set "forever")

What and how?

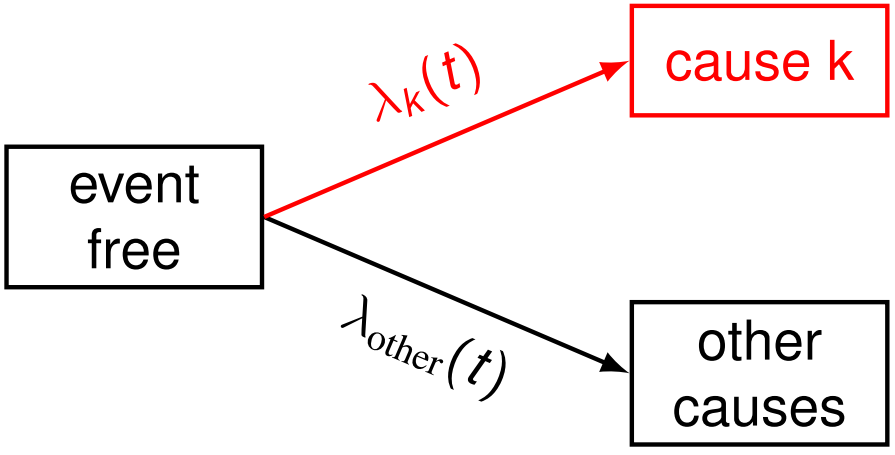
- What is a competing risk?

  A competing risk is an event that
  i) prevents the event of interest to occur
  or
  ii) changes the risk of the event of interest

- How do I deal with individuals with competing risk?
  - Marginal analysis: treat as censored and interpret as censored
  - Competing risks, multi-state approach: treat as censored and interpret as competing risk
  - Competing risks, subdistribution approach: ignore occurrence of competing risk in estimation (keep in risk set "forever")

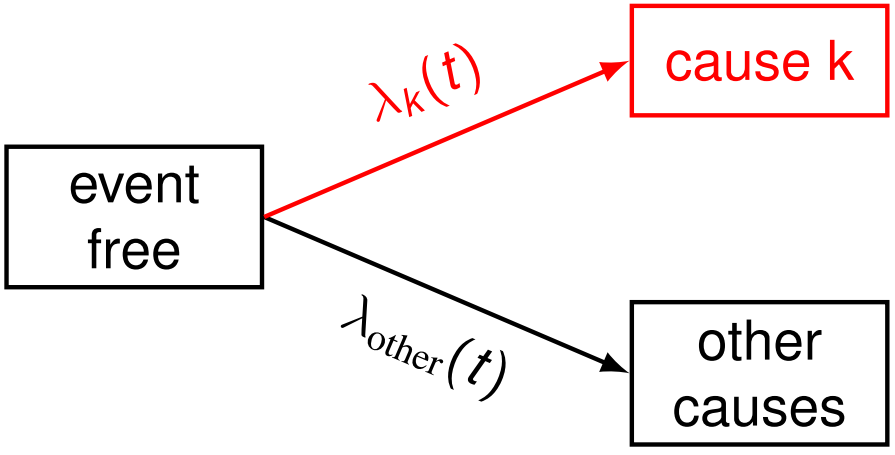
Outline

I: The multi-state approach: cause-specific hazard



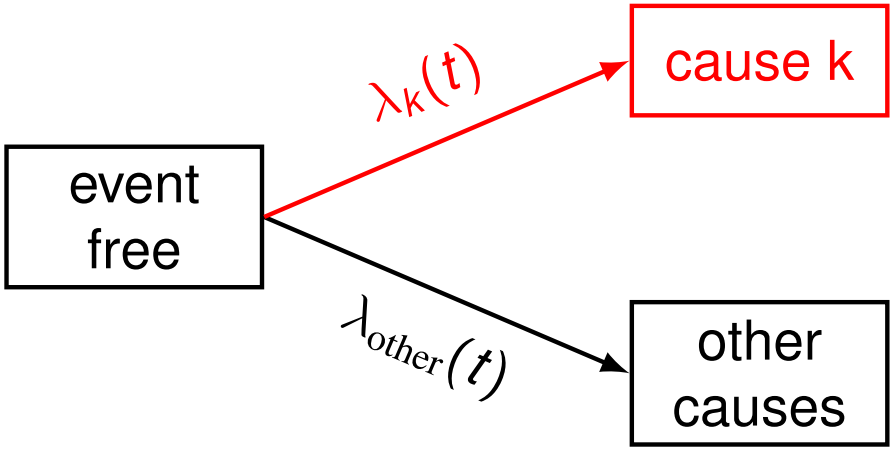
- Transition rate to cause k:
  $\lambda_k(t) = \mathrm{P}(T = t, E = k \mid T \geq t) = \frac{\mathrm{P}(T=t, E=k)}{\mathrm{P}(T \geq t)}$ *(discrete)*

I: The multi-state approach: cause-specific hazard



- Transition rate to cause k:
  $\lambda_k(t) = \mathrm{P}(T = t, E = k \mid T \geq t) = \frac{\mathrm{P}(T=t, E=k)}{\mathrm{P}(T \geq t)}$ *(discrete)*
- Estimate $\widehat{\lambda_k}(t)$: those with competing event leave risk set
  $\Longrightarrow$ treated as censored; interpretation as competing risk

I: The multi-state approach: cause-specific hazard



- Transition rate to cause k:
  $\lambda_k(t) = \mathrm{P}(T = t, E = k \mid T \geq t) = \frac{\mathrm{P}(T=t, E=k)}{\mathrm{P}(T \geq t)}$ *(discrete)*
- Estimate $\widehat{\lambda_k}(t)$: those with competing event leave risk set
  $\Longrightarrow$ treated as censored; interpretation as competing risk
- Risk $\mathrm{P}(T \leq t, E = k)$: Aalen-Johansen estimator;
  combines $\widehat{\lambda_k}(t)$ with Kaplan-Meier based on overall hazard
  $h(t) = \sum_{e=1}^{K} \lambda_e(t)$ (e.g. $K = 4$)

Outline

Estimation with complete follow-up (artificial data)

week	0-1	1-2	2-3	3-4	4-5	5-6	6-7	>7
infection	1	2	6	11	9	11	2	18
cumulative	1	3	9	20	29	40	42	60
discharge	5	9	6	6	9	12	4	35
cumulative	5	14	20	26	35	47	51	86

- Crude risk. Estimated as frequency of events:

  $\widehat{P}(\text{infection} \leq 6\,\text{weeks}) = 40/146$

  $\widehat{P}(\text{discharge} \leq 6\,\text{weeks}) = 47/146$

  Individuals with competing event remain in denominator, competing event ignored in estimation

II: The subdistribution approach

T_k^* time to event of type k: $T_k^* = T \times I\{E = k\} + \infty \times I\{E \neq k\}$

- $\mathrm{P}(T_k^* \leq t) = \mathrm{P}(T \leq t, E = k)$

II: The subdistribution approach

T_k^* time to event of type k: $T_k^* = T \times I\{E = k\} + \infty \times I\{E \neq k\}$

- $\mathrm{P}(T_k^* \leq t) = \mathrm{P}(T \leq t, E = k)$
- Subdistribution hazard:

$$\begin{aligned} h_k(t) &= \mathrm{P}\{T_k^* = t \mid T_k^* \geq t\} \\ &= \mathrm{P}\{T = t, E = k \mid T \geq t \text{ or } (T < t, E \neq k)\} \end{aligned}$$

Denominator: event free or with earlier competing event
Individuals with competing event remain in risk set "forever"

II: The subdistribution approach

T_k^* time to event of type k: $T_k^* = T \times I\{E = k\} + \infty \times I\{E \neq k\}$

- $\mathrm{P}(T_k^* \leq t) = \mathrm{P}(T \leq t, E = k)$
- Subdistribution hazard:

$$\begin{aligned} h_k(t) &= \mathrm{P}\{T_k^* = t \mid T_k^* \geq t\} \\ &= \mathrm{P}\{T = t, E = k \mid T \geq t \text{ or } (T < t, E \neq k)\} \end{aligned}$$

Denominator: event free or with earlier competing event
Individuals with competing event remain in risk set "forever"

- One-to-one relation with crude risk

$$\widehat{\mathrm{P}}(T_k^* > t) = \prod_{t_i \leq t} \left\{ 1 - \widehat{h_k}(t_i) \right\}$$

Outline

Regression on hazard

- Cause-specific hazard
  - Standard Cox model
  - Interpretation: cause-specific event rate among event-free individuals
  - Not a marginal hazard, unless progression to competing risks unrelated

Regression on hazard

- Cause-specific hazard
  - Standard Cox model
  - Interpretation: cause-specific event rate among event-free individuals
  - Not a marginal hazard, unless progression to competing risks unrelated
- Subdistribution hazard
  - Fine and Gray model: proportional subdistribution hazards
  - Interpretation: direct relation with cause-specific cumulative incidence

Outline

Rates and risks in competing risks setting

	hazard		cumulative
competing risks	marginal		net risk marginal survival function marginal cumulative incidence
	cause-specific	λ_k	no corresponding quantity
	subdistribution	h_k	crude risk $\mathrm{P}(T_k^* \leq t)$ cause-specific cumulative incidence
combined	overall	h	overall risk $\mathrm{P}(T \leq t)$ overall survival function overall cumulative incidence

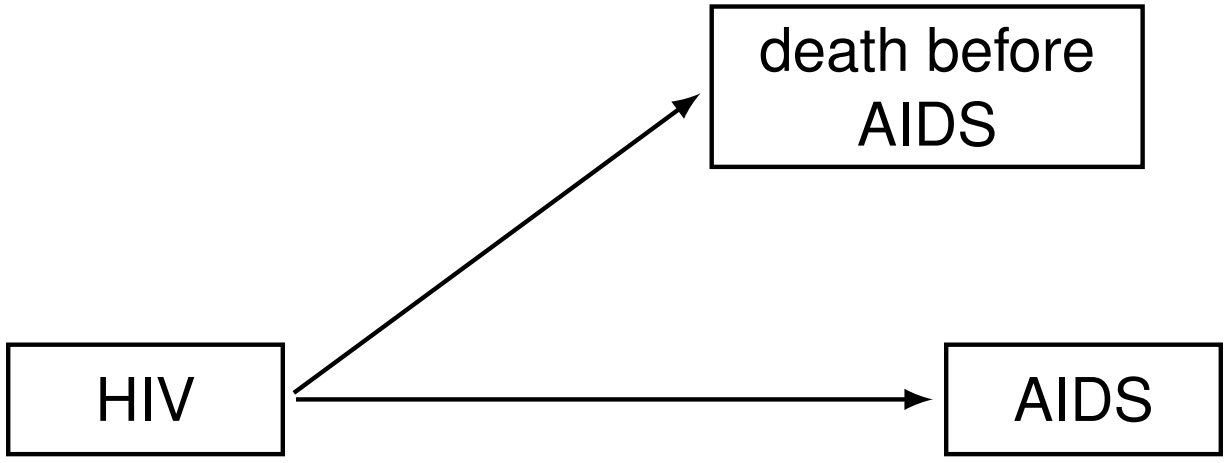
Marginal distribution

- Hypothetical world in which competing events are absent
- Standard hazard estimate. Basis for Kaplan-Meier estimate of cumulative incidence/net risk

Marginal distribution

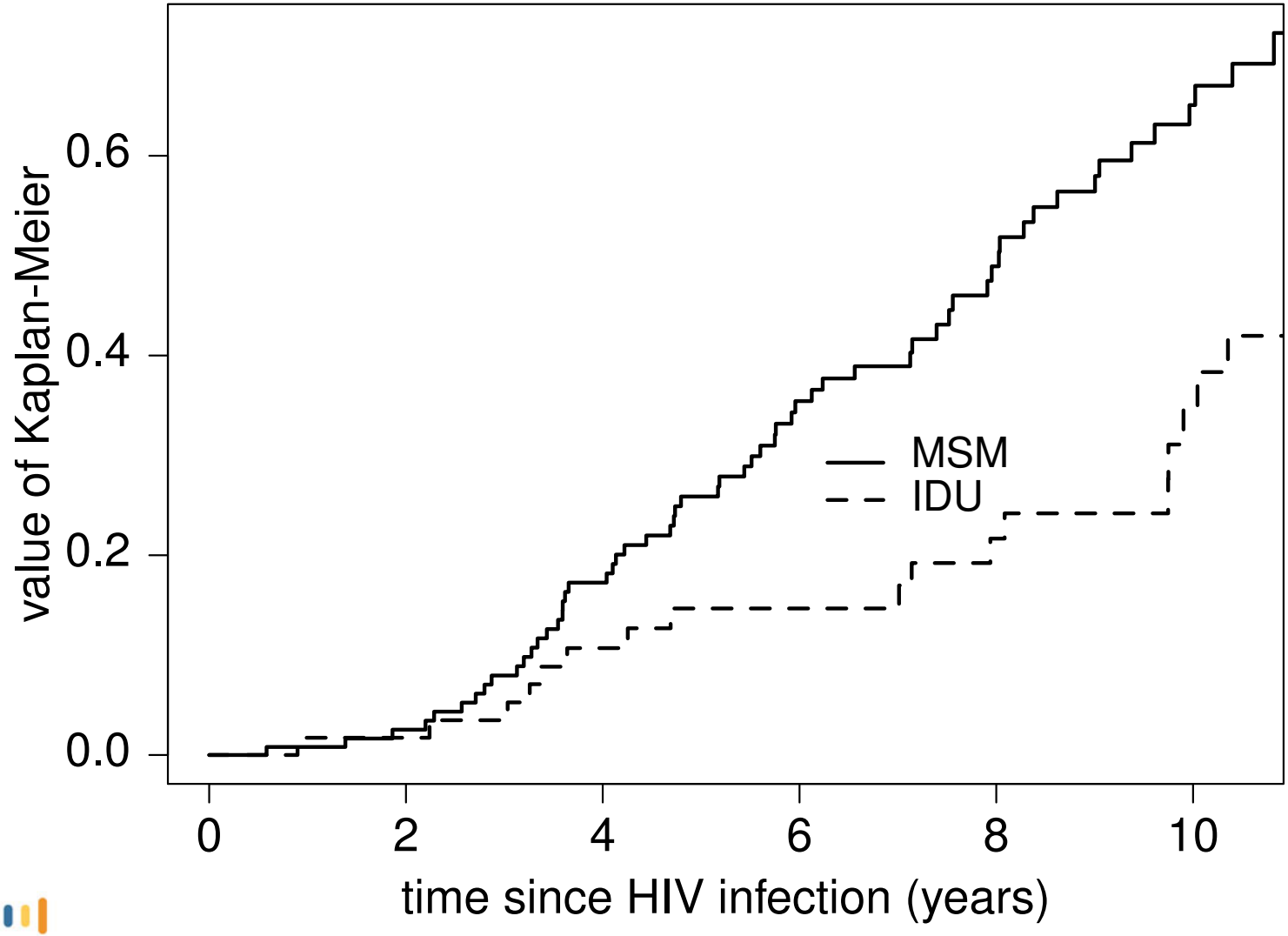
- Hypothetical world in which competing events are absent
- Standard hazard estimate. Basis for Kaplan-Meier estimate of cumulative incidence/net risk
- Assumption: Individuals that are censored can be represented by the ones that remain at risk. Mechanisms unrelated
- Otherwise: Kaplan-Meier has no meaning.
- Extra information may allow to show informative/dependent censoring, but independence can never be tested for

III: Time from HIV infection to AIDS



- Compare men who have sex with men (MSM) and injecting drug users (IDU)
- Interest in time to AIDS if there were no pre-AIDS death. Interest in etiology and marginal distribution
- Kaplan-Meier: censor at death (before AIDS)
  Assumption: those alive without AIDS represent pre-AIDS deaths

Kaplan-Meier: IDU much slower progression ($p = 0.001$)



Explanation: dependent censoring

- Information on cause of death before AIDS

	IDU	MSM
Reason of death	Number	
HIV related infections	3	0
overdose/suicide	6	0
violence/accident	2	0
liver cirrhosis	2	0
cancer	0	1
heart attack	0	1
unknown	4	3

- Some causes of pre-AIDS death in IDU related to AIDS progression, i.e. close to AIDS. Net risk estimate for IDU biased downwards

Rates and risks in competing risks setting

	hazard		cumulative	
competing risks	marginal		net risk	
			marginal survival function	
			marginal cumulative incidence	
	cause-specific	λ_k	no corresponding quantity	
	subdistribution	h_k	crude risk	$\mathrm{P}(T_k^* \leq t)$
			cause-specific cumulative incidence	
combined	overall	h	overall risk	$\mathrm{P}(T \leq t)$
			overall survival function	
			overall cumulative incidence	

Competing risks

- Competing risk is seen as separate event

Competing risks

- Competing risk is seen as separate event
- Individuals censored by competing event don't have to be represented by the ones that remain at risk.
  Other censoring must be independent

Competing risks

- Competing risk is seen as separate event
- Individuals censored by competing event don't have to be represented by the ones that remain at risk.
  Other censoring must be independent
- Cause-specific hazard
  - Basis for Aalen-Johansen estimator of cause-specific cumulative incidence/crude risk

Competing risks

- Competing risk is seen as separate event
- Individuals censored by competing event don't have to be represented by the ones that remain at risk.
  Other censoring must be independent
- Cause-specific hazard
  - Basis for Aalen-Johansen estimator of cause-specific cumulative incidence/crude risk
  - If censoring due to competing event is independent, then marginal and cause-specific hazard are equal. Cumulative quantities different: Kaplan-Meier versus Aalen-Johansen

Competing risks

- Competing risk is seen as separate event
- Individuals censored by competing event don't have to be represented by the ones that remain at risk.
  Other censoring must be independent
- Cause-specific hazard
  - Basis for Aalen-Johansen estimator of cause-specific cumulative incidence/crude risk
  - If censoring due to competing event is independent, then marginal and cause-specific hazard are equal. Cumulative quantities different: Kaplan-Meier versus Aalen-Johansen
- Subdistribution hazard: one-to-one relation with crude risk

Marginal or competing risks?

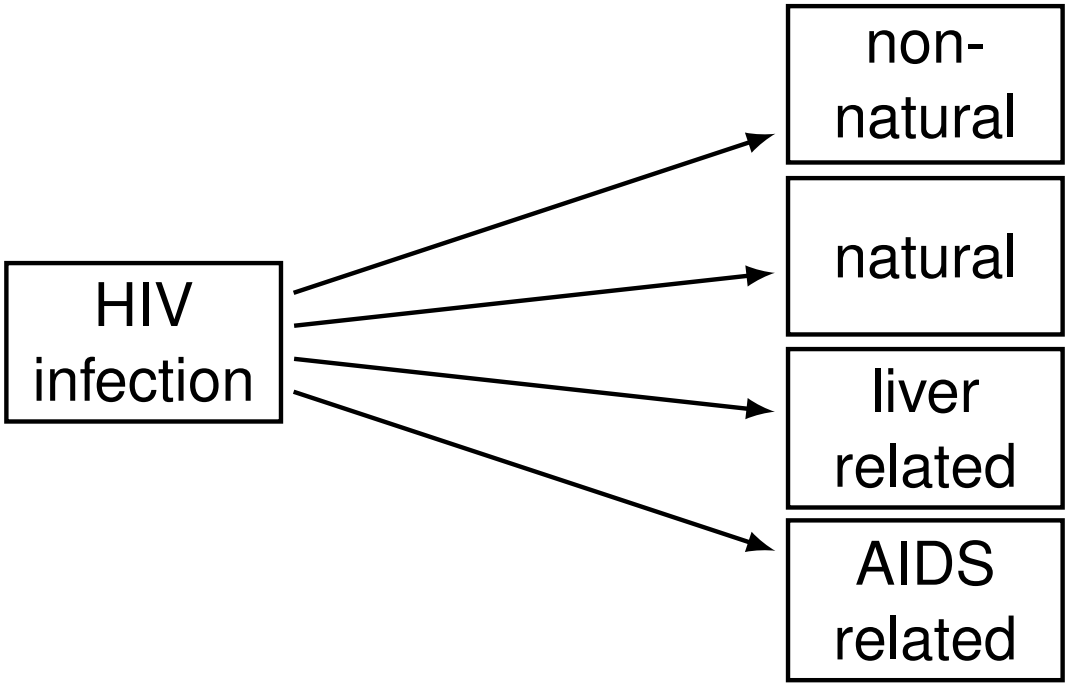
- Example I: spectrum in COD
  Competing risks; marginal analysis completely hypothetical
  No interest in change in AIDS-related death in world in which other COD's do not exist
- Example II: staphylococcus infection in hospital
  - Marginal: what if everyone would stay in hospital
  - Competing risks: how many infections are observed in hospital
- Example III: difference in natural history between IDU en MSM
  Marginal analysis (but not feasible with ACS data)

Outline

Which rate or risk?

- Both Cox model (cause-specific hazard) and Fine and Gray model (subdistribution hazard) make sense in presence of competing risks

Example I: Causes of death (COD) after HIV infection



- Has the spectrum in causes of death changed after the introduction of cART (combination anti-retroviral therapy)
- Different hazards can be chosen

Which rate or risk?

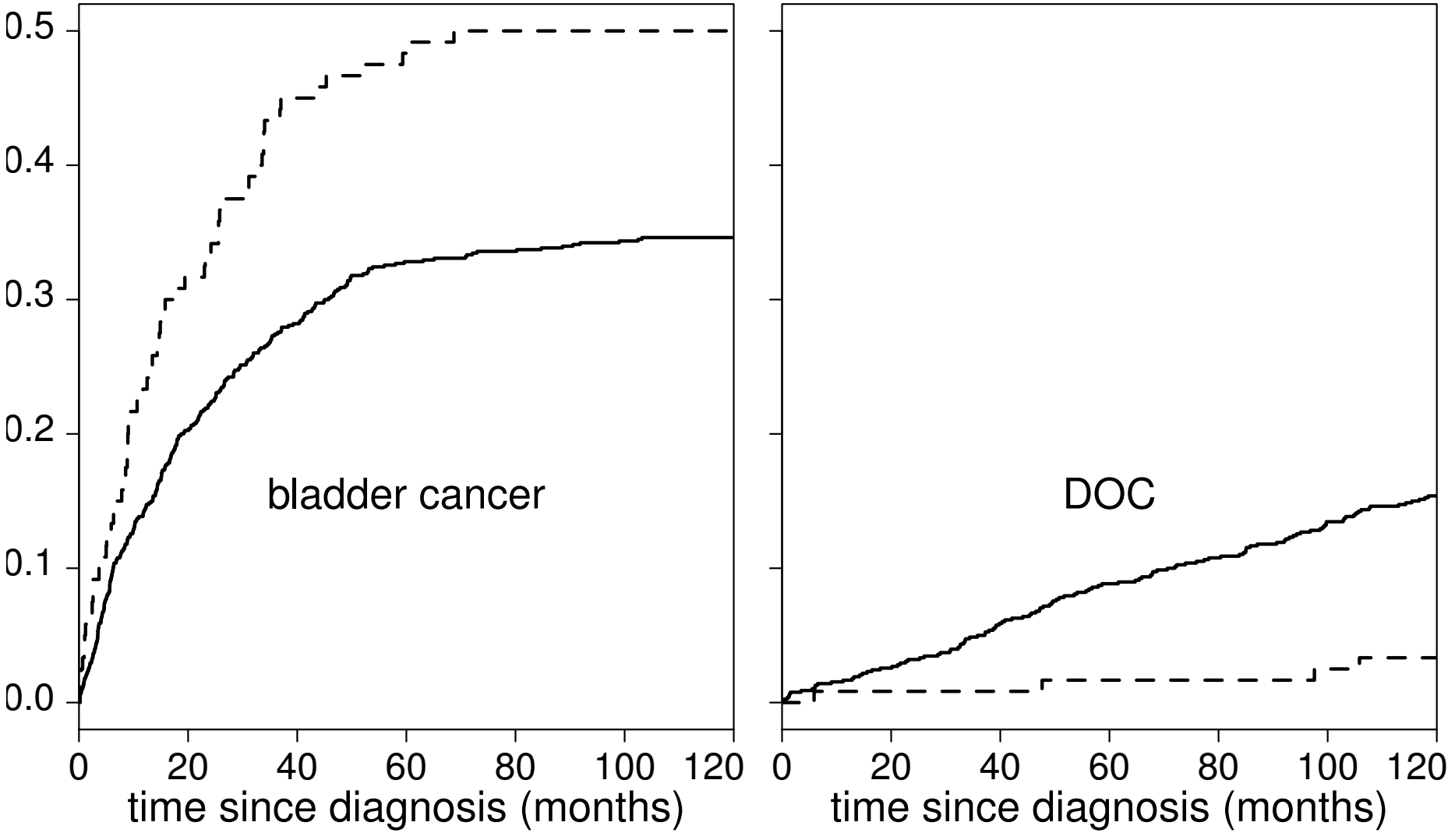
- Both Cox model (cause-specific hazard) and Fine and Gray model (subdistribution hazard) make sense in presence of competing risks
- Example
  - AIDS-specific mortality reduced by cART
  - Other COD's: more frequent, even if cART has no side effects. No change in cause-specific hazard, but subdistribution hazard increases ("in the end we all die")
  - Subdistribution hazard includes impact on other event types

Which rate or risk?

- Both Cox model (cause-specific hazard) and Fine and Gray model (subdistribution hazard) make sense in presence of competing risks
- Example
  - AIDS-specific mortality reduced by cART
  - Other COD's: more frequent, even if cART has no side effects. No change in cause-specific hazard, but subdistribution hazard increases ("in the end we all die")
  - Subdistribution hazard includes impact on other event types
- Prediction: crude risk
  Based on regression model for cause-specific or subdistribution hazard, but only latter has one-to-one relation with cause-specific cumulative incidence/crude risk

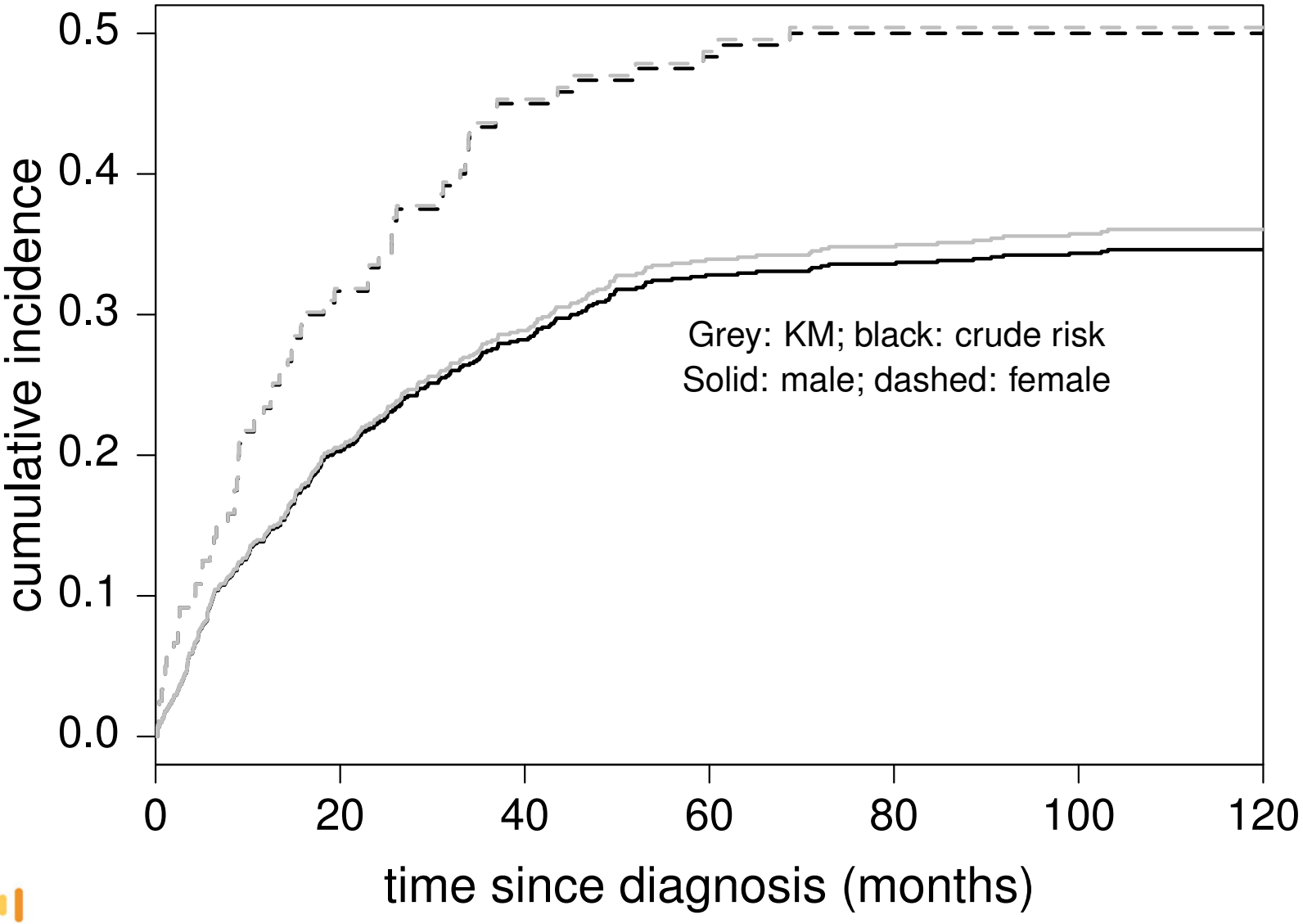
Outline

Bladder cancer; relapse, DOC competing

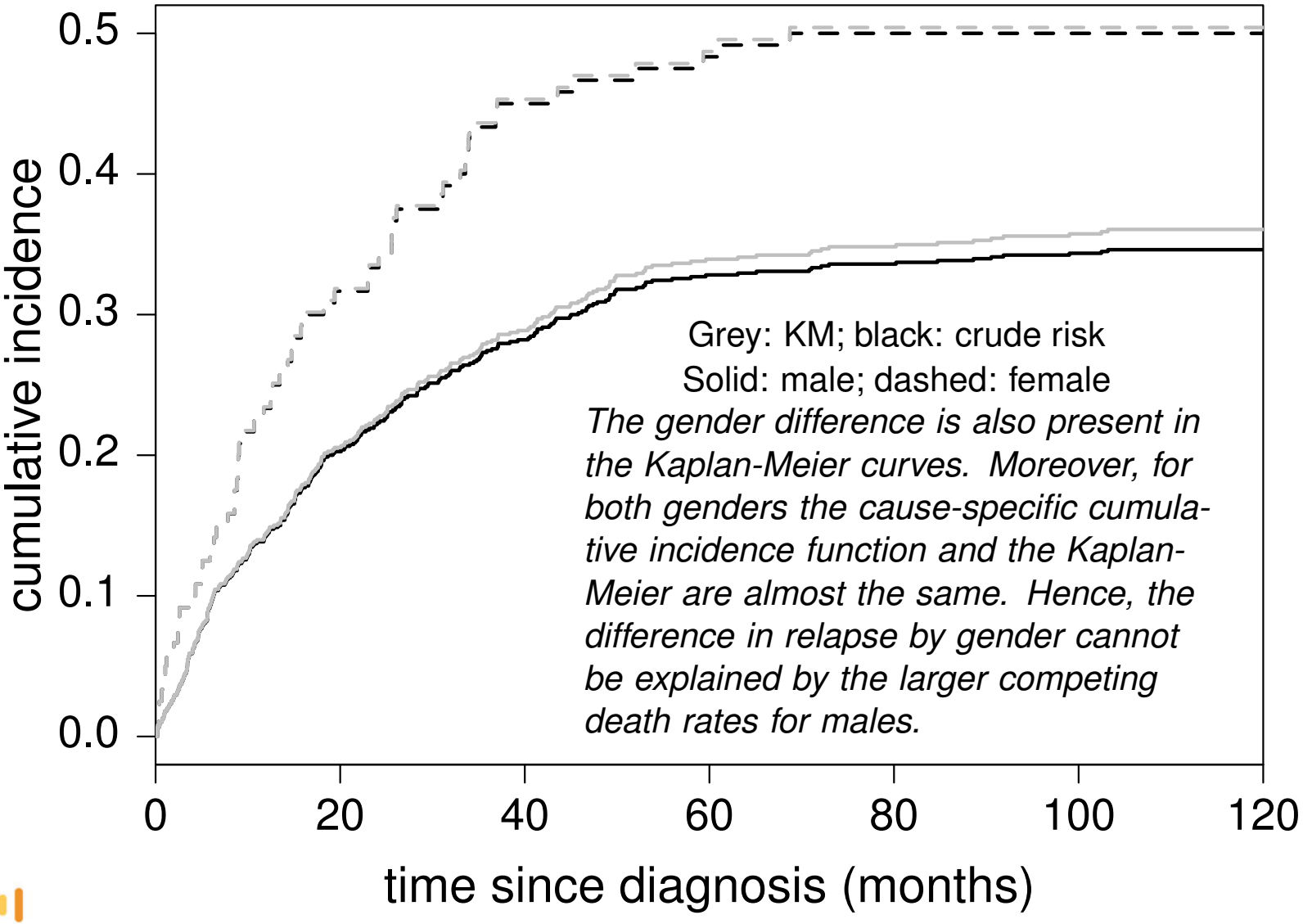


Left: relapse. Solid: male; dashed: female

Bladder cancer; relapse, DOC competing



Bladder cancer; relapse, DOC competing



Answer

- The Kaplan-Meier tries to compare the marginal distribution of time to relapse for males and females. Only valid if DOC is noninformative for relapse.
- Estimates almost equal because there is little mortality due to other causes, at least during the first 40 months.

Answer

- The Kaplan-Meier tries to compare the marginal distribution of time to relapse for males and females. Only valid if DOC is noninformative for relapse.
- Estimates almost equal because there is little mortality due to other causes, at least during the first 40 months.
- If we combine both event times, the curves for males and females will become similar. Would estimate marginal hazard if every person that died would have progressed on the next day.

Answer

- The Kaplan-Meier tries to compare the marginal distribution of time to relapse for males and females. Only valid if DOC is noninformative for relapse.
- Estimates almost equal because there is little mortality due to other causes, at least during the first 40 months.
- If we combine both event times, the curves for males and females will become similar. Would estimate marginal hazard if every person that died would have progressed on the next day.
- All we can conclude is that females have a higher relapse-specific cumulative incidence than males. And females have a lower DOC-specific incidence than males.

THANKS!

References I


J. P. Fine and R. J. Gray (1999).
A proportional hazards model for the subdistribution of a competing risk.
J. Am. Statist. Assoc. **94,** 496–509.


R. B. Geskus (2011).
Cause-specific cumulative incidence estimation and the Fine and Gray model under both left truncation and right censoring.
Biometrics **67**, 39–49


Ronald B. Geskus (2015).
Data analysis with competing risks and intermediate states.
CRC Press, Boca Raton.


H. Putter, M. Fiocco, and R. B. Geskus (2007).
Tutorial in biostatistics: competing risks and multi-state models.
Stat Med, 26(11):2389–2430.


J. van der Helm, R. B. Geskus, C. Sabin *et al.* (2013).
Effect of HCV infection on cause-specific mortality after HIV seroconversion, before and after 1997.
Gastroenterology, 144:751–760, 2013.