# Competing risks, when and how to incorporate them in the analysis

Ronald Geskus

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

October 12, 2023

oucru

wellcome

UNIVERSITY OF OXFORD

# Where I work

- Head of biostatistics group at OUCRU
  (Oxford University Clinical Research Unit)
  with units in Vietnam, Nepal and Indonesia
- Established in 1991 in Ho Chi Minh City, connected to Hospital for Tropical Diseases, founded in 1862

# Where I work

- Head of biostatistics group at OUCRU
  (Oxford University Clinical Research Unit)
  with units in Vietnam, Nepal and Indonesia
- Established in 1991 in Ho Chi Minh City, connected to Hospital for Tropical Diseases, founded in 1862
- Research institute on infectious diseases
  - Malaria: $<$10/3000 deaths per year in Vietnam/Indonesia
  - Tuberculosis: 17,000/100,000 deaths per year in Vietnam/Indonesia
  - Dengue: $<$50/500 deaths in Vietnam/Indonesia
  - Tetanus, Diphteria, Measles, Hand Foot Mouth Disease, ...
  - Drug resistant infections: $>$40,000/$>$100,000 deaths per year in Vietnam/Indonesia

# Outline

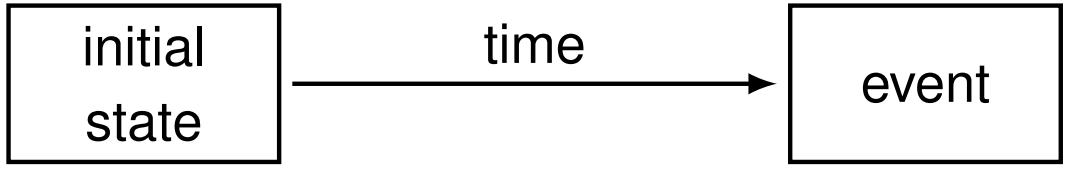
**Competing Risks; When**

Competing Risks; How

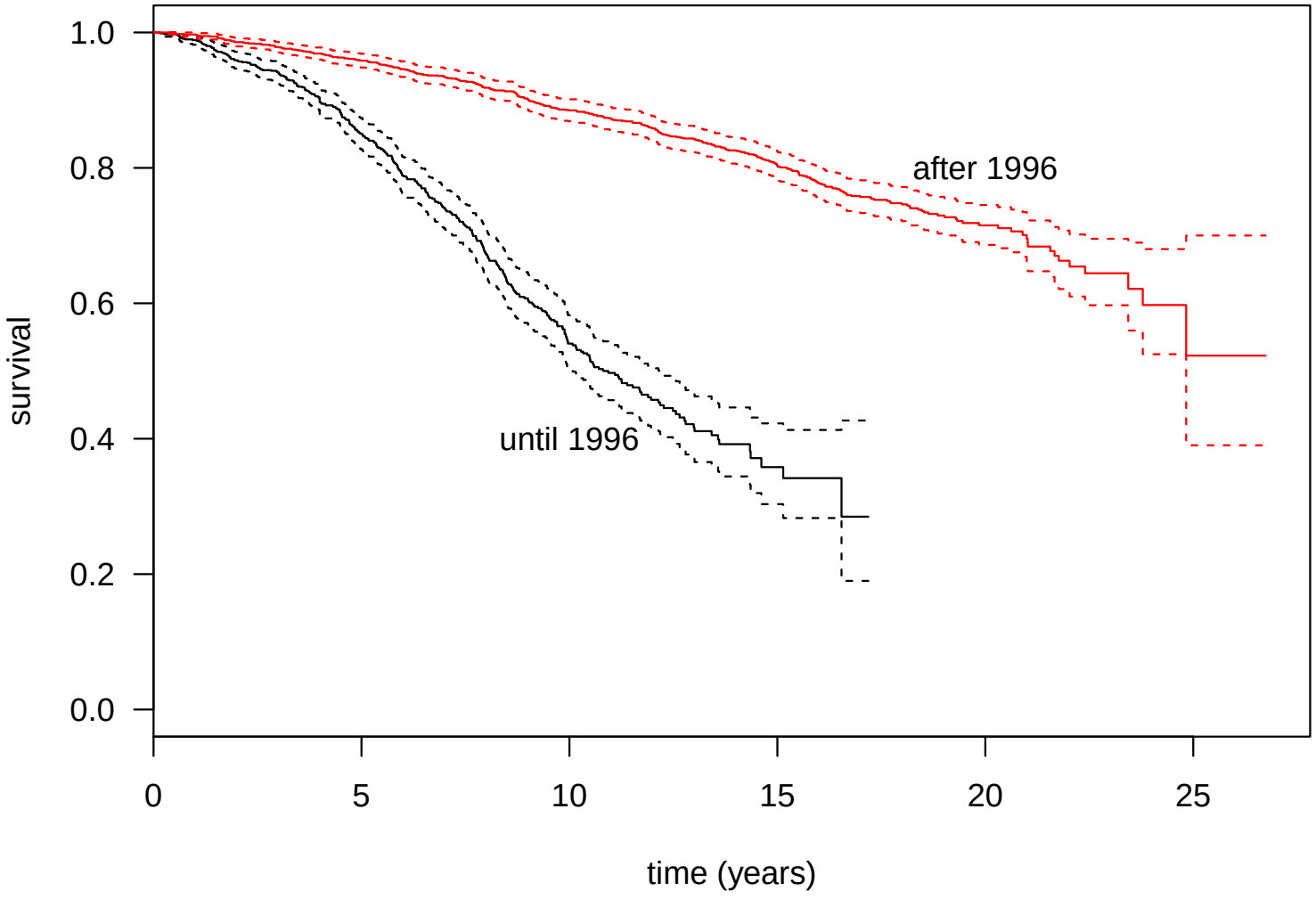
Which Approach to Choose?

# Beyond classical survival analysis

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



# Example: death after HIV infection, powerful therapy since 1996




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):

$$\lambda(t_l) = \mathrm{P}(T = t_l \,|\, T \geq t_l) \qquad \textit{discrete}$$

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

## 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):

$$\lambda(t_l) = \mathrm{P}(T = t_l \,|\, T \geq t_l) \qquad \textit{discrete}$$

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

- One-to-one relation via

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

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

## 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):

$$\lambda(t_l) = \mathrm{P}(T = t_l \,|\, T \geq t_l) \qquad \textit{discrete}$$

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

- One-to-one relation via

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

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

- Hazard basis for
  - Kaplan-Meier

# 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 |

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

Assumption: censored individuals represented by those at risk

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

## 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):

$$\lambda(t_l) = \mathrm{P}(T = t_l \,|\, T \geq t_l) \qquad \textit{discrete}$$

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

- One-to-one relation via

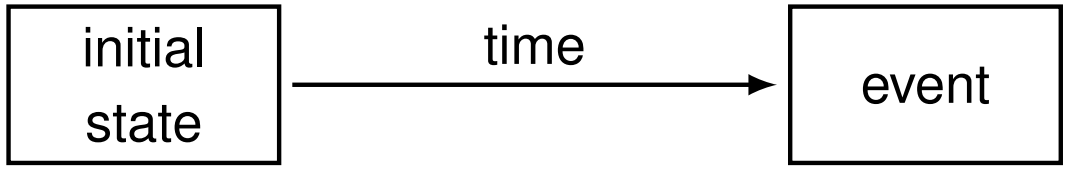
$$\mathrm{P}(T > t) = \prod_{t_i \leq t} \{1 - \lambda(t_i)\} \qquad \textit{discrete}$$

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

- Hazard basis for
  - Kaplan-Meier
  - Cox model $\lambda(t) = \lambda_0(t) \exp\{\beta_1 X_1 + \ldots + \beta_p X_p\}$

# Beyond classical survival analysis

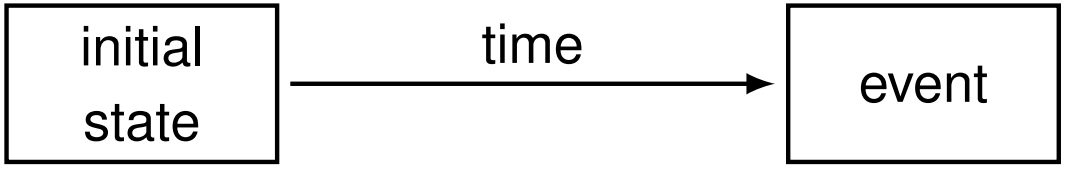
- Classical: transition between two states, one event type.
  "We all die, but not all at the same age"




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

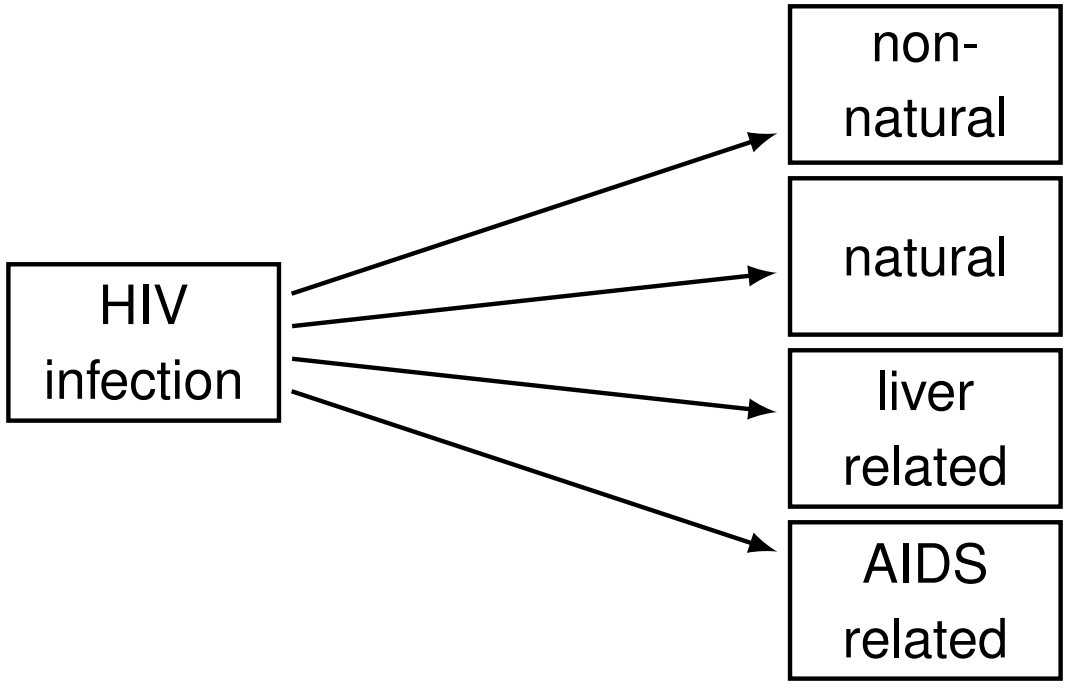
# Beyond classical survival analysis

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



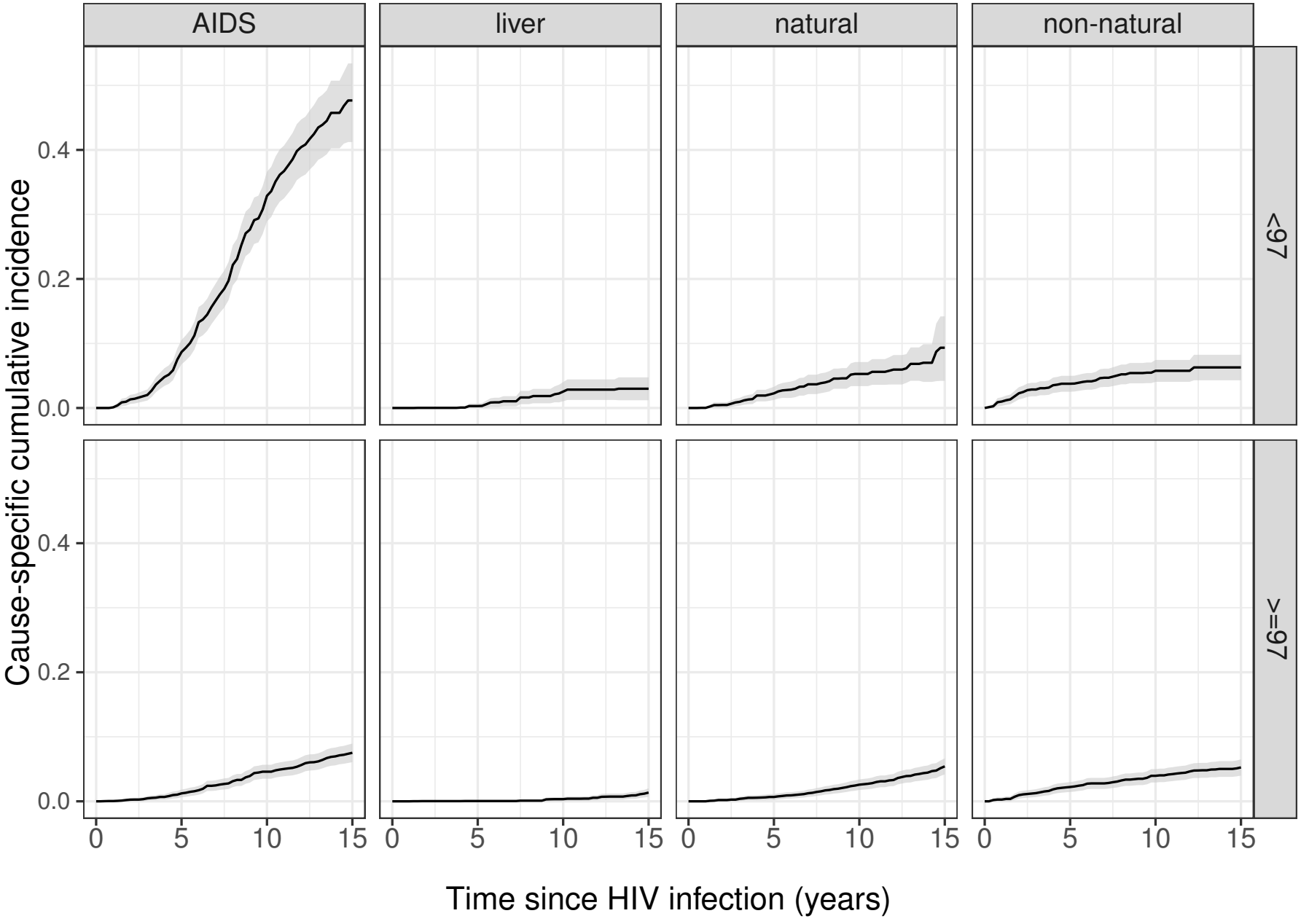

- Life and death are richer than that
  1. Multiple causes of death. Competing risks:
     Event-type outcomes that are mutually exclusive
     "we all die, but not all at the same age and from the same cause"

# 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



# Role of competing risks

- All event types of interest
  - Impact of cART on causes of death
  - Effect of dexamethasone on death and recovery after SARS-CoV-2 infection (binary outcome, complimentary levels)

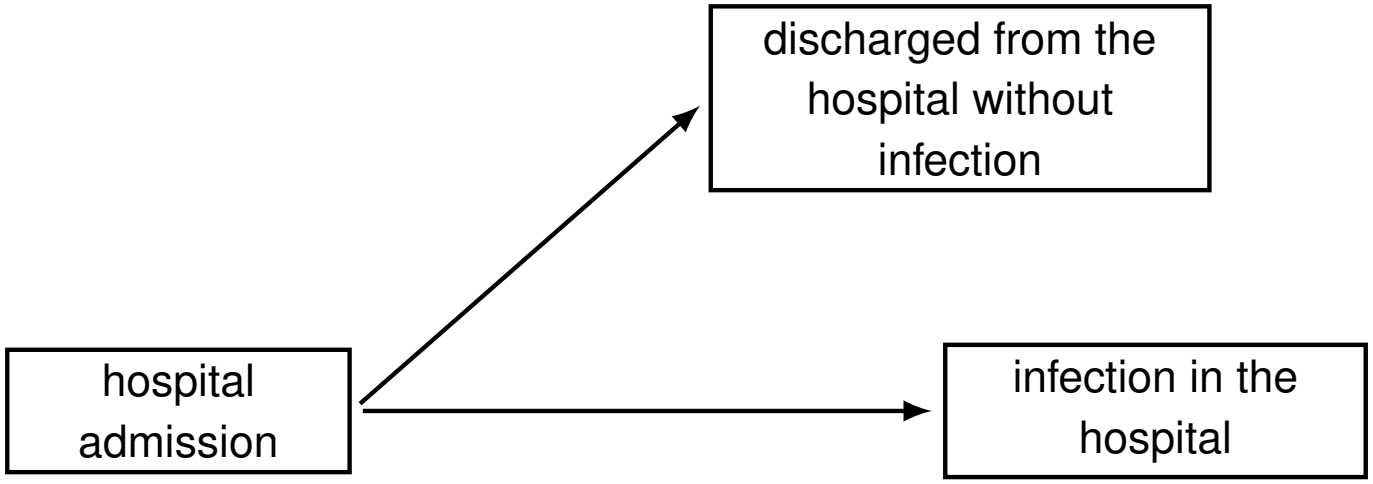
# Role of competing risks

- All event types of interest
  - Impact of cART on causes of death
  - Effect of dexamethasone on death and recovery after SARS-CoV-2 infection (binary outcome, complimentary levels)
- One event type of interest, other event prevents it from occurring
  - Intervention (transplantation/treatment) competing risk for natural disease course
  - Discharge competing risk for hospital-acquired infection
  - Death competing risk of non-fatal event (e.g. disease onset)

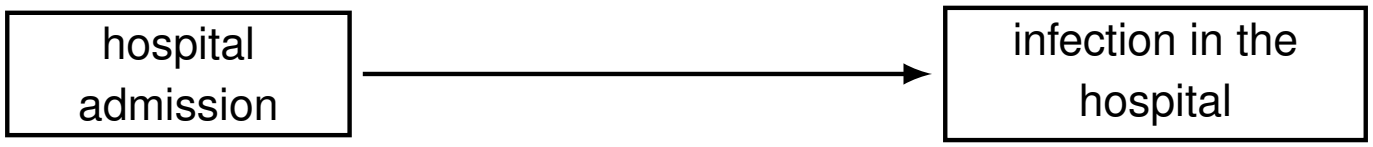
# Role of competing risks

- All event types of interest
  - Impact of cART on causes of death
  - Effect of dexamethasone on death and recovery after SARS-CoV-2 infection (binary outcome, complimentary levels)
- One event type of interest, other event prevents it from occurring
  - Intervention (transplantation/treatment) competing risk for natural disease course
  - Discharge competing risk for hospital-acquired infection
  - Death competing risk of non-fatal event (e.g. disease onset)

# II: Time to staphylococcus infection during hospital stay

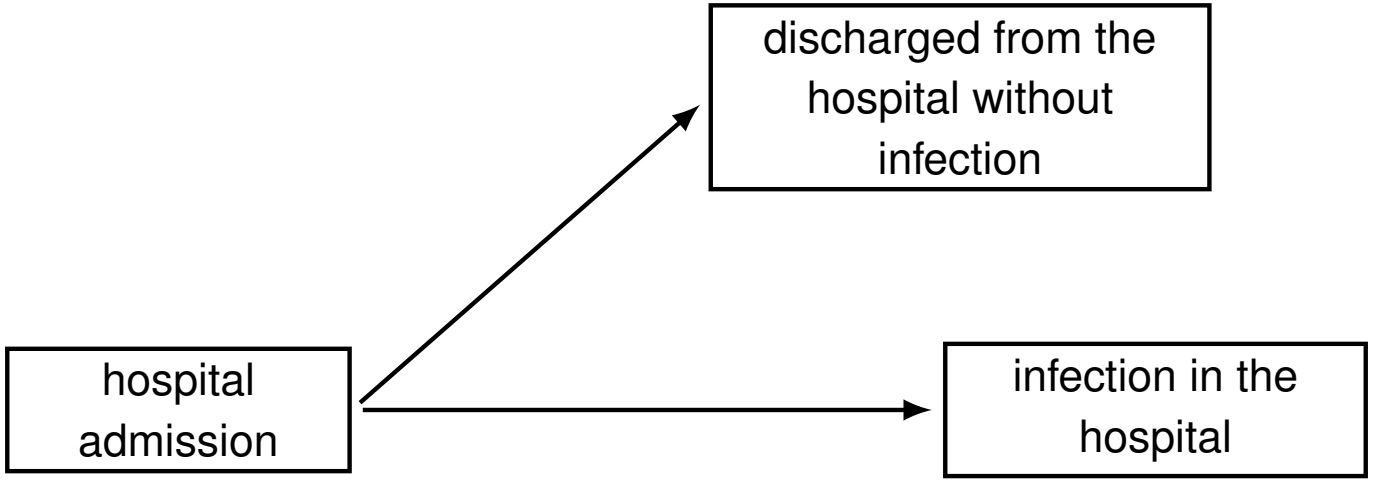


# II: Time to staphylococcus infection during hospital stay



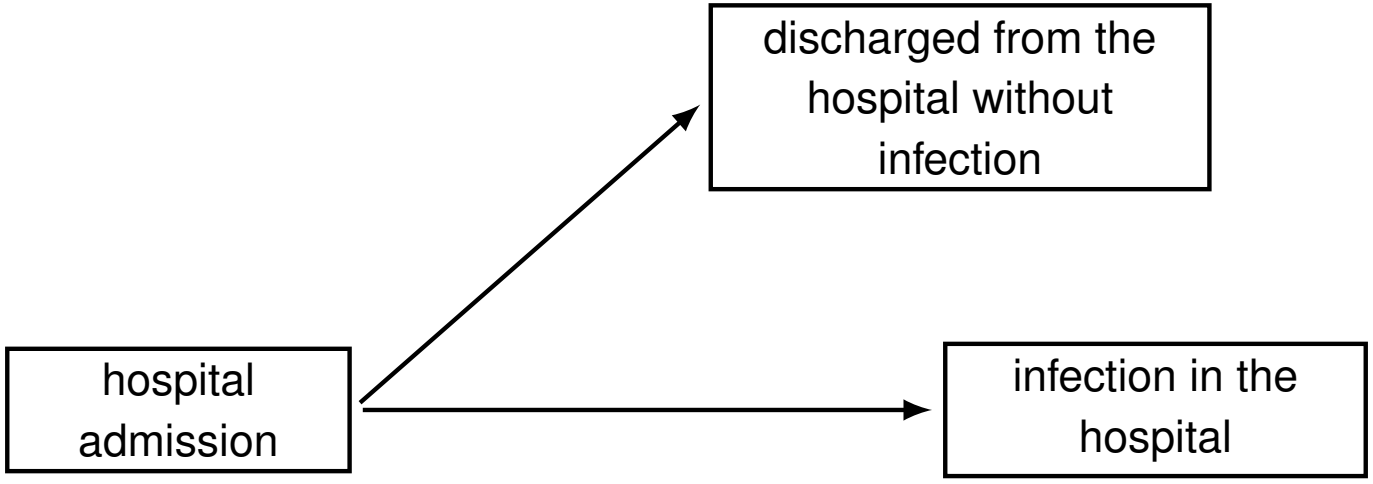

- Etiology (biological question): infection risk in hospital. What would happen if everyone stayed in hospital?
  - marginal distribution/net risk

# 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 /subdistribution/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 /subdistribution/crude risk

# Estimation, no censored 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, no censored 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 |

- Subdistribution. Estimated as frequency of events:

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

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

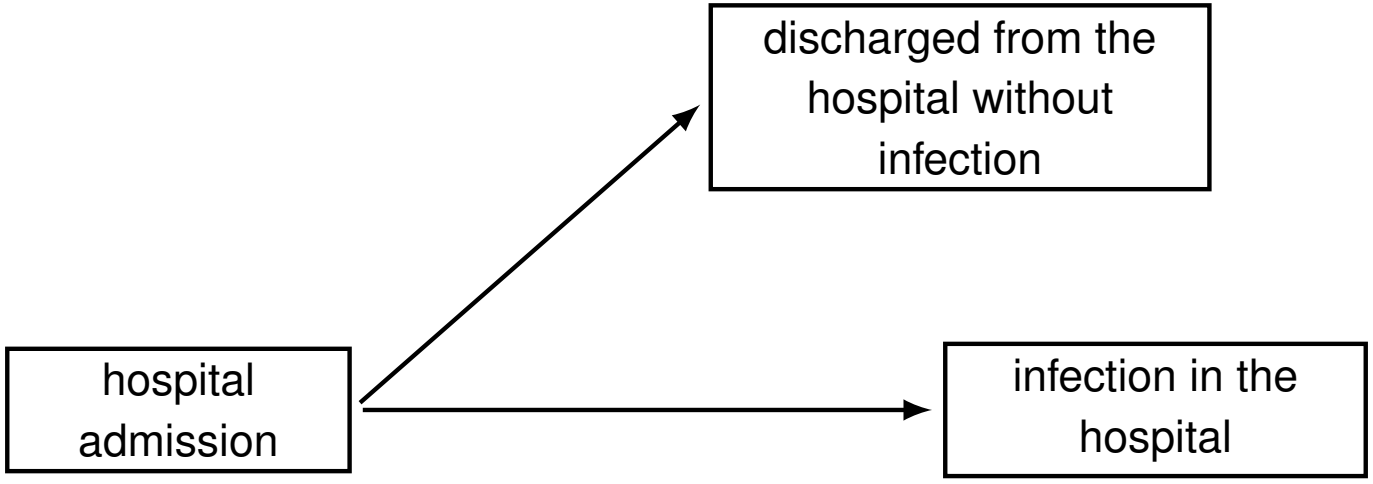
  Individuals with competing event remain in denominator, competing event ignored in estimation

# Estimation, no censored 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 |

- Subdistribution. 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
- Marginal distribution. Discharged individuals treated and interpreted as censored; assumes independent censoring

## II: Time to staphylococcus infection during hospital stay

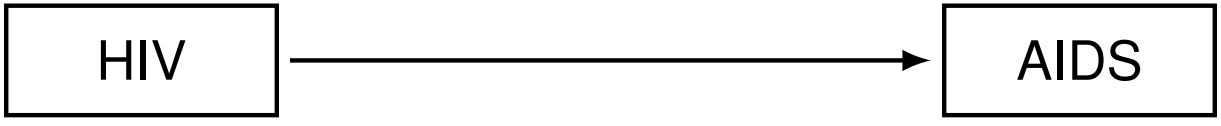



- Etiology (biological question): infection risk in hospital. What would happen if everyone stayed in hospital?
  - marginal distribution/net risk
- Predict (clinical question): burden due to infection while in hospital; discharge prevents event to occur
  - cause-specific cumulative incidence /subdistribution/crude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

# Role of competing risks

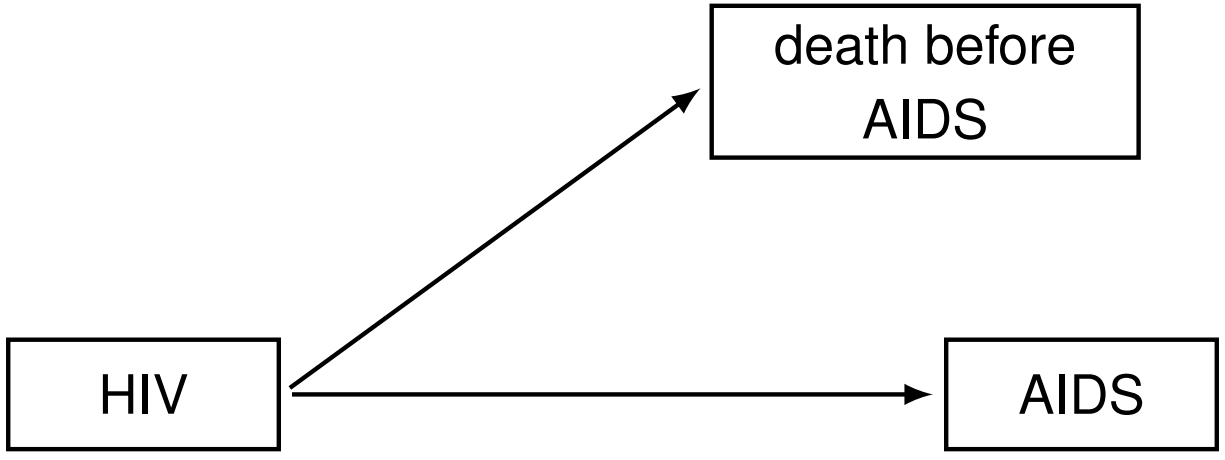
- All event types of interest
  - Impact of cART on causes of death
  - Effect of dexamethasone on death and recovery after SARS-CoV-2 infection (binary outcome, complimentary levels)
- One event type of interest, other event prevents it from occurring
  - Intervention (transplantation/treatment) competing risk for natural disease course
  - Discharge competing risk for hospital-acquired infection
  - Death competing risk of non-fatal event (e.g. disease onset)

# III: Time from HIV infection to AIDS



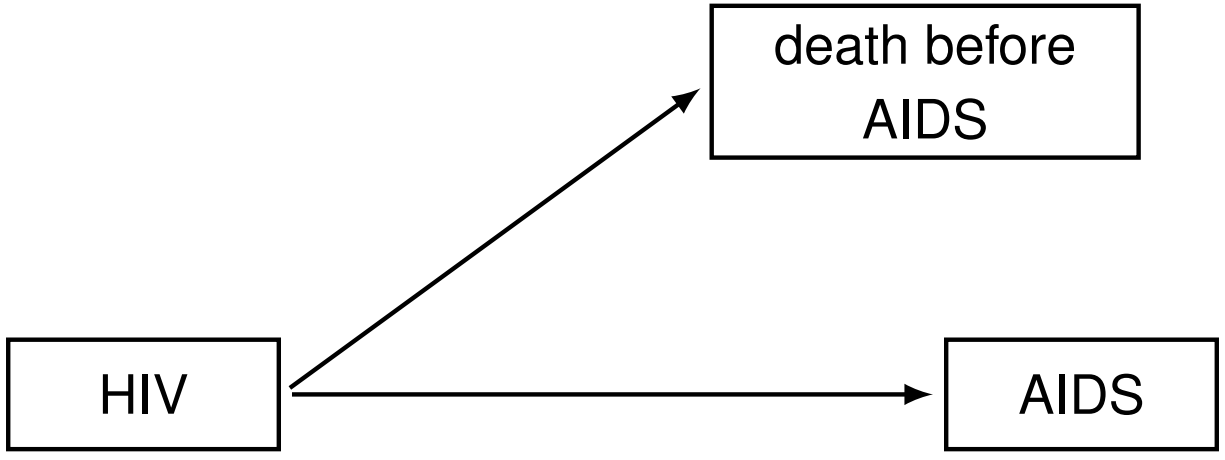

- Compare men who have sex with men (MSM) and injecting drug users (IDU)
- IDUs expected to have faster “natural” progression to AIDS

# III: Time from HIV infection to AIDS



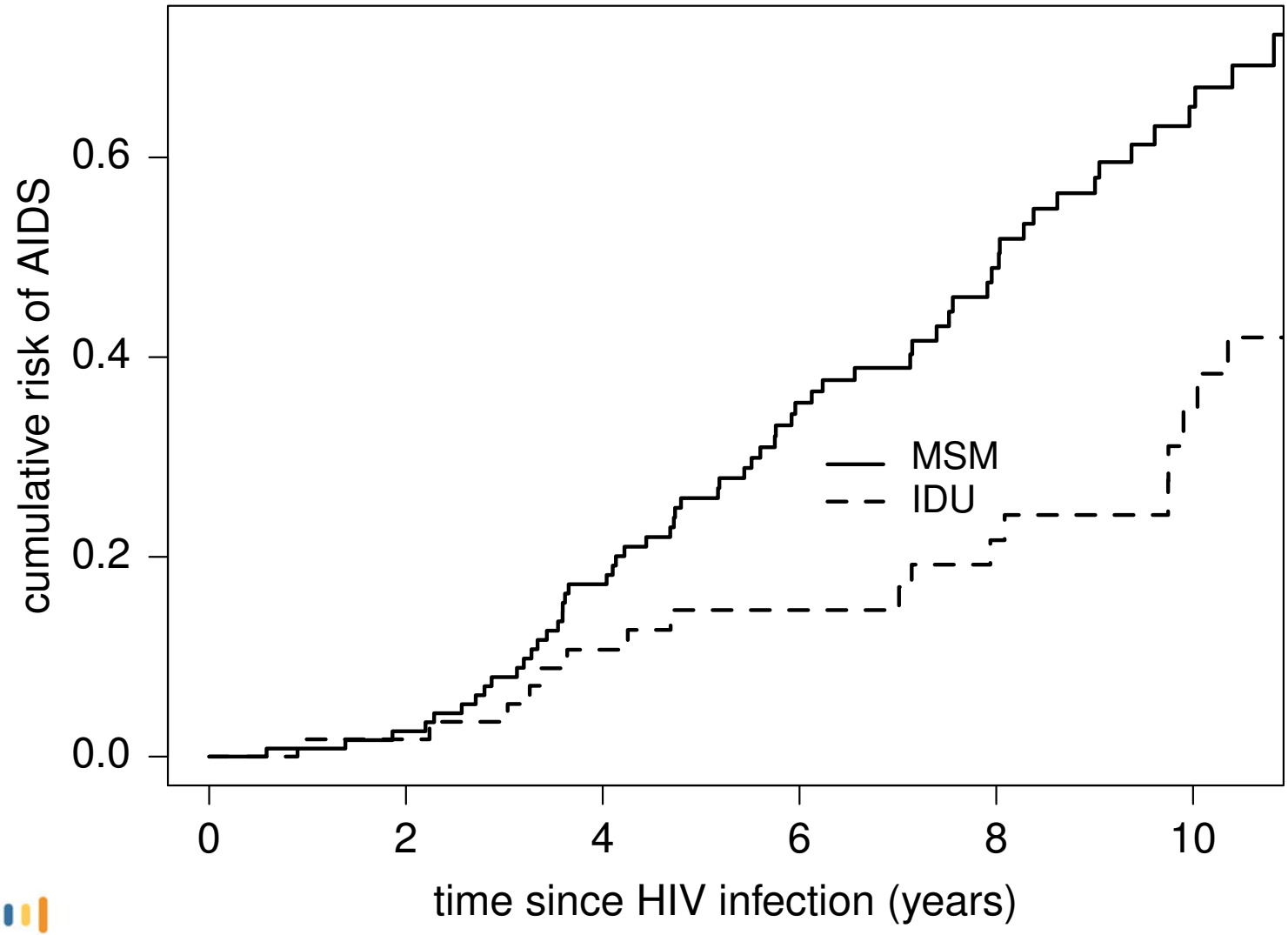

- Compare men who have sex with men (MSM) and injecting drug users (IDU)
- IDUs expected to have faster “natural” progression to AIDS
- Data from Amsterdam Cohort Studies: 99 IDU; 127 MSM

# III: Time from HIV infection to AIDS

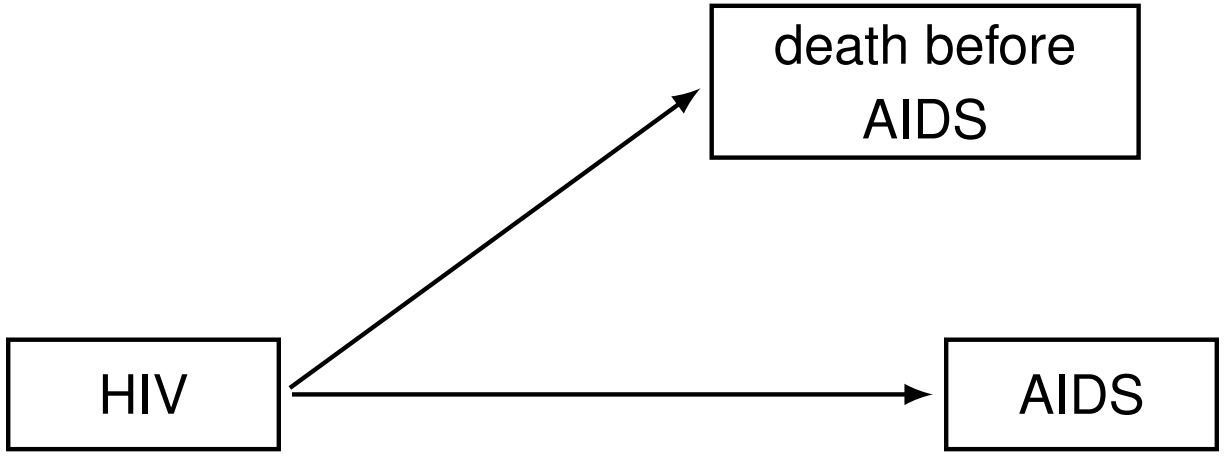



- Compare men who have sex with men (MSM) and injecting drug users (IDU)
- IDUs expected to have faster “natural” progression to AIDS
- Data from Amsterdam Cohort Studies: 99 IDU; 127 MSM
- Interest in time to AIDS if there were no pre-AIDS death.
  Interest in etiology and marginal distribution

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

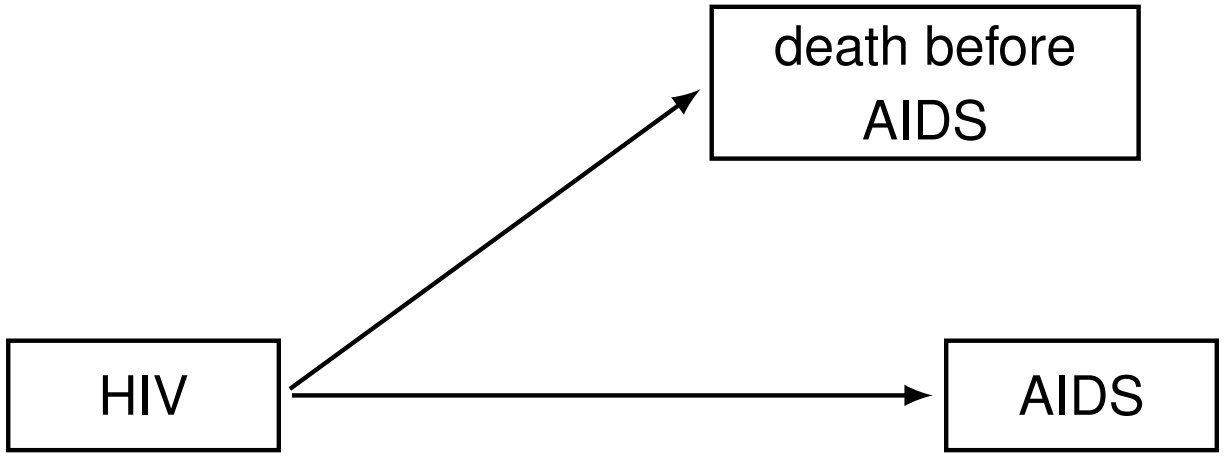


# III: Time from HIV infection to AIDS




- Compare men who have sex with men (MSM) and injecting drug users (IDU)
- IDUs expected to have faster “natural” progression to AIDS
- Data from Amsterdam Cohort Studies: 99 IDU; 127 MSM
- Interest in time to AIDS if there were no pre-AIDS death. Interest in etiology and marginal distribution
- Kaplan-Meier: leave risk set at death before AIDS

# III: Time from HIV infection to AIDS




- Compare men who have sex with men (MSM) and injecting drug users (IDU)
- IDUs expected to have faster “natural” progression to AIDS
- Data from Amsterdam Cohort Studies: 99 IDU; 127 MSM
- Interest in time to AIDS if there were no pre-AIDS death. Interest in etiology and marginal distribution
- Kaplan-Meier: leave risk set at death before AIDS
  Assumption: deaths equal AIDS risk

# Explanation: dependent censoring

- Extra 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. Censoring close to AIDS, hence estimate for IDU too optimistic

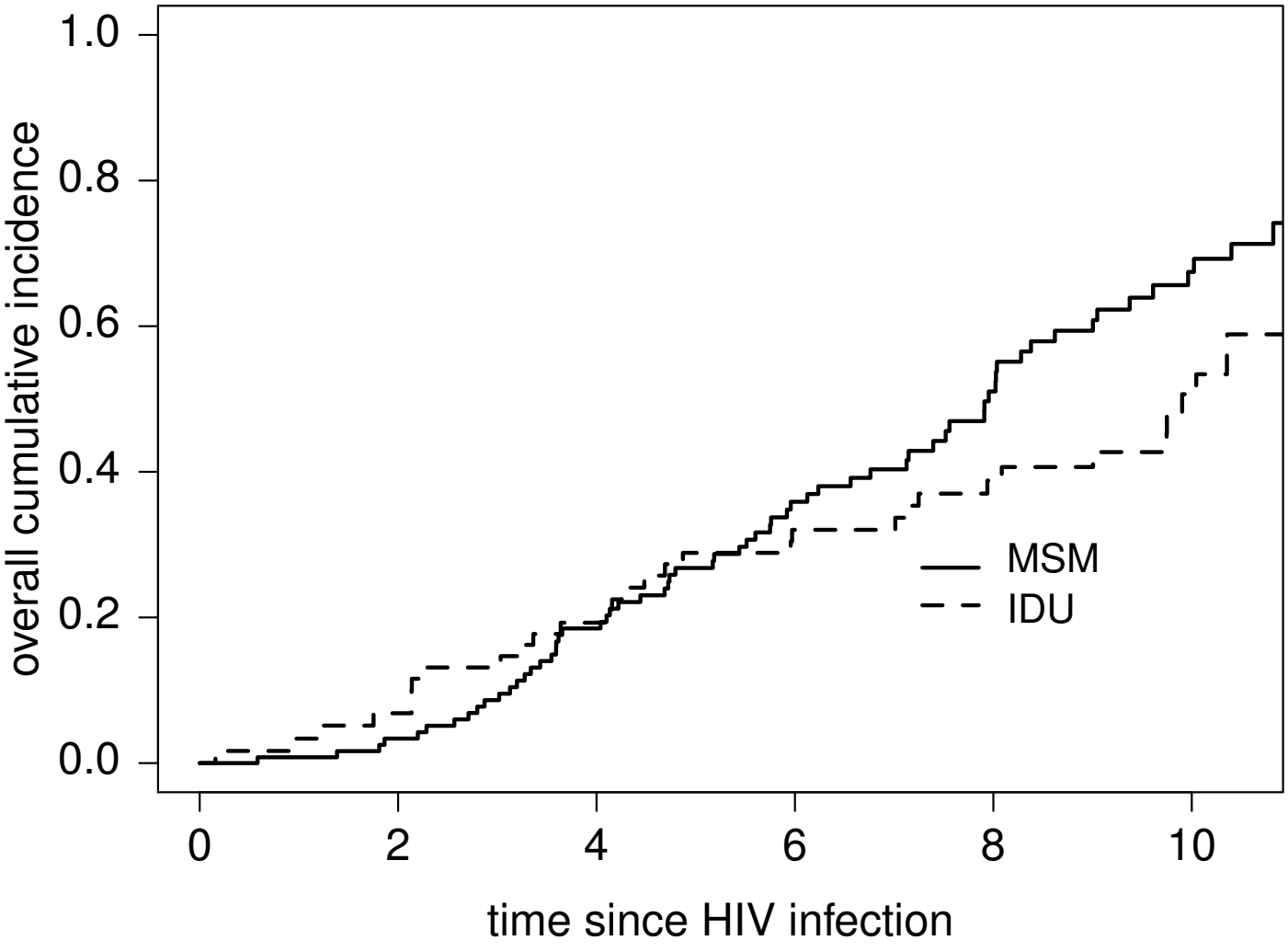
# Explanation: dependent censoring

- Extra 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. Censoring close to AIDS, hence estimate for IDU too optimistic
- What if: i) deaths would have developed AIDS right after

## i) Combine AIDS and pre-AIDS death




Overall time-to-event distribution (both event types combined)

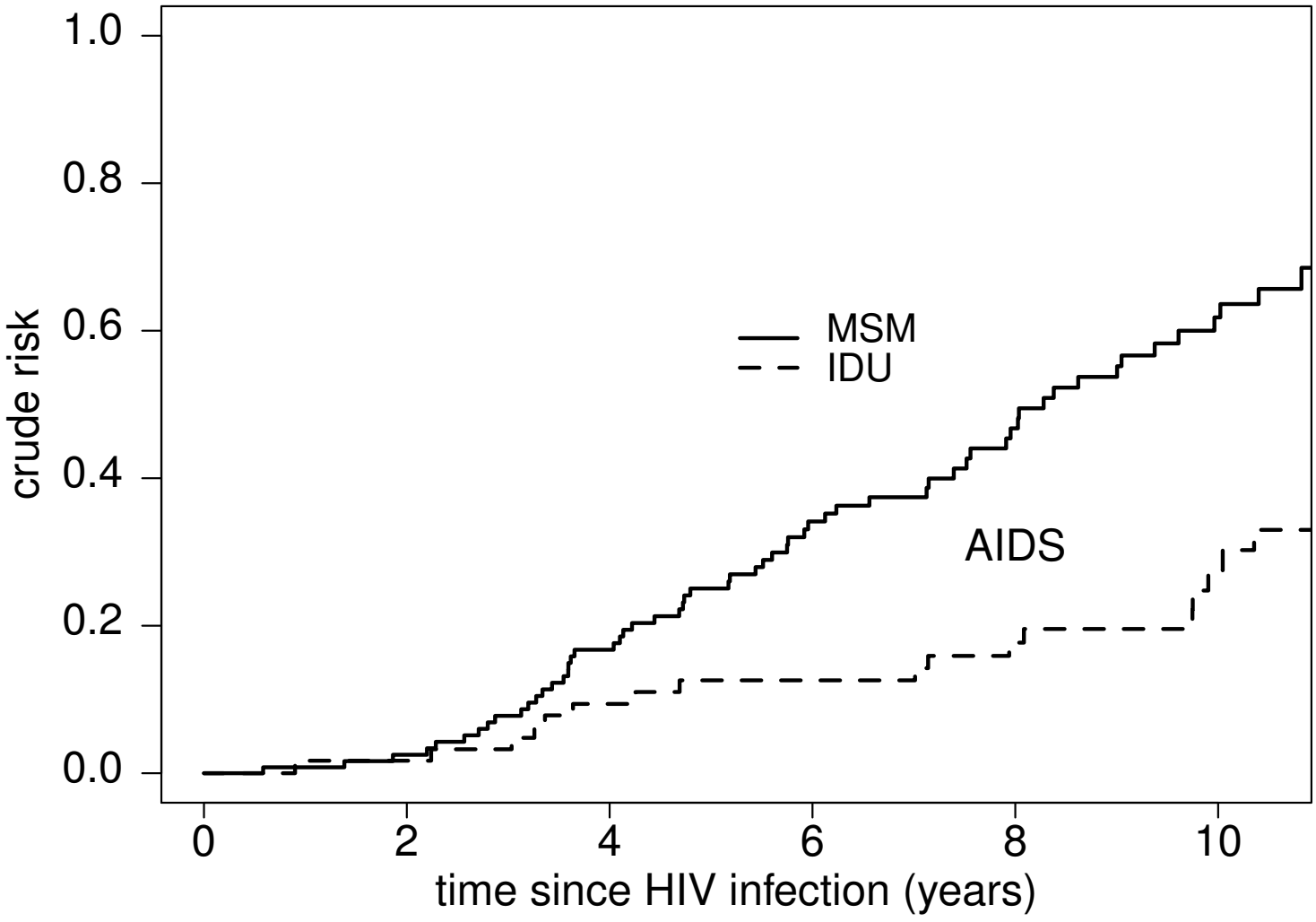
# Explanation: dependent censoring

- Extra 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 |

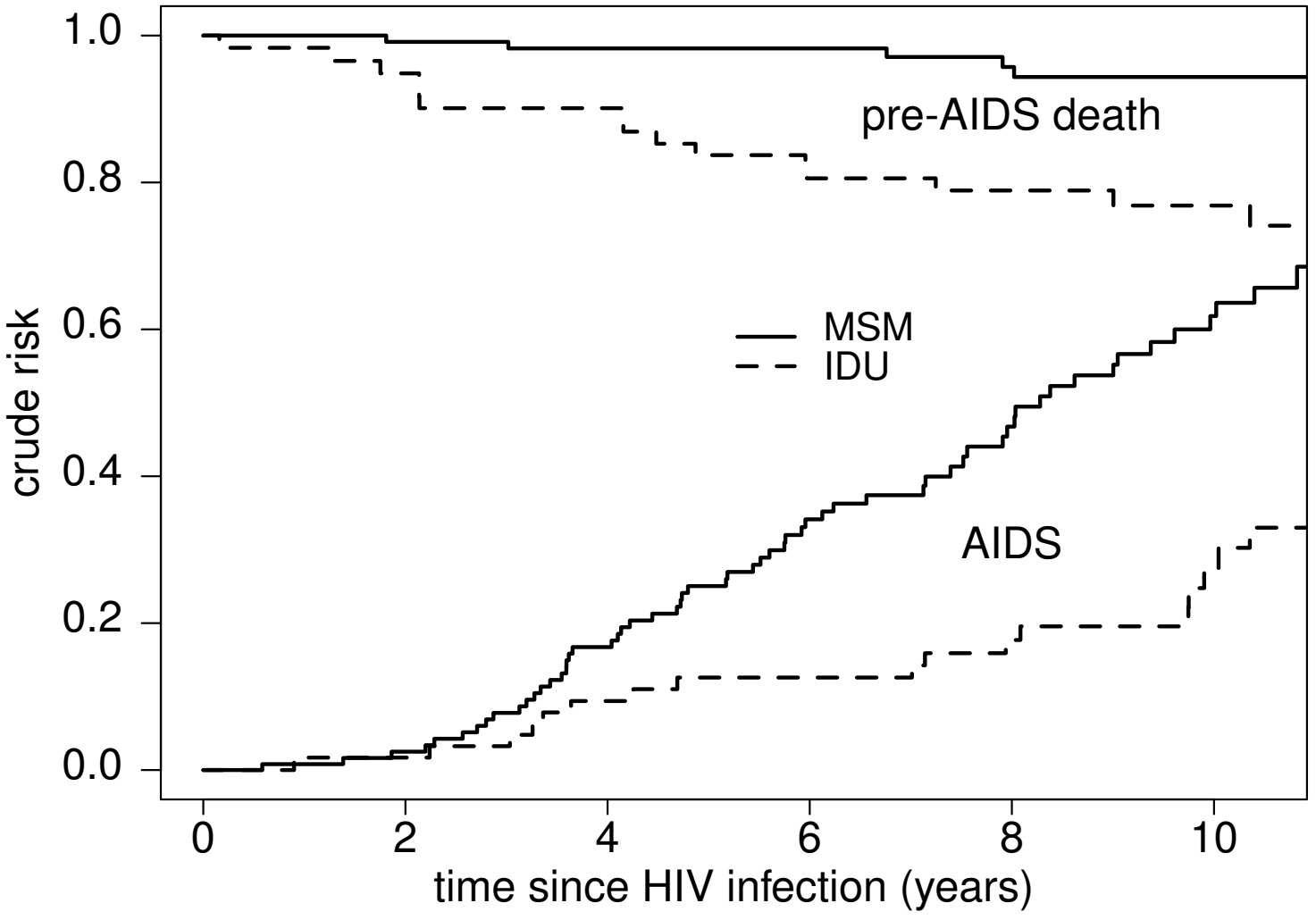
- Some causes of pre-AIDS death in IDU related to AIDS progression. Censoring close to AIDS, hence estimate for IDU too optimistic
- What if: i) deaths would have developed AIDS right after
- What if: ii) deaths would never have developed AIDS

## ii) AIDS-specific cumulative incidence



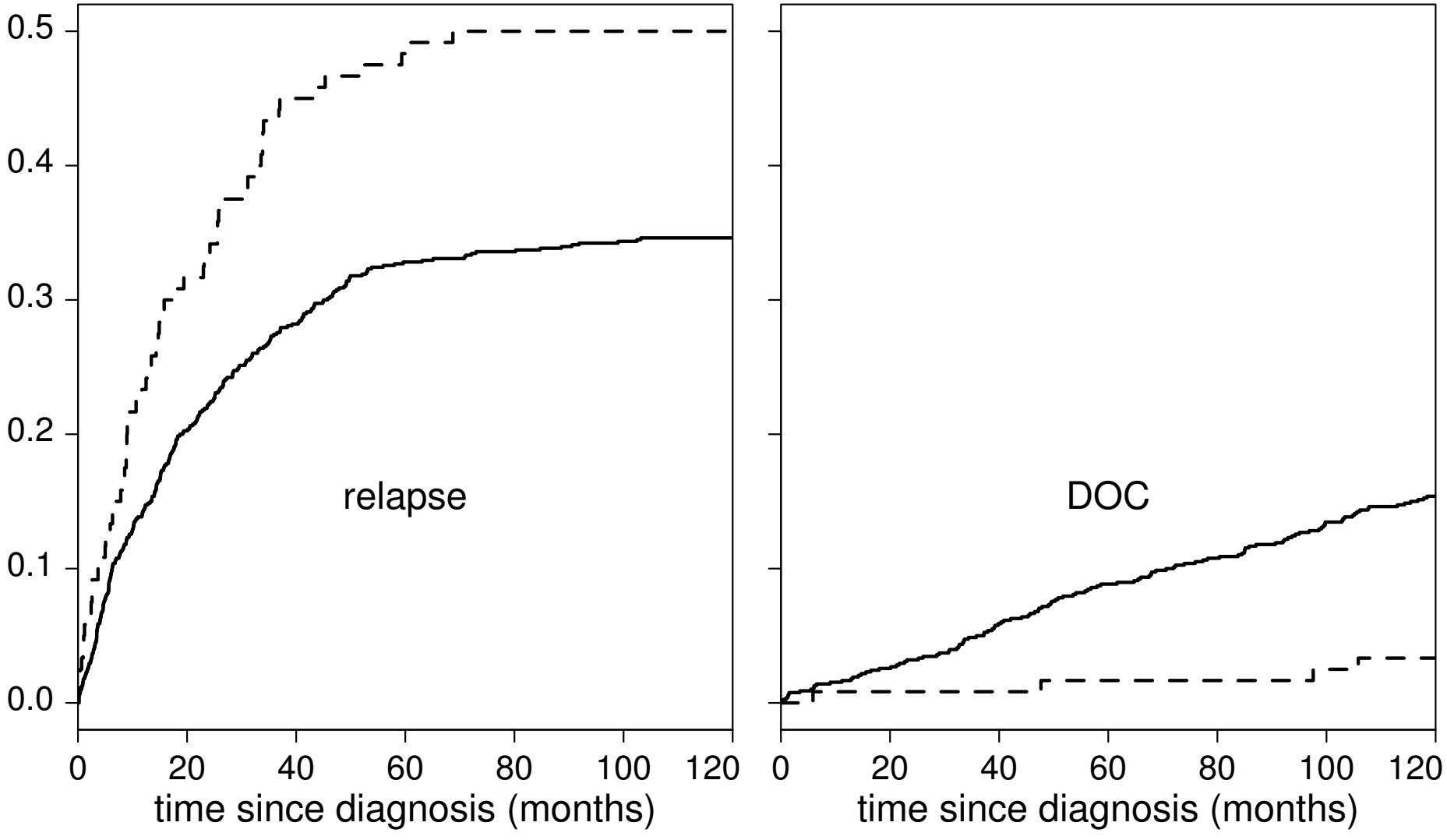

AIDS-specific cumulative incidence (pre-AIDS death prevents AIDS)

## ii) AIDS-specific cumulative incidence



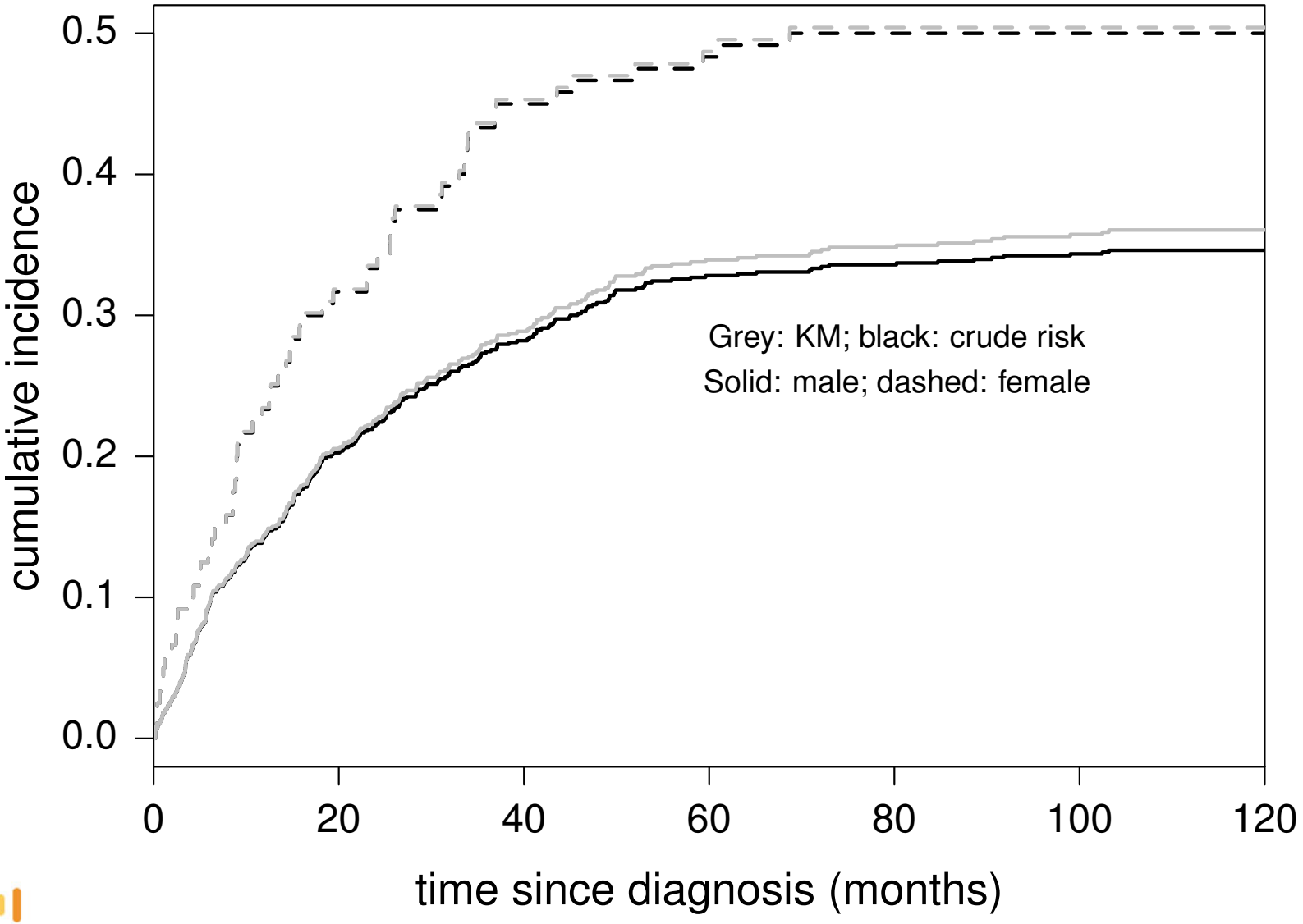

AIDS-specific cumulative incidence (pre-AIDS death prevents AIDS)

# IV: Bladder cancer; relapse, death other causes (DOC) competing

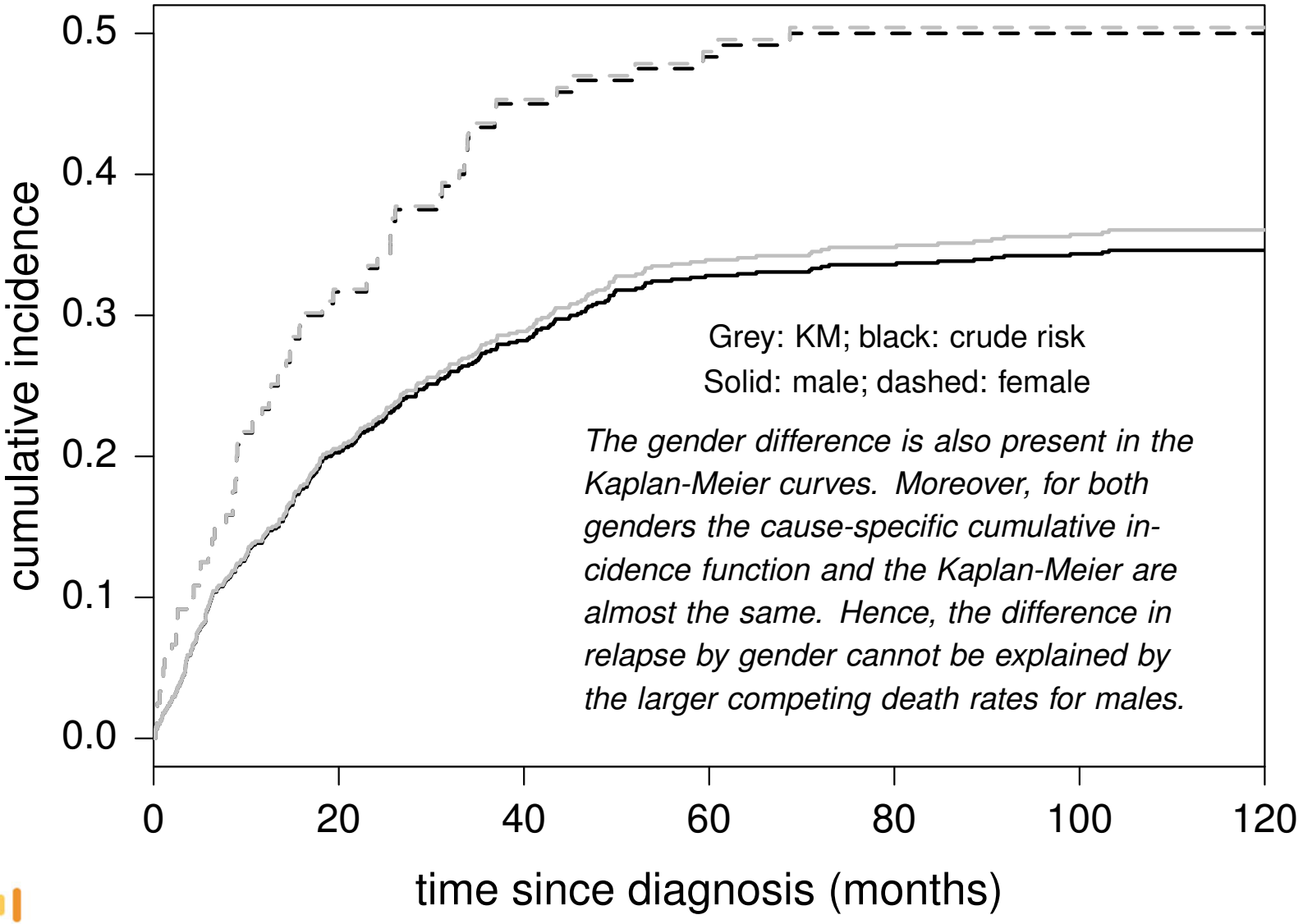



Left: relapse. Solid: male; dashed: female

# IV: Bladder cancer; relapse, DOC competing



# IV: Bladder cancer; relapse, DOC competing



# Competing risks versus marginal analysis

- Is setting without competing risk realistic or completely hypothetical? Can occurrence of competing risk be prevented?
  - Biological event: disease onset, death from “natural” cause
  - Human intervention: start of treatment, discharge, death from suicide
- Can the marginal distribution be estimated? Do those with the competing risk have the same event risk as the ones that remain event free? Is censoring due to competing risks independent?
- Examples:
  - Hospital infection and discharge
  - AIDS and pre-AIDS death
  - Relapse and DOC

# Outline

Competing Risks; When

**Competing Risks; How**

- Main Quantities
- Nonparametric Estimation
- Regression

Which Approach to Choose?

# Outline

**Competing Risks; When**

**Competing Risks; How**
- Main Quantities
- Nonparametric Estimation
- Regression

**Which Approach to Choose?**
- Marginal versus competing risks
- Which competing risks approach to choose?

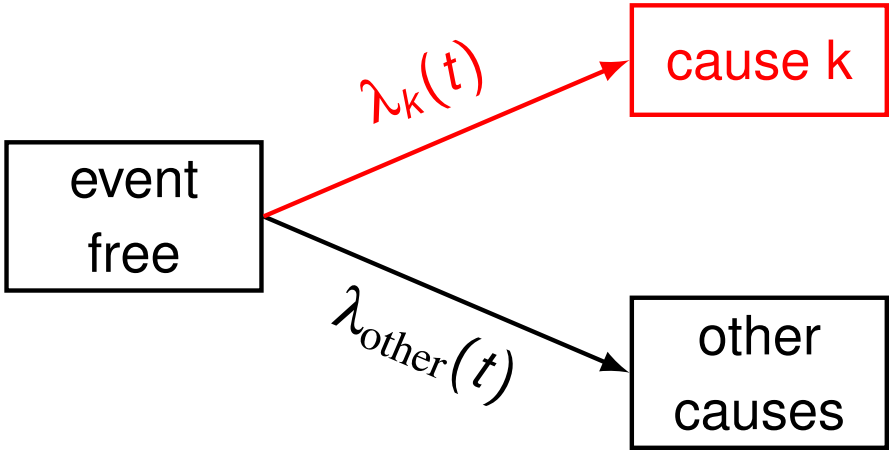
## Setup and notation

- Competing risks: type $E \in \{1, \ldots, K\}$
- $T \sim F$ time to event (of any type); $F(t) = \mathrm{P}(T \leq t)$
- Overall hazard $h$: $\mathrm{P}(T > t) = \exp\{-\int_0^t h(s)ds\}$
- Notation: $\overline{F}(t) = 1 - F(t) = \mathrm{P}(T > t)$

## Setup and notation

- Competing risks: type $E \in \{1, \ldots, K\}$
- $T \sim F$ time to event (of any type); $F(t) = \mathrm{P}(T \leq t)$
- Overall hazard $h$: $\mathrm{P}(T > t) = \exp\{-\int_0^t h(s)ds\}$
- Notation: $\overline{F}(t) = 1 - F(t) = \mathrm{P}(T > t)$
- cause-specific cumulative incidence : $F_k(t) = \mathrm{P}(T \leq t, E = k)$

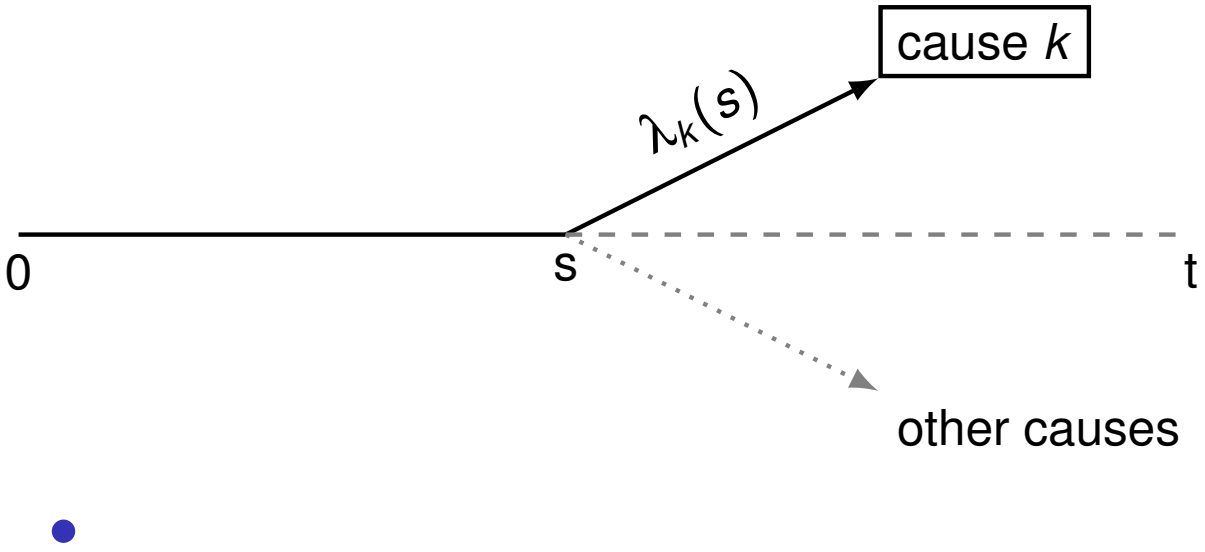
# The multi-state approach: cause-specific hazard



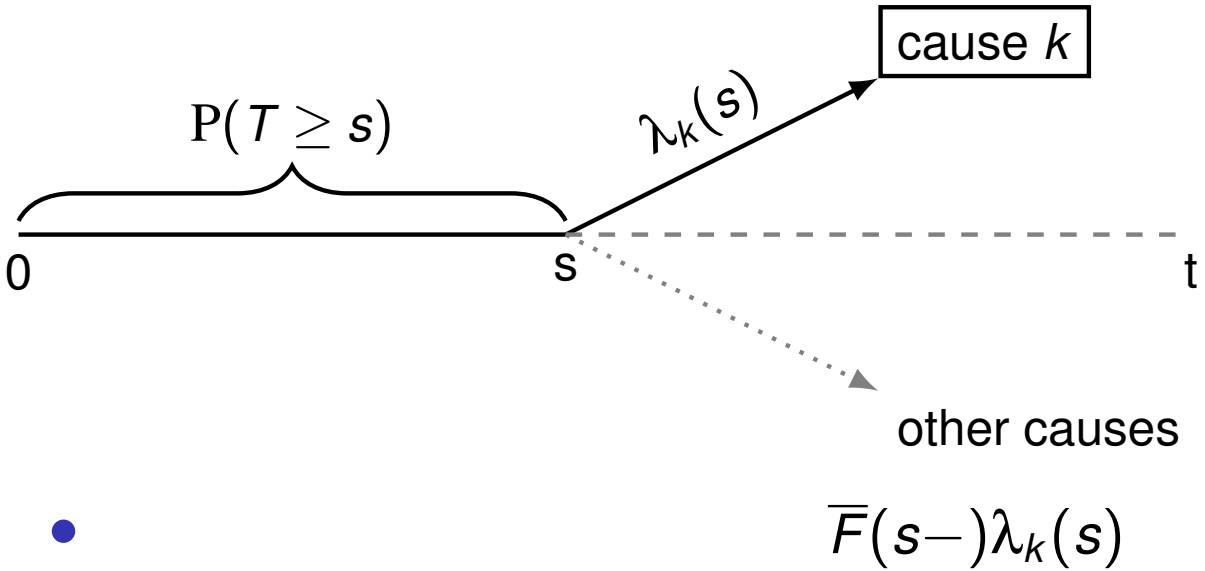

- Transition rate to cause k. For continuous distribution:

$$\lambda_k(t) = \lim_{\Delta t \downarrow 0} \frac{\mathrm{P}(t \leq T < t + \Delta t, E = k \mid T \geq t)}{\Delta t}$$

# From hazard to cumulative scale $\mathrm{P}(T \leq t, E = k)$

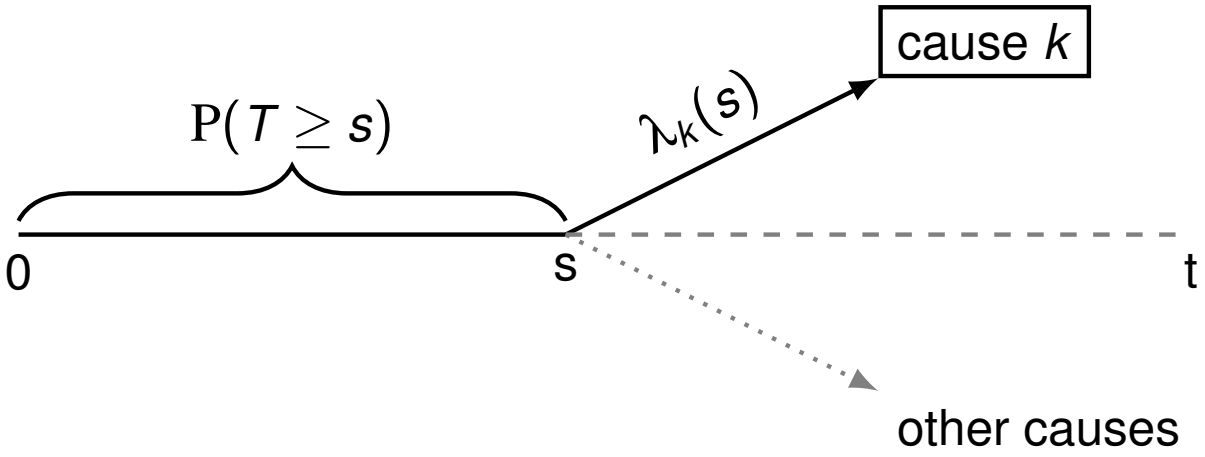


# From hazard to cumulative scale $\mathrm{P}(T \leq t, E = k)$



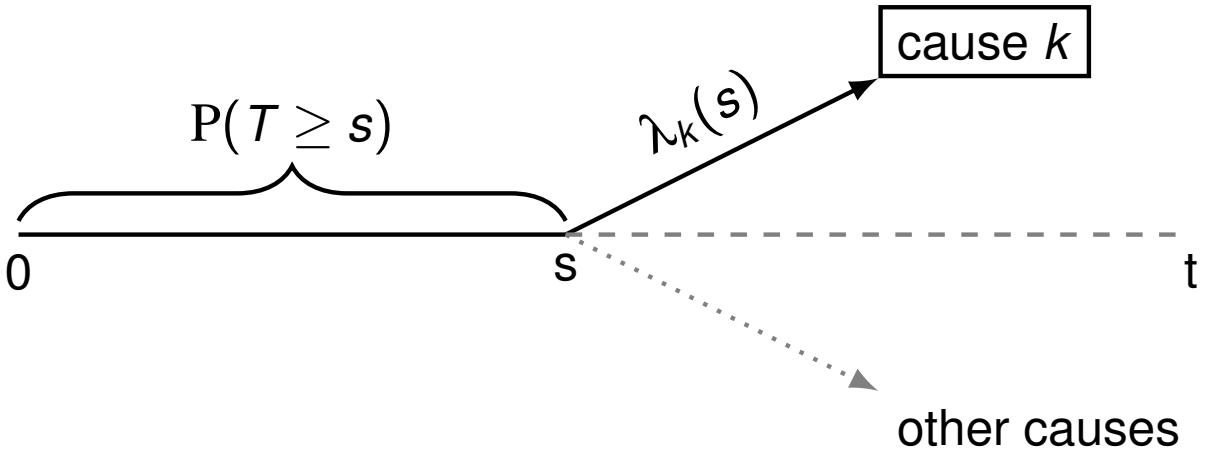

- $\overline{F}(s-)\lambda_k(s)$

# From hazard to cumulative scale $\mathrm{P}(T \leq t, E = k)$




- $F_k(t) = \mathrm{P}(T \leq t, E = k) = \int_0^t \overline{F}(s-)\lambda_k(s)ds$

# From hazard to cumulative scale $\mathrm{P}(T \leq t, E = k)$




- $F_k(t) = \mathrm{P}(T \leq t, E = k) = \int_0^t \overline{F}(s-)\lambda_k(s)ds$
- Depends on all cause-specific hazards via overall "survival"

$$\overline{F}(s) = \exp\left\{-\int_0^s h(u)du\right\} = \exp\left\{-\sum_{e=1}^{K}\int_0^s \lambda_e(u)du\right\}$$

- $\Longrightarrow \lambda_k$ does not uniquely specify $\mathrm{P}(T \leq t, E = k)$

# The subdistribution approach

# Setup and notation

- Competing risks: type $E \in \{1, \ldots, K\}$

- cause-specific cumulative incidence : $F_k(t) = \mathrm{P}(T \leq t, E = k)$
- Subdistribution random variable $T_k \sim F_k$:
  $T_k = T \times I\{E = k\} + \infty \times I\{E \neq k\}$

# Setup and notation

- Competing risks: type $E \in \{1, \ldots, K\}$

- cause-specific cumulative incidence : $F_k(t) = \mathrm{P}(T \leq t, E = k)$
- Subdistribution random variable $T_k \sim F_k$:
  $T_k = T \times I\{E = k\} + \infty \times I\{E \neq k\}$
- $\mathrm{P}(T_k \leq t) = \mathrm{P}(T \leq t, E = k), \qquad \overline{F_k}(t) = 1 - F_k(t) = \mathrm{P}(T_k > t)$

## Setup and notation

- Competing risks: type $E \in \{1, \ldots, K\}$

- cause-specific cumulative incidence : $F_k(t) = \mathrm{P}(T \leq t, E = k)$
- Subdistribution random variable $T_k \sim F_k$:
  $T_k = T \times I\{E = k\} + \infty \times I\{E \neq k\}$
- $\mathrm{P}(T_k \leq t) = \mathrm{P}(T \leq t, E = k), \qquad \overline{F_k}(t) = 1 - F_k(t) = \mathrm{P}(T_k > t)$
- Subdistribution hazard $h_k$:

$$\mathrm{P}(T_k > t) = \exp\{-\int_0^t h_k(s)ds\}$$

## Setup and notation

- Competing risks: type $E \in \{1, \ldots, K\}$
- Overall hazard $h$: $\mathrm{P}(T > t) = \exp\{-\int_0^t h(s)ds\}$
- cause-specific cumulative incidence : $F_k(t) = \mathrm{P}(T \leq t, E = k)$
- Subdistribution random variable $T_k \sim F_k$:
  $T_k = T \times I\{E = k\} + \infty \times I\{E \neq k\}$
- $\mathrm{P}(T_k \leq t) = \mathrm{P}(T \leq t, E = k), \qquad \overline{F_k}(t) = 1 - F_k(t) = \mathrm{P}(T_k > t)$
- Subdistribution hazard $h_k$:

$$\mathrm{P}(T_k > t) = \exp\{-\int_0^t h_k(s)ds\}$$

## The subdistribution approach

- Subdistribution hazard ($T_k = T \times I\{E = k\} + \infty \times I\{E \neq k\}$):

$$h_k(t) = \lim_{\Delta t \downarrow 0} \frac{1}{\Delta t} \times \mathrm{P}\{t \leq T_k < t + \Delta t \,|\, T_k \geq t\}$$

# The subdistribution approach

- Subdistribution hazard ($T_k = T \times I\{E = k\} + \infty \times I\{E \neq k\}$):

$$h_k(t) = \lim_{\Delta t \downarrow 0} \frac{1}{\Delta t} \times \mathrm{P}\{t \leq T_k < t + \Delta t \,|\, T_k \geq t\}$$

$$= \lim_{\Delta t \downarrow 0} \frac{\frac{1}{\Delta t}\mathrm{P}\{t \leq T < t + \Delta t, E = k\}}{\mathrm{P}\{T \geq t \text{ or } (T < t, E \neq k)\}}$$

- Denominator: event free or with earlier competing event

# The subdistribution approach

- Subdistribution hazard ($T_k = T \times I\{E = k\} + \infty \times I\{E \neq k\}$):

$$h_k(t) = \lim_{\Delta t \downarrow 0} \frac{1}{\Delta t} \times \mathrm{P}\{t \leq T_k < t + \Delta t \,|\, T_k \geq t\}$$

$$= \lim_{\Delta t \downarrow 0} \frac{\frac{1}{\Delta t}\mathrm{P}\{t \leq T < t + \Delta t, E = k\}}{\mathrm{P}\{T \geq t \text{ or } (T < t, E \neq k)\}}$$

- Denominator: event free or with earlier competing event
- Interpretation controversial
  - Not a rate in epidemiological sense,

# The subdistribution approach

- Subdistribution hazard ($T_k = T \times I\{E = k\} + \infty \times I\{E \neq k\}$):

$$h_k(t) = \lim_{\Delta t \downarrow 0} \frac{1}{\Delta t} \times \mathrm{P}\{t \leq T_k < t + \Delta t \,|\, T_k \geq t\}$$

$$= \lim_{\Delta t \downarrow 0} \frac{\frac{1}{\Delta t}\mathrm{P}\{t \leq T < t + \Delta t, E = k\}}{\mathrm{P}\{T \geq t \text{ or } (T < t, E \neq k)\}}$$

- Denominator: event free or with earlier competing event
- Interpretation controversial
  - Not a rate in epidemiological sense, unless we can assume that those with the competing event were immune for the event of interest (cure model)

# The subdistribution approach

- Subdistribution hazard ($T_k = T \times I\{E = k\} + \infty \times I\{E \neq k\}$):

$$h_k(t) = \lim_{\Delta t \downarrow 0} \frac{1}{\Delta t} \times \mathrm{P}\{t \leq T_k < t + \Delta t \,|\, T_k \geq t\}$$

$$= \lim_{\Delta t \downarrow 0} \frac{\frac{1}{\Delta t}\mathrm{P}\{t \leq T < t + \Delta t, E = k\}}{\mathrm{P}\{T \geq t \text{ or } (T < t, E \neq k)\}}$$

- Denominator: event free or with earlier competing event
- Interpretation controversial
  - Not a rate in epidemiological sense, unless we can assume that those with the competing event were immune for the event of interest (cure model)
- One-to-one relation with crude risk

$$\overline{F_k}(t) = \prod_{t_l \leq t} \left\{1 - h_k(t_l)\right\} \quad \text{or} \quad \overline{F_k}(t) = \exp\left\{-\int_0^t h_k(u)du\right\}$$

# Rates and risks in competing risks setting

| | hazard | | cumulative | |
|---|---|---|---|---|
| competing risks | marginal | $\lambda^m$ | net risk | $F^m(t)$ |
| | | | marginal survival function | |
| | | | marginal cumulative incidence | |
| | cause-specific | $\lambda_k$ | no corresponding quantity | |
| | subdistribution | $h_k$ | crude risk | $F_k(t)$ |
| | | | cause-specific cumulative incidence | |
| combined | overall | $h$ | overall risk | $F(t)$ |
| | | | overall survival function | |
| | | | overall cumulative incidence | |

# Rates and risks in competing risks setting

| | hazard | | cumulative | |
|---|---|---|---|---|
| competing risks | marginal | $\lambda^m$ | net risk | $F^m(t)$ |
| | | | marginal survival function | |
| | | | marginal cumulative incidence | |
| | cause-specific | $\lambda_k$ | no corresponding quantity | |
| | subdistribution | $h_k$ | crude risk | $F_k(t)$ |
| | | | cause-specific cumulative incidence | |
| combined | overall | $h$ | overall risk | $F(t)$ |
| | | | overall survival function | |
| | | | overall cumulative incidence | |

# Outline

**Competing Risks; When**

**Competing Risks; How**

- Main Quantities
- Nonparametric Estimation
- Regression

**Which Approach to Choose?**

- Marginal versus competing risks
- Which competing risks approach to choose?

# Observed data

$$\{(x_1, e_1\delta_1), \ldots, (x_N, e_N\delta_N)\}$$

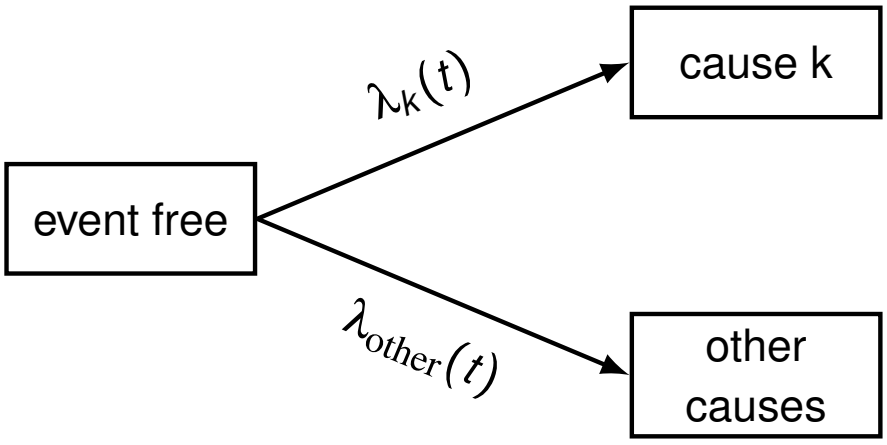
- $x_i = \min\{t_i, c_i\}$, $\delta_i = \{t_i \leq c_i\}$, $e_i \in \{1, \ldots, K\}$
- $t_{(i)}$ ordered unique event times of any type

# Observed data

$\{(x_1, e_1\delta_1), \ldots, (x_N, e_N\delta_N)\}$

- $x_i = \min\{t_i, c_i\}$, $\delta_i = \{t_i \leq c_i\}$, $e_i \in \{1, \ldots, K\}$
- $t_{(i)}$ ordered unique event times of any type
- $r(t_{(i)})$ number observed at risk
- $d_k(t_{(i)})$ number of events at $t_{(i)}$ of type $k$

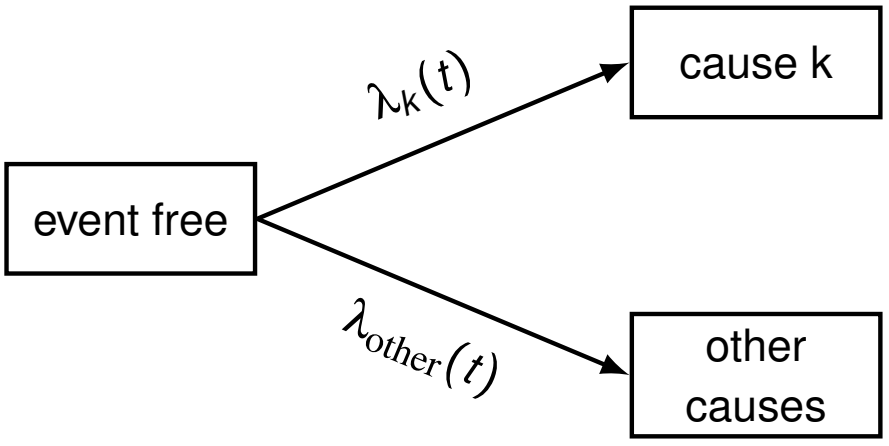
# Cause-specific hazard




- Individuals with a competing event are no longer at risk $\Longrightarrow$ leave the risk set

$$\widehat{\lambda}_k(t_{(i)}) = \frac{d_k(t_{(i)})}{r(t_{(i)})} \ .$$

# Cause-specific hazard




- Individuals with a competing event are no longer at risk $\Longrightarrow$ leave the risk set

$$\widehat{\lambda}_k(t_{(i)}) = \frac{d_k(t_{(i)})}{r(t_{(i)})} \ .$$

- Standard rate estimation. Same estimator as marginal hazard, but different interpretation, unless censoring due to competing risks is non-informative

# Aalen-Johansen estimator of $F_k$

- Plug-in estimator based on $F_k(t) = \int_0^t \mathrm{P}\{T \geq s\}\lambda_k(s)ds$:

$$\widehat{F}_k^{\mathrm{AJ}}(t) = \sum_{i:t_{(i)} \leq t} \widehat{\overline{F}}^{\mathrm{PL}}(t_{(i)}-) \times \widehat{\lambda}_k(t_{(i)}) \text{ with}$$

$$\widehat{\lambda}_k(t_{(i)}) = \frac{d_k(t_{(i)})}{r(t_{(i)})} \qquad \text{cause specific hazard}$$

# Aalen-Johansen estimator of $F_k$

- Plug-in estimator based on $F_k(t) = \int_0^t \mathrm{P}\{T \geq s\}\lambda_k(s)ds$:

$$\widehat{F}_k^{\mathrm{AJ}}(t) = \sum_{i:t_{(i)} \leq t} \widehat{\overline{F}}^{\mathrm{PL}}(t_{(i)}-) \times \widehat{\lambda}_k(t_{(i)}) \text{ with}$$

$$\widehat{\lambda}_k(t_{(i)}) = \frac{d_k(t_{(i)})}{r(t_{(i)})} \qquad \text{cause specific hazard}$$

# Aalen-Johansen estimator of $F_k$

- Plug-in estimator based on $F_k(t) = \int_0^t \mathrm{P}\{T \geq s\}\lambda_k(s)ds$:

$$\widehat{F}_k^{\mathrm{AJ}}(t) = \sum_{i:t_{(i)} \leq t} \widehat{\overline{F}}^{\mathrm{PL}}(t_{(i)}-) \times \widehat{\lambda}_k(t_{(i)}) \text{ with}$$

$$\widehat{\lambda}_k(t_{(i)}) = \frac{d_k(t_{(i)})}{r(t_{(i)})} \qquad \text{cause specific hazard}$$

$$\widehat{\overline{F}}^{\mathrm{PL}}(t_{(i)}-) = \prod_{j:t_{(j)} < t_{(i)}} \left(1 - \frac{d(t_{(j)})}{r(t_{(j)})}\right) \text{ Kaplan-Meier}$$

# Observed data

$\{(x_1, e_1\delta_1), \ldots, (x_N, e_N\delta_N)\}$

- $x_i = \min\{t_i, c_i\}$, $\delta_i = \{t_i \leq c_i\}$, $e_i \in \{1, \ldots, K\}$
- $t_{(i)}$ ordered unique event times of any type
- $r(t_{(i)})$ number observed at risk

- $d_k(t_{(i)})$ number of events at $t_{(i)}$ of type $k$
- $d(t_{(i)})$ total number of events at $t_{(i)}$

# Aalen-Johansen estimator of $F_k$

- Plug-in estimator based on $F_k(t) = \int_0^t \mathrm{P}\{T \geq s\}\lambda_k(s)ds$:

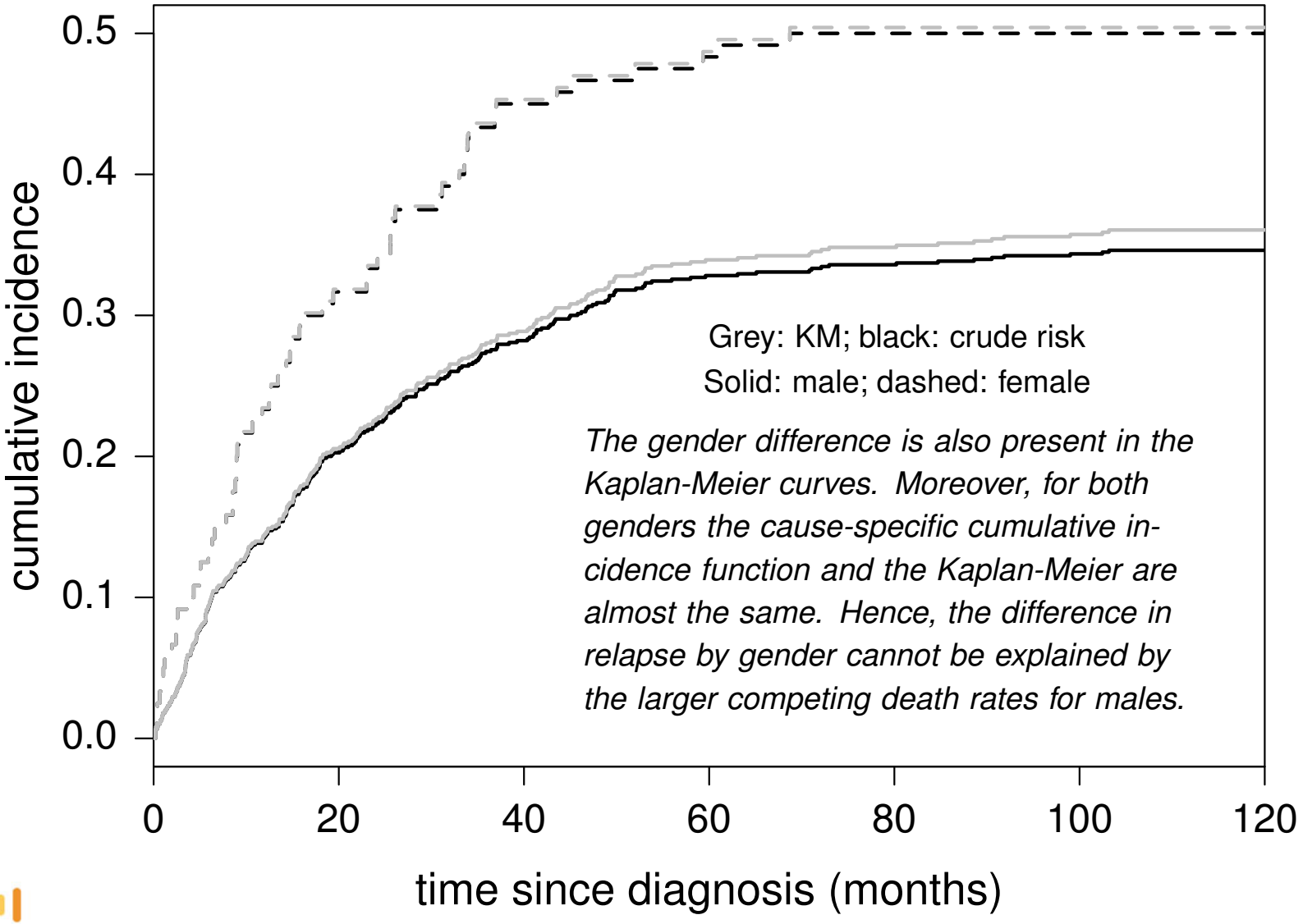
$$\widehat{F}_k^{\mathrm{AJ}}(t) = \sum_{i:t_{(i)} \leq t} \widehat{\overline{F}}^{\mathrm{PL}}(t_{(i)}-) \times \widehat{\lambda}_k(t_{(i)}) \text{ with}$$

$$\widehat{\lambda}_k(t_{(i)}) = \frac{d_k(t_{(i)})}{r(t_{(i)})} \qquad \text{cause specific hazard}$$

$$\widehat{\overline{F}}^{\mathrm{PL}}(t_{(i)}-) = \prod_{j:t_{(j)} < t_{(i)}} \left(1 - \frac{d(t_{(j)})}{r(t_{(j)})}\right) \text{ Kaplan-Meier}$$

- Single event type: equal to Kaplan-Meier
- Competing risks: same hazard estimate as for marginal distribution, but cumulative quantity different

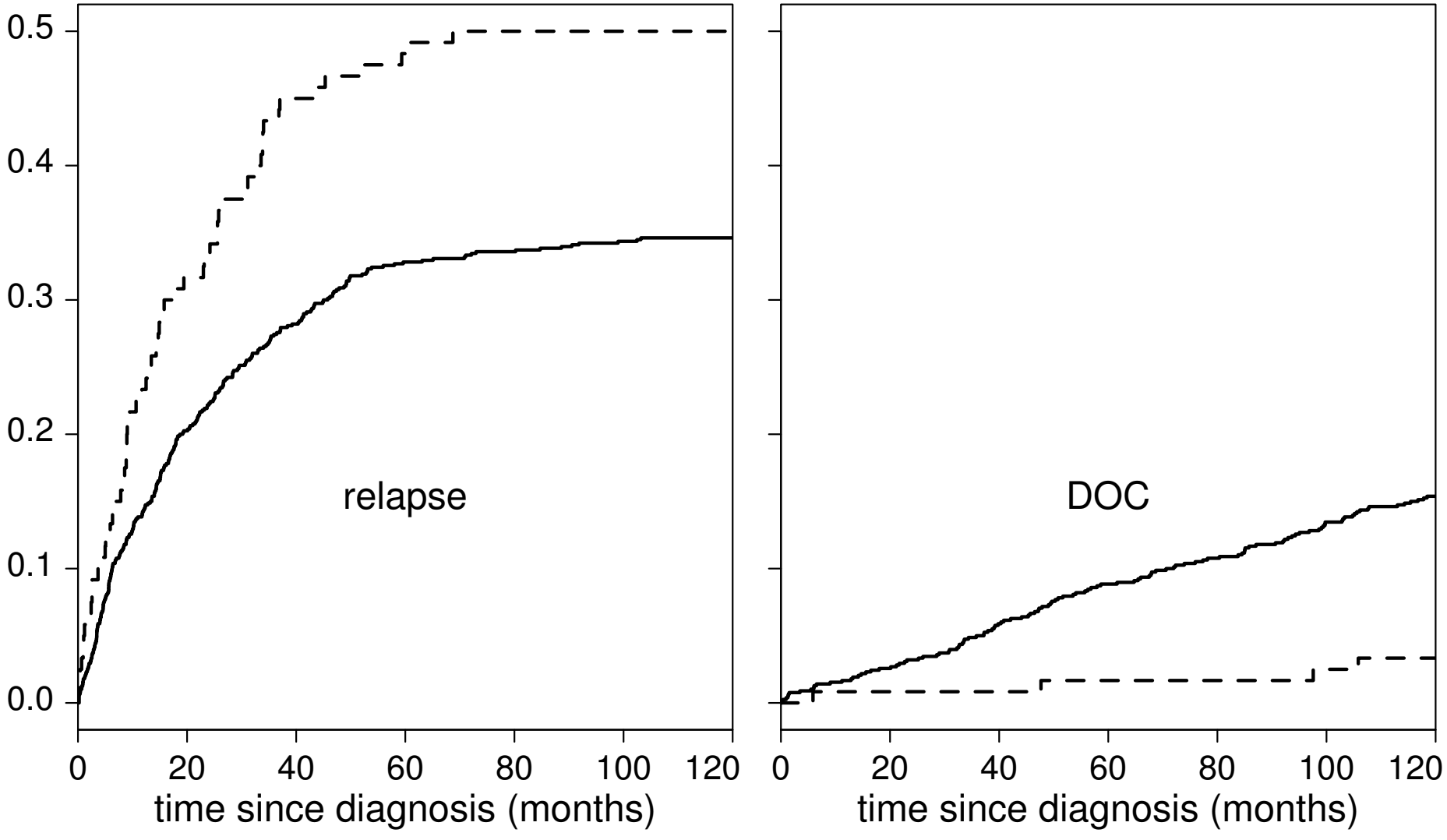
# IV: Bladder cancer; relapse, DOC competing



# IV: 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.
- Aalen-Johansen and Kaplan-Meier estimates almost equal because there is little mortality due to other causes, at least during the first 40 months.

# IV: Bladder cancer; relapse, death other causes (DOC) competing




Left: relapse. Solid: male; dashed: female

# IV: 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.
- Aalen-Johansen and Kaplan-Meier 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.

## IV: 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.
- Aalen-Johansen and Kaplan-Meier 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.

# The subdistribution approach

- Subdistribution hazard ($T_k = T \times I\{E = k\} + \infty \times I\{E \neq k\}$):

$$h_k(t) = \lim_{\Delta t \downarrow 0} \frac{1}{\Delta t} \times \mathrm{P}\{t \leq T_k < t + \Delta t \,|\, T_k \geq t\}$$

$$= \lim_{\Delta t \downarrow 0} \frac{\frac{1}{\Delta t}\mathrm{P}\{t \leq T < t + \Delta t, E = k\}}{\mathrm{P}\{T \geq t \text{ or } (T < t, E \neq k)\}}$$

- Denominator: event free or with earlier competing event
- Interpretation controversial
  - Not a rate in epidemiological sense, unless we can assume that those with the competing event were immune for the event of interest (cure model)
- One-to-one relation with crude risk

$$\overline{F_k}(t) = \prod_{t_l \leq t} \left\{1 - h_k(t_l)\right\} \quad \text{or} \quad \overline{F_k}(t) = \exp\left\{-\int_0^t h_k(u)du\right\}$$

# Observed data

$\{(x_1, e_1\delta_1), \ldots, (x_N, e_N\delta_N)\}$

- $x_i = \min\{t_i, c_i\}$, $\delta_i = \{t_i \leq c_i\}$, $e_i \in \{1, \ldots, K\}$
- $t_{(i)}$ ordered unique event times of any type
- $r(t_{(i)})$ number observed at risk
- $r^*(t_{(i)})$ number in risk set (for subdistribution hazard)
- $d_k(t_{(i)})$ number of events at $t_{(i)}$ of type $k$
- $d(t_{(i)})$ total number of events at $t_{(i)}$

## Subdistribution $\widehat{F}_k$: product-limit estimator

$$\widehat{F}_k^{\mathrm{PL}}(t) = \prod_{i:t_{(j)} \leq t} \left\{ 1 - \widehat{h}_k(t_{(j)}) \right\}$$

with

$$\widehat{h}_k(t_{(j)}) = \frac{d_k(t_{(j)})}{r^*(t_{(j)})}$$

**No censoring:** individuals with competing event remain in risk set forever. Small change in data, standard software

# Estimation, no censored 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 |

- Subdistribution. 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

## Subdistribution $\widehat{F}_k$: product-limit estimator

$$\widehat{F}_k^{\mathrm{PL}}(t) = \prod_{i:t_{(j)} \leq t} \left\{ 1 - \widehat{h}_k(t_{(j)}) \right\}$$

with

$$\widehat{h}_k(t_{(j)}) = \frac{d_k(t_{(j)})}{r^*(t_{(j)})}$$

**No censoring:** individuals with competing event remain in risk set forever. Small change in data, standard software

**Administrative censoring:** individuals with competing event leave risk set at their date of censoring. Small change in data, standard software

## Subdistribution $\widehat{F}_k$: product-limit estimator

$$\widehat{F}_k^{\mathrm{PL}}(t) = \prod_{i:t_{(j)} \le t} \left\{ 1 - \widehat{h}_k(t_{(j)}) \right\}$$

with

$$\widehat{h}_k(t_{(j)}) = \frac{d_k(t_{(j)})}{r^*(t_{(j)})}$$

**No censoring:** individuals with competing event remain in risk set forever. Small change in data, standard software

**Administrative censoring:** individuals with competing event leave risk set at their date of censoring. Small change in data, standard software

**General censoring:** Estimate time-to-censoring distribution, then

- multiply impute censoring times for those with competing event
- reweight them by probability to remain uncensored

## Right Censored Data

$$\widehat{h_k}(t_{(i)}) = \frac{d_k(t_{(i)})}{r^*(t_{(i)})}$$

Contribution $\omega_l(t_{(i)})$ of individual $l$ to the risk set $r^*(t_{(i)})$ is:

- censored or event of type $k$ before $t_{(i)}$: 0

## Right Censored Data

$$\widehat{h_k}(t_{(i)}) = \frac{d_k(t_{(i)})}{r^*(t_{(i)})}$$

Contribution $\omega_l(t_{(i)})$ of individual $l$ to the risk set $r^*(t_{(i)})$ is:

- censored or event of type $k$ before $t_{(i)}$: 0
- still at risk at $t_{(i)}$: 1

## Right Censored Data

$$\widehat{h_k}(t_{(i)}) = \frac{d_k(t_{(i)})}{r^*(t_{(i)})}$$

Contribution $\omega_l(t_{(i)})$ of individual $l$ to the risk set $r^*(t_{(i)})$ is:

- censored or event of type $k$ before $t_{(i)}$: 0
- still at risk at $t_{(i)}$: 1
- competing event at $t_{(j)}$ before $t_{(i)}$:

estimate of $\mathrm{P}\{C \geq t_{(i)} | C \geq t_{(j)}\}$ :

# Right Censored Data

$$\widehat{h_k}(t_{(i)}) = \frac{d_k(t_{(i)})}{r^*(t_{(i)})}$$

Contribution $\omega_l(t_{(i)})$ of individual $l$ to the risk set $r^*(t_{(i)})$ is:

- censored or event of type $k$ before $t_{(i)}$: 0
- still at risk at $t_{(i)}$: 1
- competing event at $t_{(j)}$ before $t_{(i)}$:

$$\text{estimate of } \mathrm{P}\{C \geq t_{(i)} | C \geq t_{(j)}\} : \widehat{\overline{\Gamma}}(t_{(i)}-)/\widehat{\overline{\Gamma}}(t_{(j)}-)$$

## Right Censored Data

$$\widehat{h_k}(t_{(i)}) = \frac{d_k(t_{(i)})}{r^*(t_{(i)})}$$

Contribution $\omega_l(t_{(i)})$ of individual $l$ to the risk set $r^*(t_{(i)})$ is:

- censored or event of type $k$ before $t_{(i)}$: 0
- still at risk at $t_{(i)}$: 1
- competing event at $t_{(j)}$ before $t_{(i)}$:

$$\text{estimate of } \mathrm{P}\{C \geq t_{(i)} | C \geq t_{(j)}\} : \widehat{\bar{\Gamma}}(t_{(i)}-)/\widehat{\bar{\Gamma}}(t_{(j)}-)$$

- $\widehat{\bar{\Gamma}}$: reverse role of event time $T_i$ and censoring $C_i$:

$$\widehat{\bar{\Gamma}}(t) = \prod_{j:c_{(j)} \leq t} \left\{ 1 - \frac{m_j}{r(c_{(j)})} \right\}$$

$m_j$: number of censorings at $c_{(j)}$

# Subdistribution $\widehat{F}_k$: ECDF estimator

- Without censoring

$$\widehat{F}_k^{\mathrm{EC}}(t) = \frac{\#\{t_i \leq t\}}{N} = \sum_{t_{(j)} \leq t} \frac{d_k(t_{(j)})}{N}.$$

# Estimation, no censored 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 |

- Subdistribution. 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

# Subdistribution $\widehat{F}_k$: ECDF estimator

- Without censoring

$$\widehat{F}_k^{\mathrm{EC}}(t) = \frac{\#\{t_i \leq t\}}{N} = \sum_{t_{(j)} \leq t} \frac{d_k(t_{(j)})}{N}.$$

- Right censored data

$$\widehat{F}_k^{\mathrm{EC}}(t) = \frac{1}{N} \sum_{t_{(j)} \leq t} \frac{d_k(t_{(j)})}{\widehat{\overline{\Gamma}}(t_{(j)}-)}$$

# Equivalence

If $\widehat{\overline{\Gamma}}$ based on the PL-form, then we have

$$\widehat{F}_k^{\,\mathrm{AJ}} \equiv \widehat{F}_k^{\,\mathrm{PL}} \equiv \widehat{F}_k^{\,\mathrm{EC}}$$

(Geskus 2011, Biometrics)

$$\widehat{F}_k^{\,\mathrm{AJ}}(t) = \sum_{i:t_{(i)} \le t} \widehat{\overline{F}}^{\mathrm{PL}}(t_{(i)}-) \times \widehat{\lambda}_k(t_{(i)})$$

$$\widehat{\overline{F}_k}^{\mathrm{PL}}(t) = \prod_{i:t_{(j)} \le t} \left\{ 1 - \widehat{h}_k(t_{(j)}) \right\}$$

$$\widehat{F}_k^{\,\mathrm{EC}}(t) = \frac{1}{N} \sum_{t_{(j)} \le t} \frac{d_k(t_{(j)})}{\widehat{\overline{\Gamma}}(t_{(j)}-)}$$

# Estimators in competing risks setting

| hazard | estimate | estimate cumulative incidence |
|---|---|---|
| marginal | $\left\{\widehat{\lambda}(t) = d_k(t)/r(t)\right\}$ | $\left\{\widehat{\overline{F^m}}^{\mathrm{PL}}(t) = \prod_{t_{(j)} \le t}\left[1 - \widehat{\lambda}(t_{(j)})\right]\right\}$ |
| cause-specific | $\widehat{\lambda_k}(t) = d_k(t)/r(t)$ | $\widehat{F_k}^{\mathrm{AJ}}(t) = \sum_{t_{(j)} \le t} \widehat{\overline{F}}^{\mathrm{PL}}(t_{(j)}-)\widehat{\lambda_k}(t_{(j)})$ |
| subdistribution | $\widehat{h_k}(t) = d_k(t)/r^*(t)$ | $\widehat{\overline{F_k}}^{\mathrm{PL}}(t) = \prod_{t_{(j)} \le t}\left[1 - \widehat{h_k}(t_{(j)})\right]$ |
| | | $\widehat{F_k}^{\mathrm{EC}}(t) = \frac{1}{N}\sum_{t_{(j)} \le t} \frac{d_k(t_{(j)})}{\widehat{\overline{\Gamma}}(t_{(j)}-)}$ |
| overall | $\widehat{h}(t) = d(t)/r(t)$ | $\widehat{\overline{F}}^{\mathrm{PL}}(t) = \prod_{t_{(j)} \le t}\left[1 - \widehat{h}(t_{(j)})\right]$ |

# Outline

**Competing Risks; When**

**Competing Risks; How**

- Main Quantities
- Nonparametric Estimation
- Regression

**Which Approach to Choose?**

- Marginal versus competing risks
- Which competing risks approach to choose?

# Cause-specific hazard

- Completely standard, e.g. Cox model

$$\lambda_k(t \mid \mathbf{Z}_i) = \lambda_{k,0}(t) \exp(\beta_k^\top \mathbf{Z}_i).$$

  - Censor individuals with a competing event
  - SE completely standard as well
  - Interpretation is different: cause-specific event rate among event-free individuals
  - Not a marginal hazard, unless competing risks independent conditionally on covariables

# Subdistribution hazard

- Proportional subdistribution hazards (Fine and Gray model)

$$h_k(t \mid \mathbf{Z}_i) = h_{k,0}(t) \exp(\beta_k^\top \mathbf{Z}_i) .$$

  - Estimation: those with competing event remain in risk set, with weight reflecting probability to remain uncensored
  - SE completely standard
    Simulation study (Geskus 2011, Biometrics): sandwich estimator as suggested by Fine and Gray performed worse
  - Standard software with time-varying individual weights
  - Interpretation difficult, but direct relation with cause-specific cumulative incidence

# Cause-specific cumulative incidence

- (Proportional odds) logistic regression

$$\text{logit}[\text{P}(T_k \leq t \mid \mathbf{Z}_i)] = \log[\pi_{k,0}(t)] + \beta_k^\top \mathbf{Z}_i$$

translates to the cause-specific cumulative incidence as

$$\text{P}(T_k \leq t \mid \mathbf{Z}_i) = \frac{\pi_{k,0}(t)\exp(\beta_k^\top \mathbf{Z}_i)}{1+\pi_{k,0}(t)\exp(\beta_k^\top \mathbf{Z}_i)}.$$

- Estimation: estimating equations, with weights reflecting probability to remain uncensored
  - Scheike *et al.* (2008): event weighted as in $\widehat{F}_k^{\text{EC}}$
  - Eriksson *et al.* (2015): weights as in $\widehat{h}_k$
  - Blanche *et al.* (2023): individuals weighted as in $\widehat{F}_k^{\text{EC}}$
- SE based on sandwich estimator

# Outline

Competing Risks; When

Competing Risks; How

**Which Approach to Choose?**

- Marginal versus competing risks
- Which competing risks approach to choose?

# Outline

**Competing Risks; When**

**Competing Risks; How**

- Main Quantities
- Nonparametric Estimation
- Regression

**Which Approach to Choose?**

- Marginal versus competing risks
- Which competing risks approach to choose?

# Marginal distribution

- Often via (marginal) hazard, basis for Kaplan-Meier estimate of cumulative incidence/net risk
- Assumption: Censored individuals (including due to competing risk) can be represented by the ones that remain at risk. Being censored should give no information on residual time to event

# Marginal distribution

- Often via (marginal) hazard, basis for Kaplan-Meier estimate of cumulative incidence/net risk
- Assumption: Censored individuals (including due to competing risk) can be represented by the ones that remain at risk. Being censored should give no information on residual time to event
- Otherwise Kaplan-Meier has no meaning.
  Does not describe survival in (hypothetical) world with competing event removed, unless we know that censoring is independent
- Extra information may allow to show informative/dependent censoring (IDU and pre-AIDS death), but independence can never be tested for

# Competing risks

- Competing risk is a separate event
  - Individuals censored by competing event don't have to be represented by the ones that remain at risk.
    Other censoring (administrative/loss to follow-up) must be independent

# Competing risks

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

# Competing risks

- Competing risk is a separate event
  - Individuals censored by competing event don't have to be represented by the ones that remain at risk.
    Other censoring (administrative/loss to follow-up) 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 a separate event
  - Individuals censored by competing event don't have to be represented by the ones that remain at risk.
    Other censoring (administrative/loss to follow-up) 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
- 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 difficult to perform
- Example IV: bladder cancer relapse and DOC
  Marginal interpretation, but does it make sense?
  Can we eliminate DOC? Is their mechanism completely separate from mechanism leading to relapse?

# Outline

**Competing Risks; When**

**Competing Risks; How**

- Main Quantities
- Nonparametric Estimation
- Regression

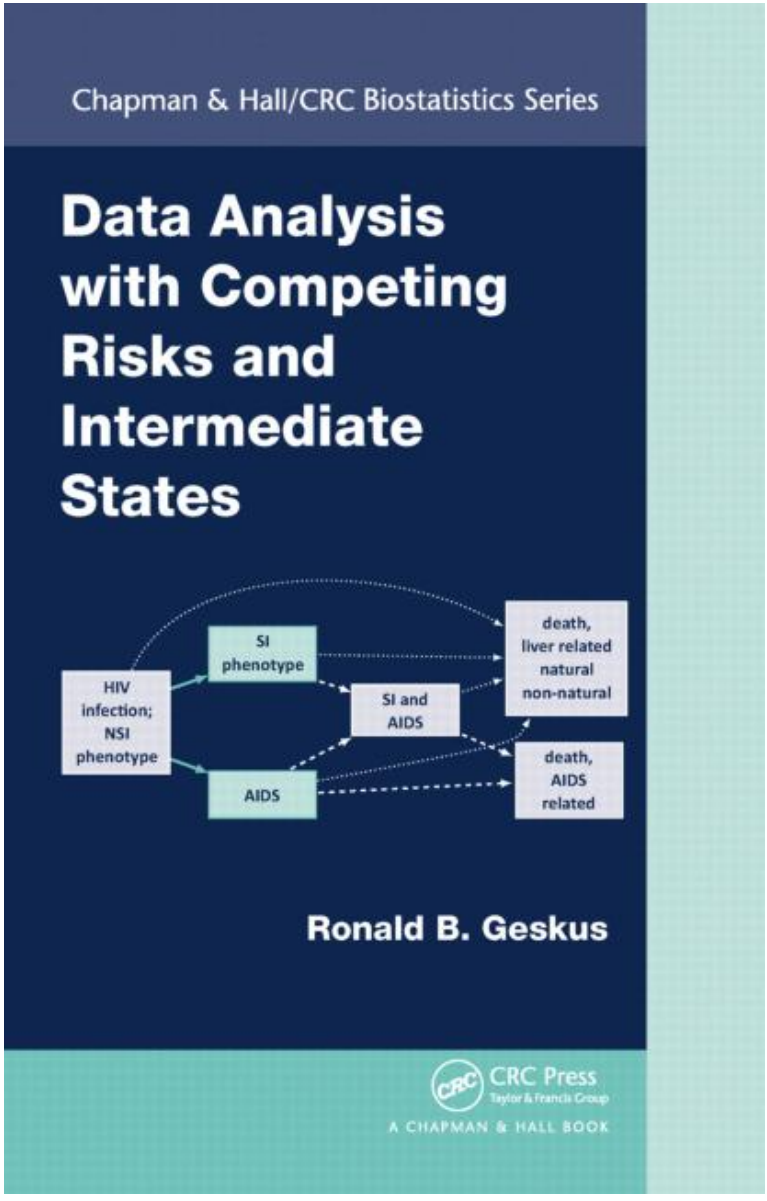
**Which Approach to Choose?**

- Marginal versus competing risks
- Which competing risks approach to choose?

# Cause-specific hazard, subdistribution hazard or crude risk?

- Cause-specific may often be closer to causal effects, but any hazard is problematic as causal estimand
- Cumulative scale combination of etiology and effect on other event types.
  “In the end we all die, but cause of death may be different”
  “Impact” instead of “effect”
- Cumulative estimate final aim for prediction, but each may serve as basis for model
- Subdistribution hazard directly relates to cumulative scale
  Regression: impact on cumulative scale via parameters
- Some suggest to quantify both hazards

# Published by CRC Press, 2015



# THANKS!

# References I

Blanche PF, Holt A, Scheike T. (2023).

On logistic regression with right censored data, with or without competing risks, and its use for estimating treatment effects.

*Lifetime Data Analysis* 29(2):441–482

Eriksson F, Li J, Scheike T, Zhang MJ. (2015).

The proportional odds cumulative incidence model for competing risks.

*Biometrics* 71(3):687–695

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.

R. B. Geskus (2024).

Competing Risks: Concepts, Methods and Software.

*Annual Reviews of Statistics and its Application*

Scheike TH, Zhang MJ, Gerds TA. (2008).

Predicting cumulative incidence probability by direct binomial regression.

*Biometrika* 95(1):205–220

oucru

wellcome

UNIVERSITY OF OXFORD