SESSION 2: HAZARD FUNCTIONS, COMPETING RISKS, CAUSE-SPECIFIC HAZARDS, AND AND CUMULATIVE INCIDENCE

Module 20: Survival Analysis for Observational Data

Summer Institute in Statistics for Clinical Research
University of Washington
July, 2017

Barbara McKnight, Ph.D.

OUTLINE

- Nonparametric estimation of hazard functions
- Competing risks:
  - Definition: when there is more than one cause of death/failure
  - Cumulative functions:
    - Event-free survival
    - Cumulative Incidence estimator
  - Cause-specific and sub-distribution hazards
  - Fine-Gray and Cox regression models
  - Interpretation subtleties

OUTLINE

- **Nonparametric estimation of hazard functions**
- Competing risks:
  - Definition: when there is more than one cause of death/failure
  - Cumulative functions:
    - Event-free survival
    - Cumulative Incidence estimator
  - Cause-specific and sub-distribution hazards
  - Fine-Gray and Cox regression models
  - Interpretation subtleties

HAZARD FUNCTION

$$\lambda(t) = \lim_{\Delta t \to 0} \frac{1}{\Delta t} \Pr[t \leq T < t + \Delta t | T \geq t]$$

- Instantaneous rate at which death occurs at t in those who are alive at t
- Examples:
  - Age-specific death rate
  - Age-specific disease incidence rate

CUMULATIVE HAZARD FUNCTION

$$\Lambda(t) = \int_0^t \lambda(s)ds$$

= area under the hazard function curve between 0 and t.

= amount of “hazard" accumulated between 0 and t.

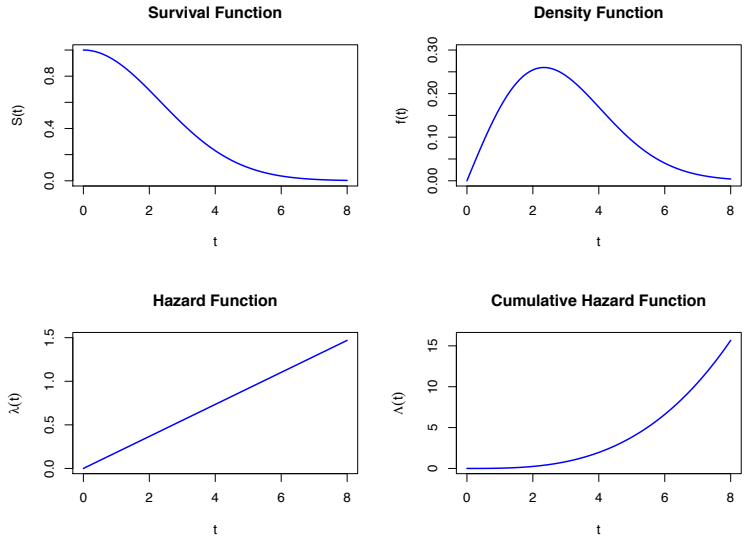
$= -\log(S(t))$

Not usually of interest *per se*, but estimates useful for diagnostics.

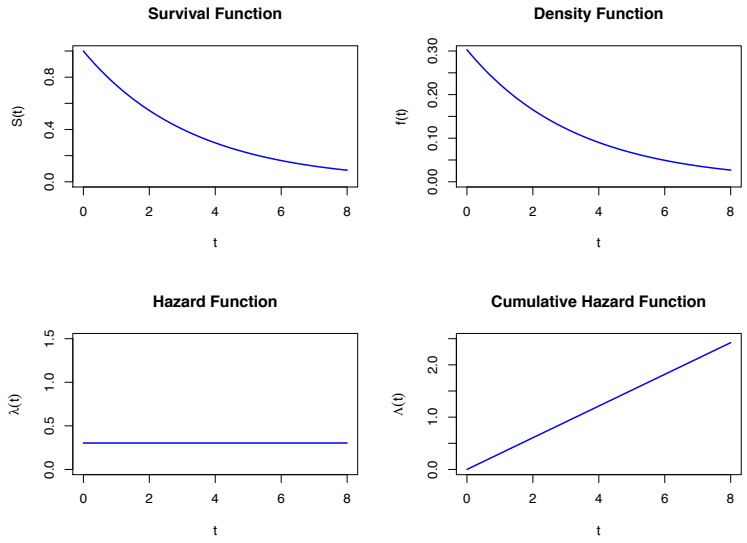
EQUIVALENT CHARACTERIZATIONS

- Any <u>one</u> of these four functions is enough to determine the survival distribution.
- They are each functions of each other:
  - $S(t) = \int_t^\infty f(s)ds = e^{-\int_0^t \lambda(s)ds}$
  - $f(t) = -\frac{d}{dt}S(t) = \lambda(t)e^{-\int_0^t \lambda(s)ds}$
  - $\lambda(t) = \frac{f(t)}{S(t)}$
  - $\Lambda(t) = \int_0^t \lambda(s)ds = -\log(S(t))$

EQUIVALENT CHARACTERIZATIONS



EQUIVALENT CHARACTERIZATIONS



CUMULATIVE HAZARD

- Nelson - Aalen estimator:

  $\hat{\Lambda}(t) = \sum_{j:t_{(j)} \le t} \frac{D_{(j)}}{N_{(j)}}$

- Variance:

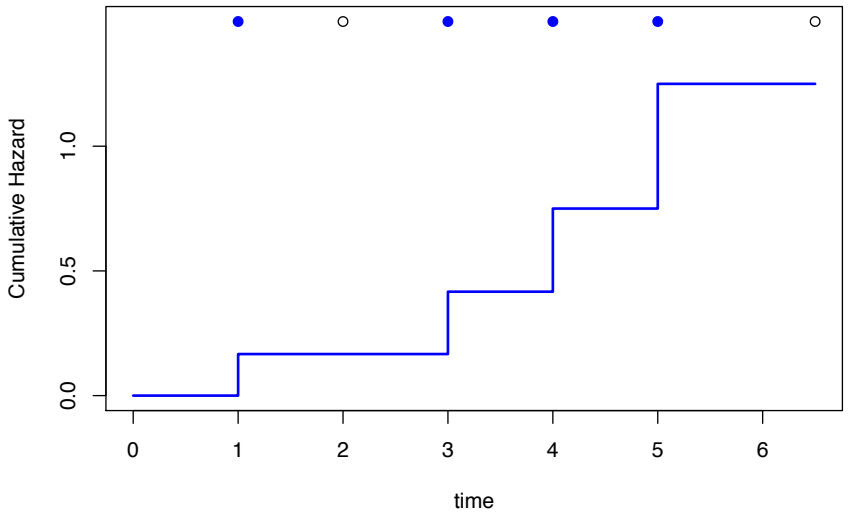
  $\widehat{Var}(\hat{\Lambda}(t)) = \sum_{j:t_{(j)} \le t} \frac{D_{(j)} S_{(j)}}{[N_{(j)}]^3}$

- Standard error:

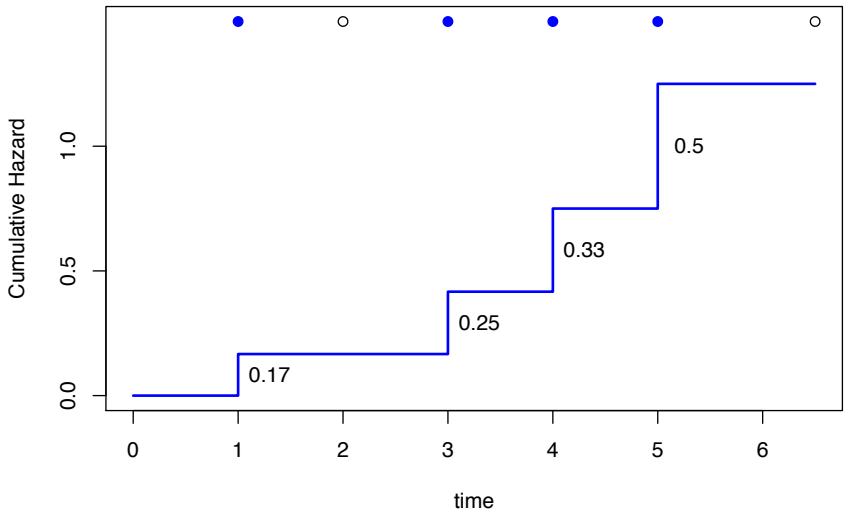
  $\sqrt{\widehat{Var}(\hat{\Lambda}(t))}$ can be used to form pointwise CI's.

- or could use $\hat{\Lambda}(t) = -\log(\hat{S}(t))$

CUMULATIVE HAZARD FUNCTION



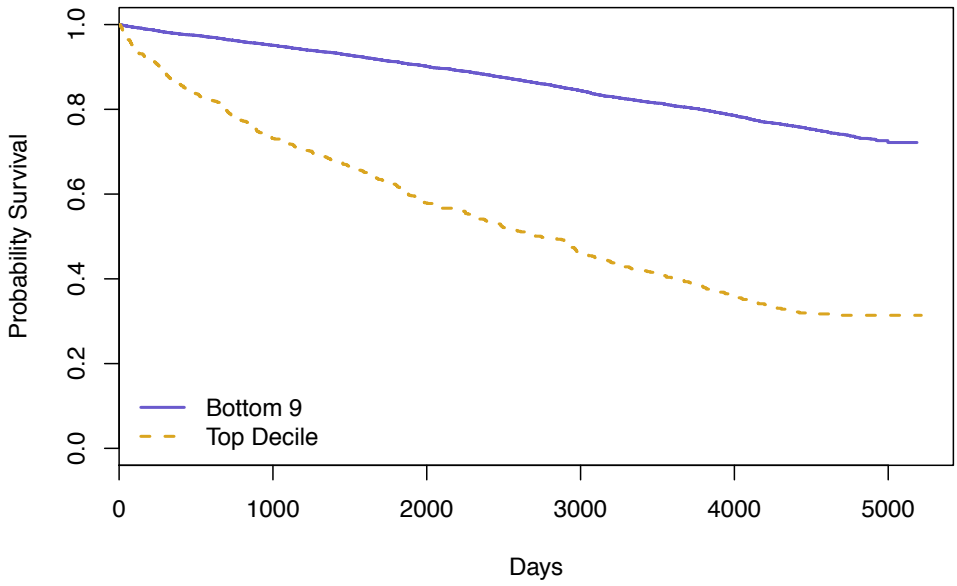
CUMULATIVE HAZARD FUNCTION



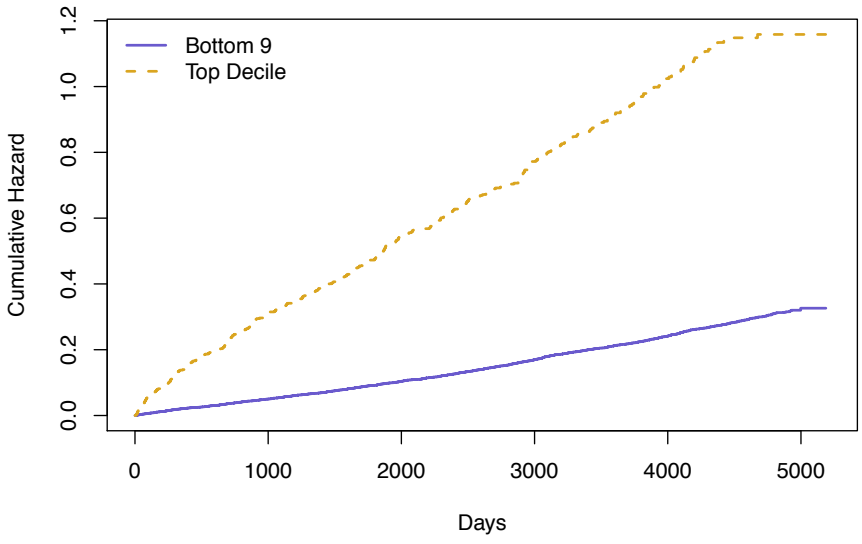
SURVIVAL AND IG

- Random subset of the data from A. Dispenzieri, J. Katzmann, R. Kyle, D. Larson, T. Therneau, C. Colby, R. Clark, G. Mead, S. Kumar, L.J. Melton III, and S.V. Rajkumar. Use of monclonal serum immunoglobulin free light chains to predict overall survival in the general population. Mayo Clinic Proc, 87:512–523, 2012.
- Are high free-chain Ig levels associated with survival?
  - Population-based Olmstead County example
  - Men and women 50+ years of age

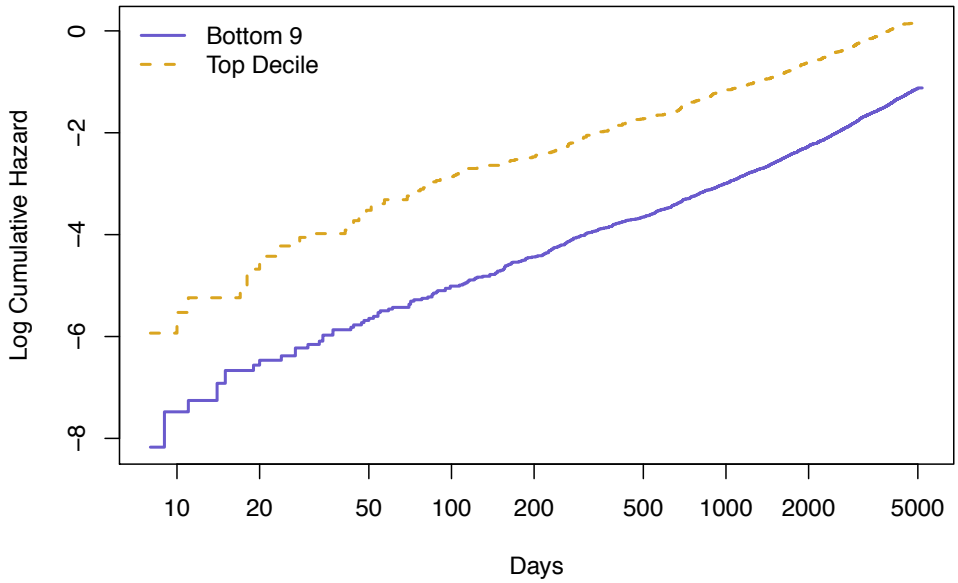
TOP DECILE FLC



FLC EXAMPLE



FLC EXAMPLE



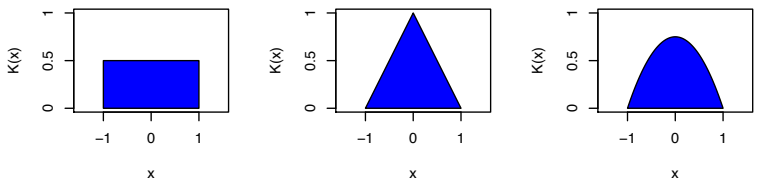
HAZARD FUNCTION

IDEA:

At each time t, let estimate of $\lambda(t)$ be a weighted average of jumps in $\hat{\Lambda}(t)$ at nearby times.

Steps:

1. Choose a "bandwidth" $\pm b$ outside of which observations are not averaged.
2. Choose a "kernel" or weight function $K(\cdot)$.

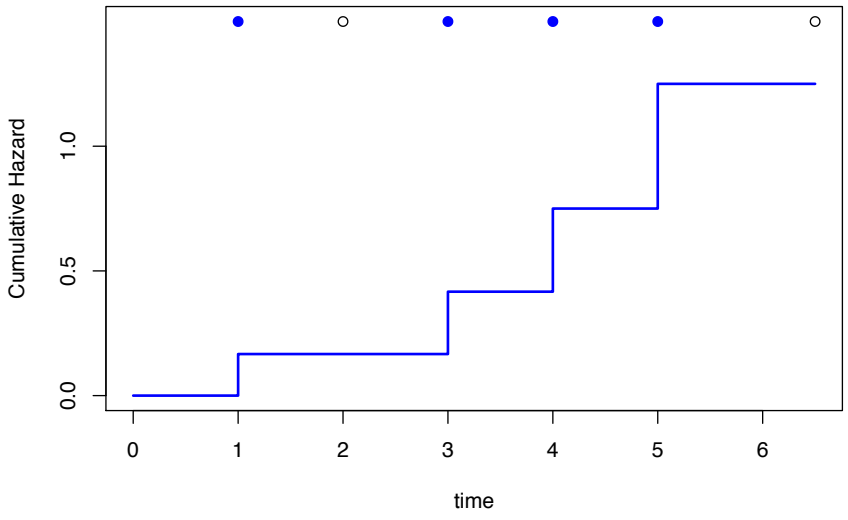


HAZARD FUNCTION

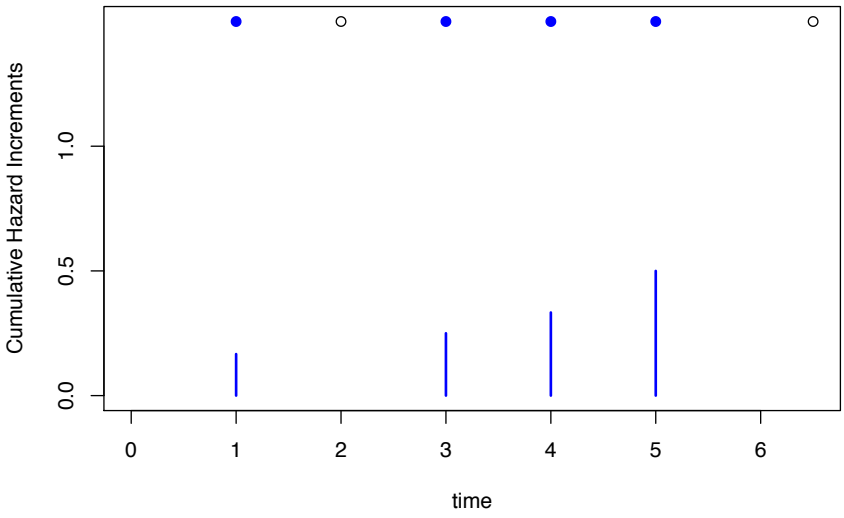
3. Calculate the estimate:

- $\hat{\lambda}(t) = \frac{1}{b}\sum_{j=1}^{J} K(\frac{t-t_{(j)}}{b})\frac{D_{(j)}}{N_{(j)}}$
- $se(\hat{\lambda}(t)) = \frac{1}{b}\left\{\sum_{j=1}^{J} K^2(\frac{t-t_{(j)}}{b})\frac{D_{(j)}}{N_{(j)}^2}\right\}^{\frac{1}{2}}$

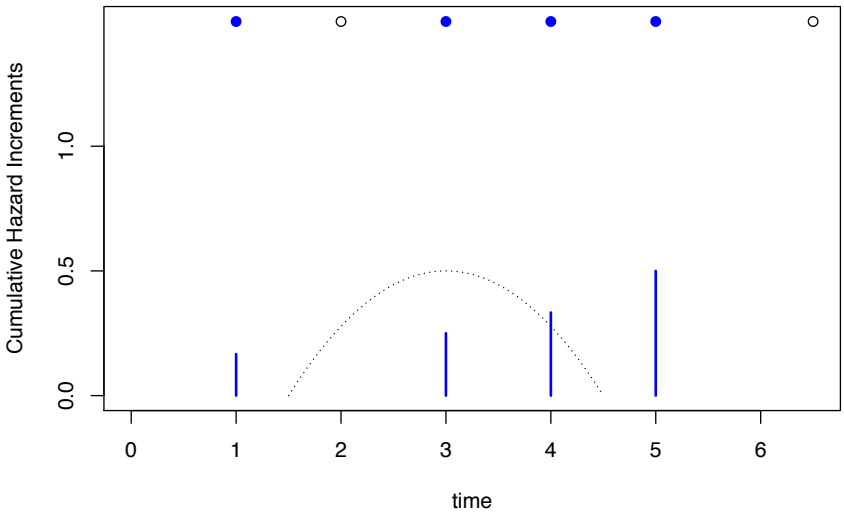
CUMULATIVE HAZARD FUNCTION



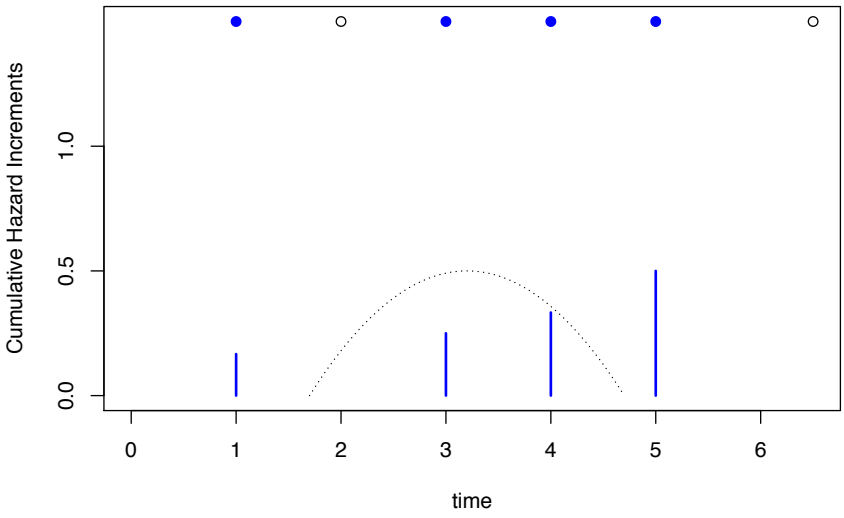
KERNEL HAZARD ESTIMATE



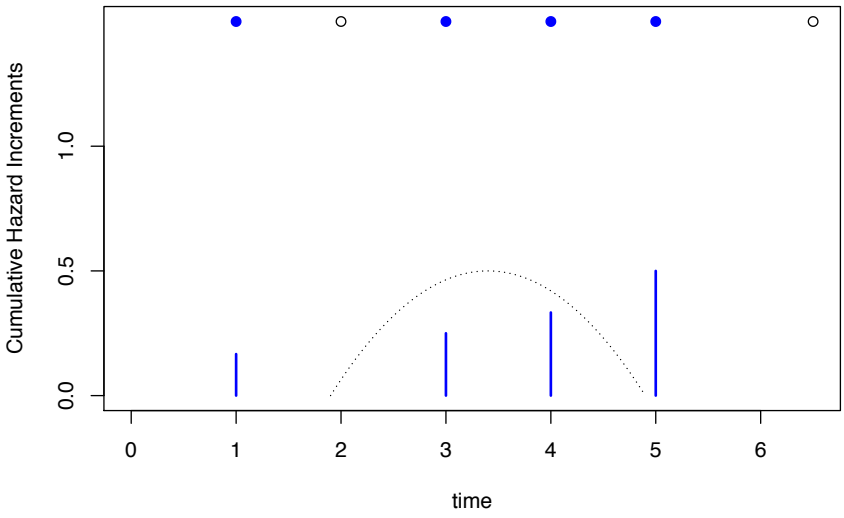
KERNEL HAZARD ESTIMATE



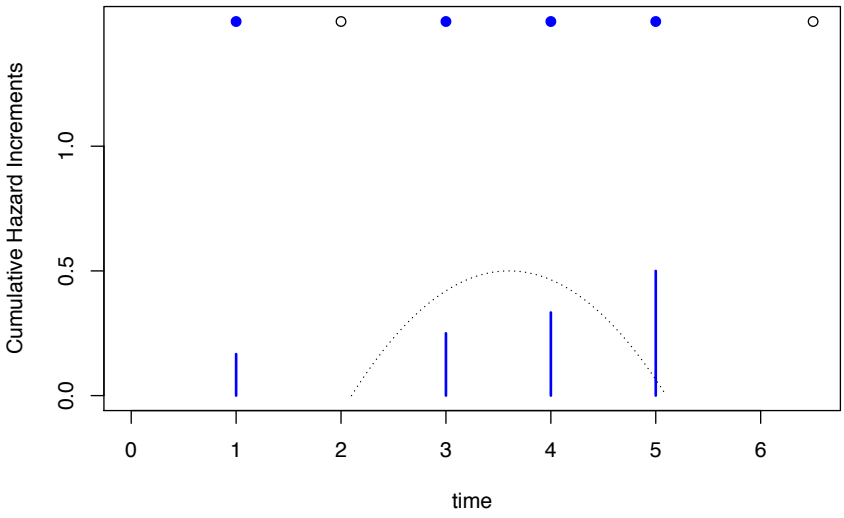
KERNEL HAZARD ESTIMATE



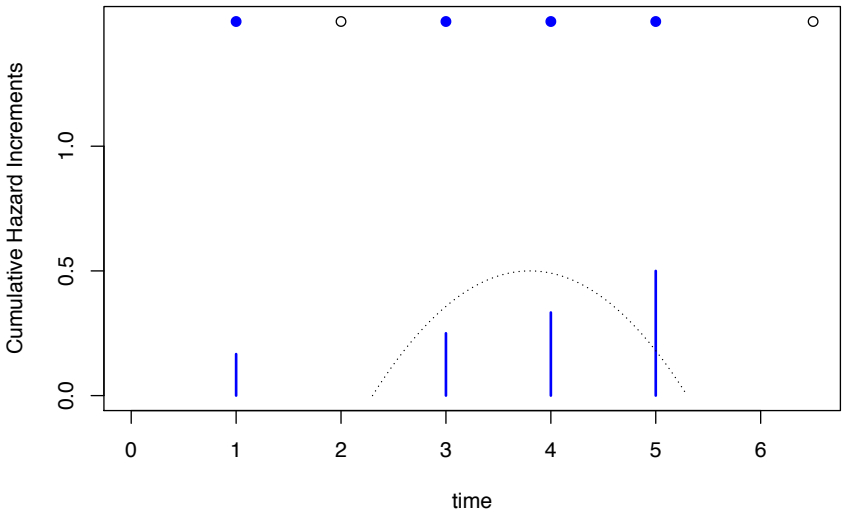
KERNEL HAZARD ESTIMATE



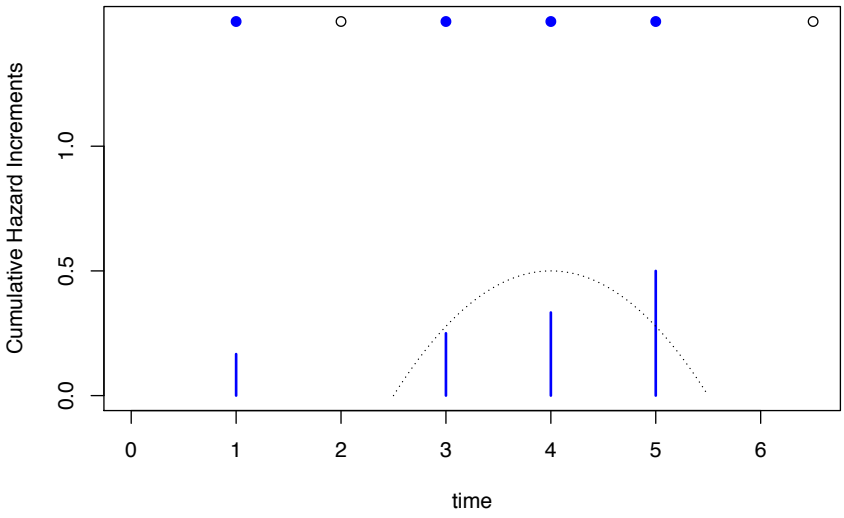
KERNEL HAZARD ESTIMATE



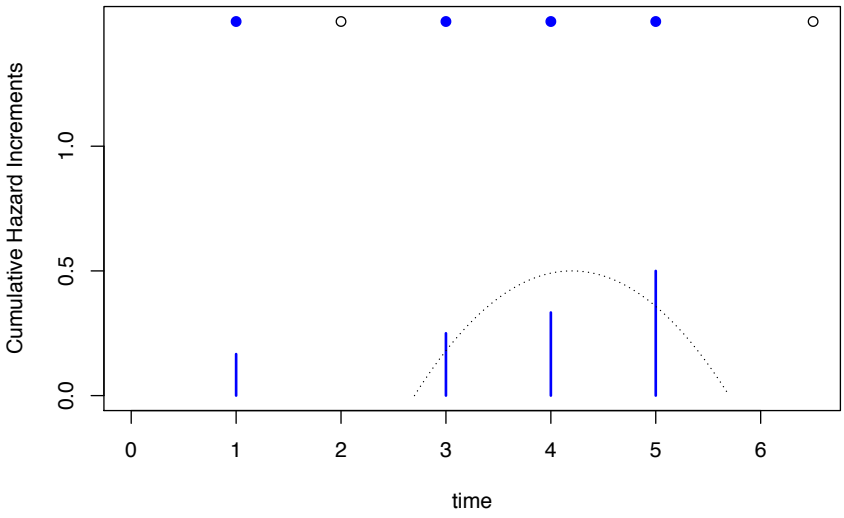
KERNEL HAZARD ESTIMATE



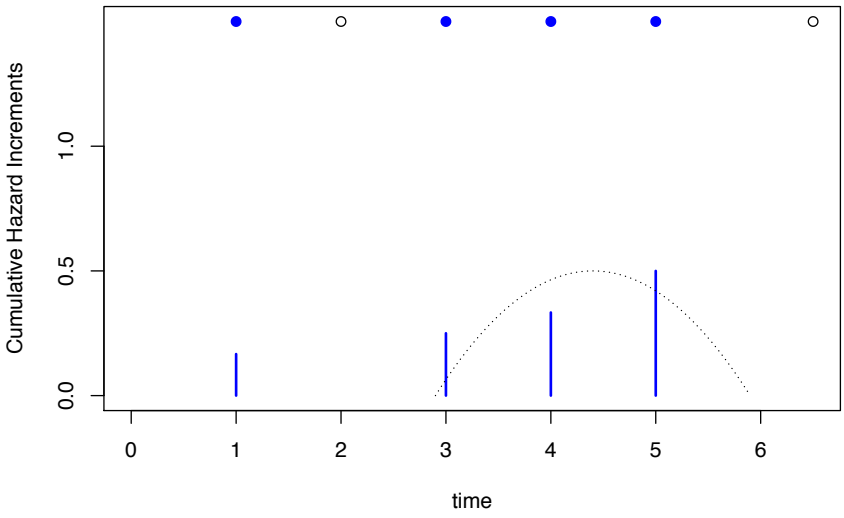
KERNEL HAZARD ESTIMATE



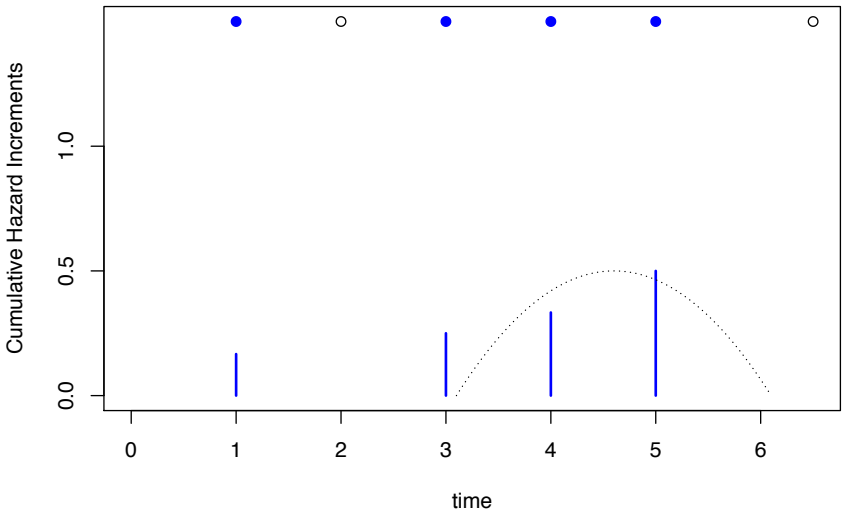
KERNEL HAZARD ESTIMATE



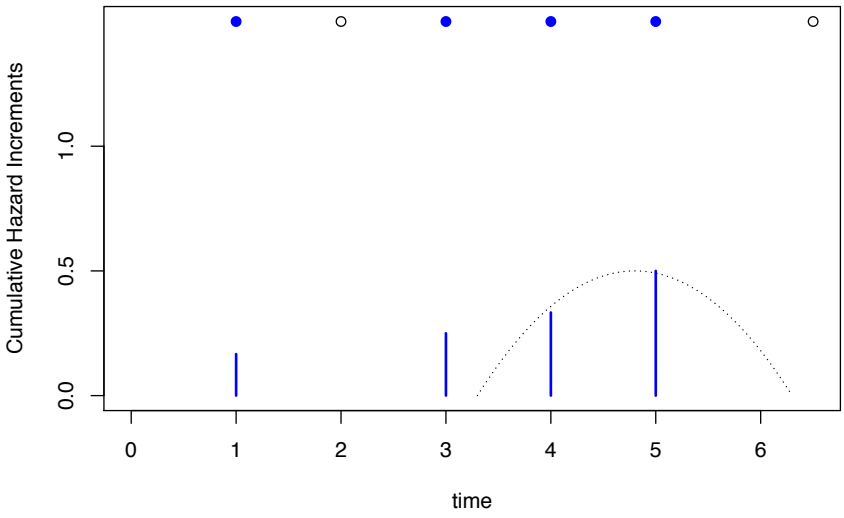
KERNEL HAZARD ESTIMATE



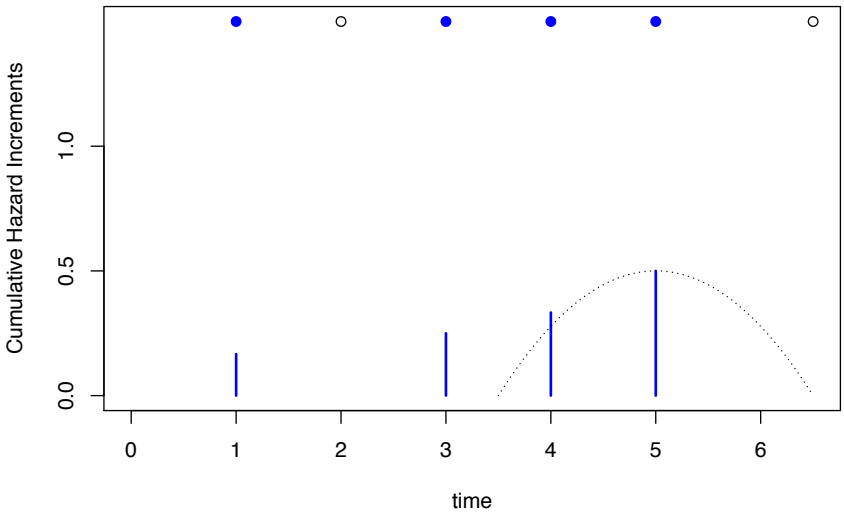
KERNEL HAZARD ESTIMATE



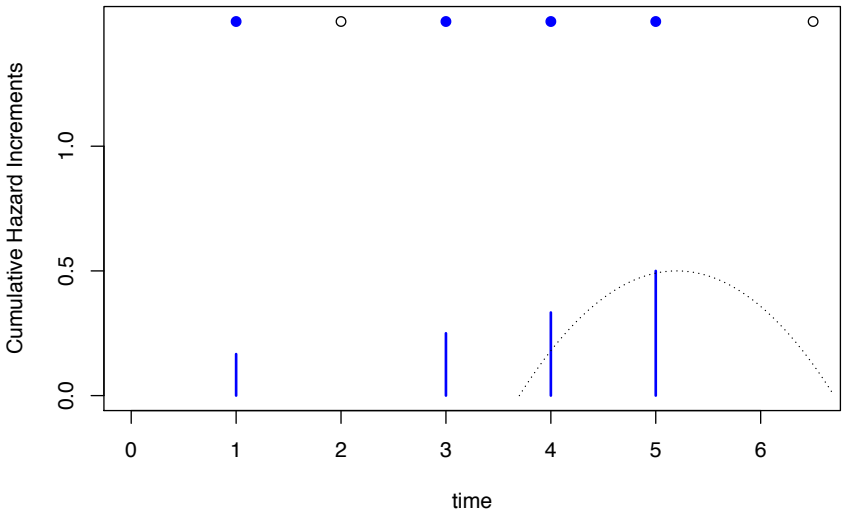
KERNEL HAZARD ESTIMATE



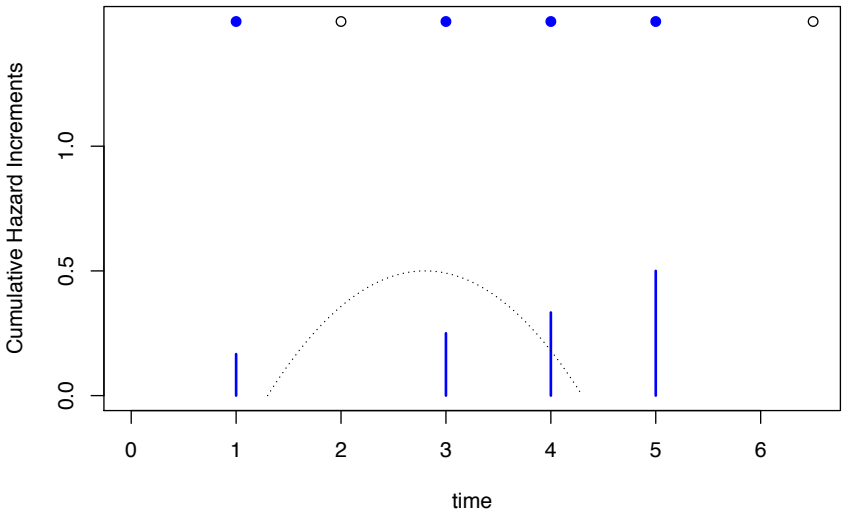
KERNEL HAZARD ESTIMATE



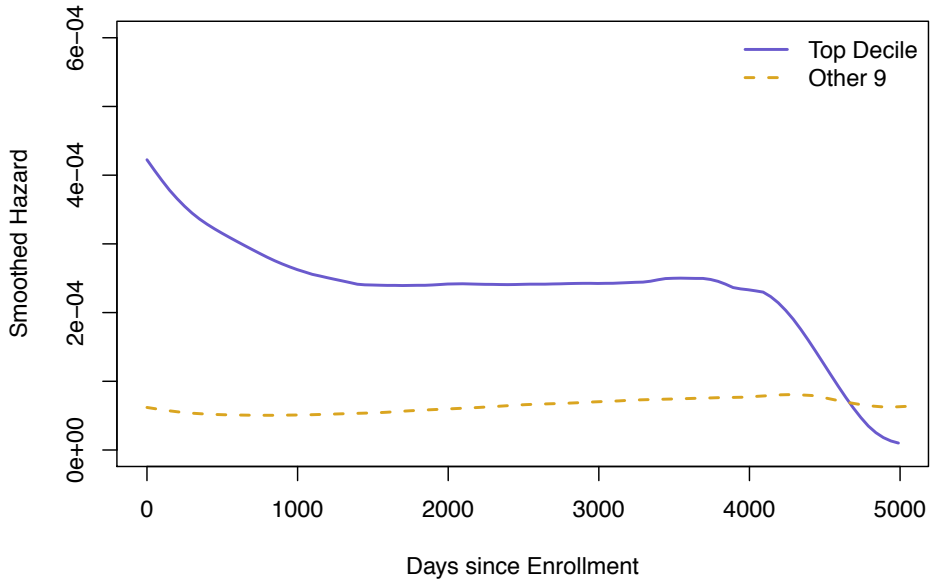
KERNEL HAZARD ESTIMATE



KERNEL HAZARD ESTIMATE



FLC EXAMPLE



OUTLINE

- Nonparametric estimation of hazard functions
- **Competing risks:**
  - **Definition: when there is more than one cause of death/failure**
  - **Cumulative functions:**
    - **Event-free survival**
    - **Cumulative Incidence estimator**
  - Cause-specific and sub-distribution hazards
  - Fine-Gray and Cox regression models
  - Interpretation subtleties

COMPETING RISKS

- When there is more than one cause of failure:
  - Recurrence or death before recurrence
  - MI, stroke, PE or death from other causes
- The different types of failure are called "competing risks".
  - They "compete" to be the first to make subjects experience an event

MONOCLONAL GAMMOPATHY

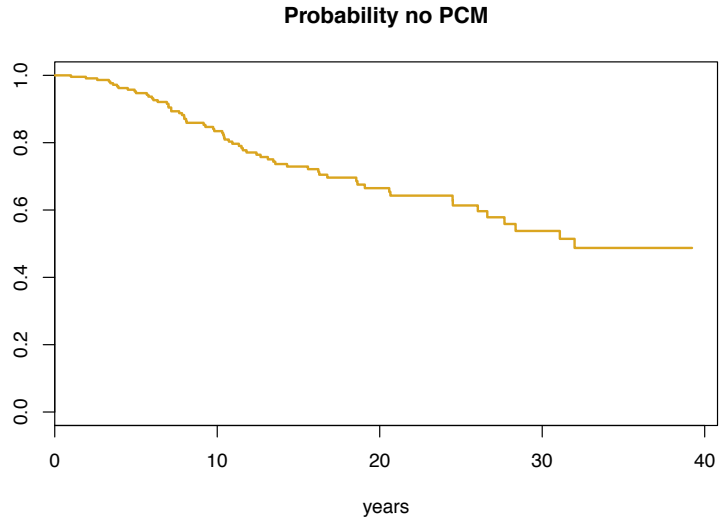
- 241 Mayo Clinic Patients (Monoclonal Gammopathy of Undetermined Significance)
- 20-40 years of follow-up after Dx
- 64 developed plasma cell malignancy (PCM), 163 died without it.
- PCM and death without PCM are competing risks

R Kyle, Benign monoclonal gammopathy – after 20 to 35 years of follow-up, Mayo Clinic Proc 1993; 68:26-36

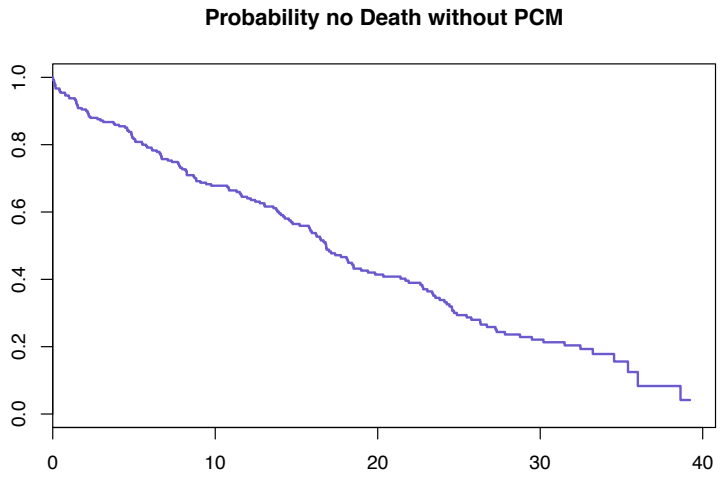
MONOCLONAL GAMMOPATHY

- In situations like this, it has been common practice to apply the KM method to estimate "survival" functions:
  - Probability of avoiding PCM over time
  - Probabilty of avoiding death w/o PCM
- For PCM curve, treat deaths w/o PCM as censored
- For death w/o PCM, treat PCMs as censored
- Conceptual difficulties in describing distribution of the time of an event that may not occur.

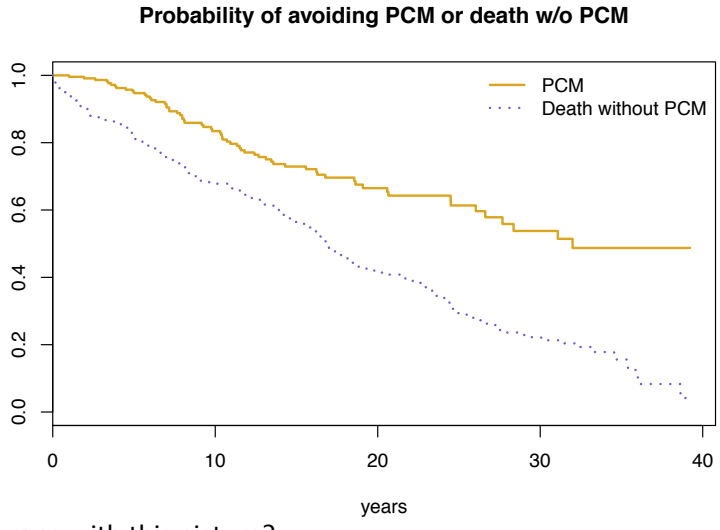
MONOCLONAL GAMMOPATHY



MONOCLONAL GAMMOPATHY

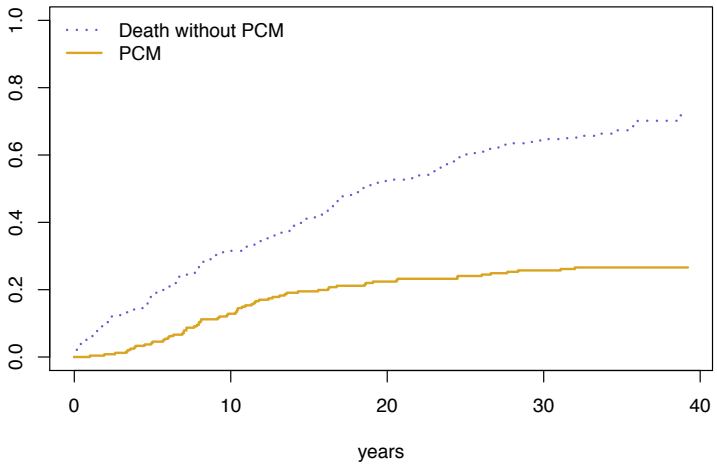


MONOCLONAL GAMMOPATHY

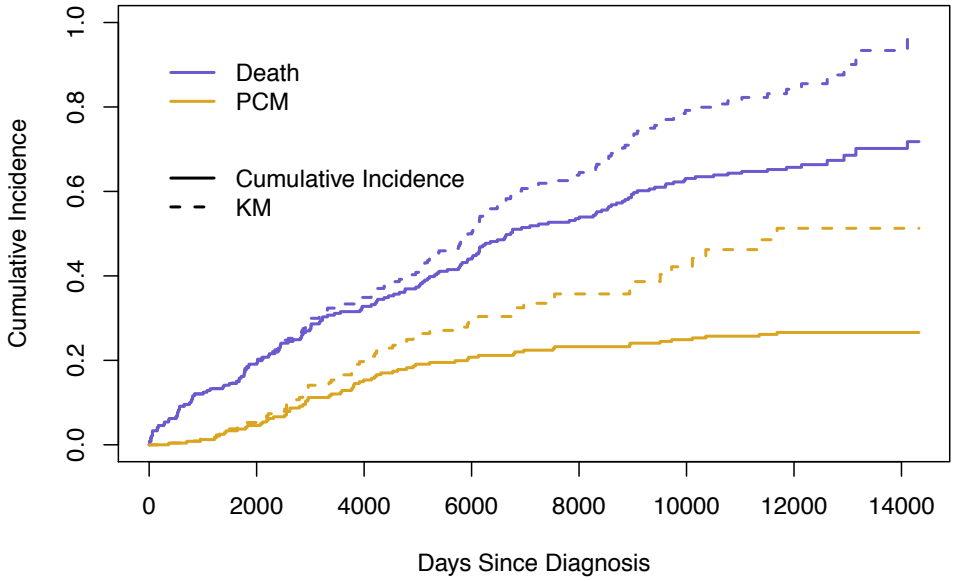


What is wrong with this picture?

CUMULATIVE INCIDENCE



CUMULATIVE INCIDENCE



ESTIMATING CUMULATIVE INCIDENCE

- We can write

$$1 - \hat{S}^{(k)}(t) = \sum_{j:t_{(j)} \le t} \frac{D_{(j)}^{(k)}}{N_{(j)}} \hat{S}^{(k)}(t_{(j-1)})$$

- At the second failure time of type k,

$$1 - \hat{S}^{(k)}(t_{(2)}) = 1 - \frac{N_{(1)} - D_{(1)}^{(k)}}{N_{(1)}} \cdot \frac{N_{(2)} - D_{(2)}^{(k)}}{N_{(2)}} = \frac{D_{(1)}^{(k)}}{N_{(1)}} + \frac{D_{(2)}^{(k)}}{N_{(2)}} \cdot \frac{N_{(1)} - D_{(1)}^{(k)}}{N_{(1)}}$$

- If any failures of another type have occurred between $t_{(1)}$ and $t_{(2)}$, *the* $\frac{N_{(1)} - D_{(1)}^{(k)}}{N_{(1)}}$ term is too big.
- This bias will accumulate and get larger, as we move to larger and larger $t_{(j)}$.

ESTIMATING CUMULATIVE INCIDENCE

- Letting $D_{(j)}^{(\bar{k})}$ = the number of failures of types other than k at $t_{(j)}$, an unbiased estimate of $F^{(k)}(t)$ is given by

$$\sum_{j:t_{(j)} \le t} \frac{D_{(j)}^{(k)}}{N_{(j)}} \prod_{i=1}^{j-1} \frac{N_{(i)} - D_{(i)}^{(k)} - D_{(i)}^{(\bar{k})}}{N_{(i)}} = \sum_{j:t_{(j)} \le t} \frac{D_{(j)}^{(k)}}{N_{(j)}} \prod_{i=1}^{j-1} \frac{N_{(i)} - D_{(i)}^{(k)}}{N_{(i)}} \cdot \frac{N_{(i)} - D_{(i)}^{(\bar{k})}}{N_{(i)}}$$

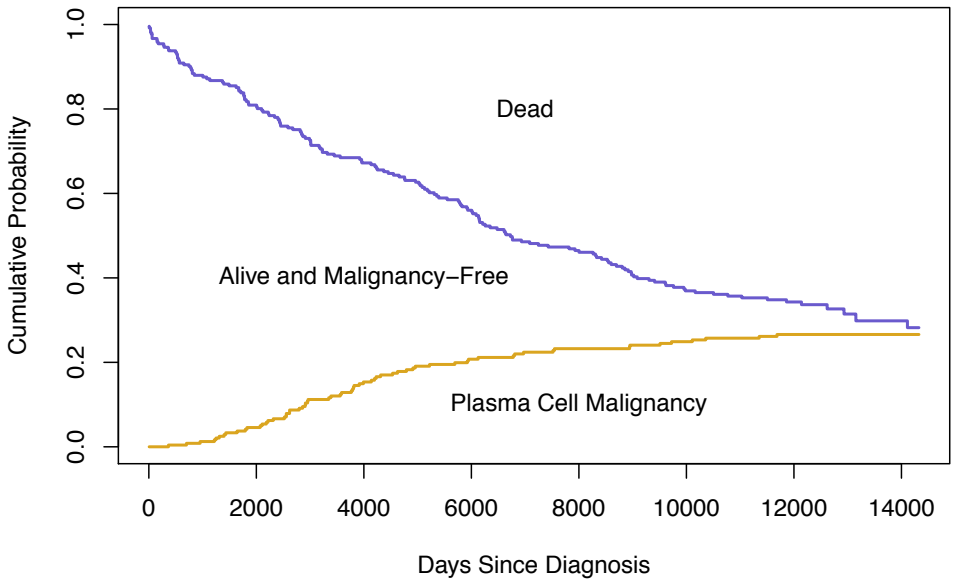
↑
no ties between failures of different types

- Compare to biased upward

$$1 - \hat{S}^{(k)}(t) = \sum_{j:t_{(j)} \le t} \frac{D_{(j)}^{(k)}}{N_{(j)}} \hat{S}^{(k)}(t_{(j-1)}) = \sum_{j:t_{(j)} \le t} = \frac{D_{(j)}^{(k)}}{N_{(j)}} \prod_{i=1}^{j-1} \frac{N_{(i)} - D_{(i)}^{(k)}}{N_{(i)}}$$

.

PREFERRED: TOGETHER



CHOICE OF OUTCOME EVENT

- May not want cause-specific event as primary outcome.
- Interpretation cloudy, particularly for survival curves, since those who die early may count as “survivors” of a competing event.

CUMULATIVE FUNCTIONS

- When there are competing risks, functions that describe the distribution of the event-specific time T_k do not make sense:
  - If the subject fails of another cause before t, T_k is not defined.
- Some other cumulative functions do make sense, depending on context:
  - The probability that a subject is alive and event-of-interest-free at t.
    - This means re-defining the event of interest to be the original event of interest or death.
  - The probability that an event of type k has (or has not) occurred by time t.
    - It has not occurred if the subject dies before t.

CUMULATIVE FUNCTIONS

Event-free Survival:

Estimating the probability a subject is alive and event-of-interest-free at time t is easy:

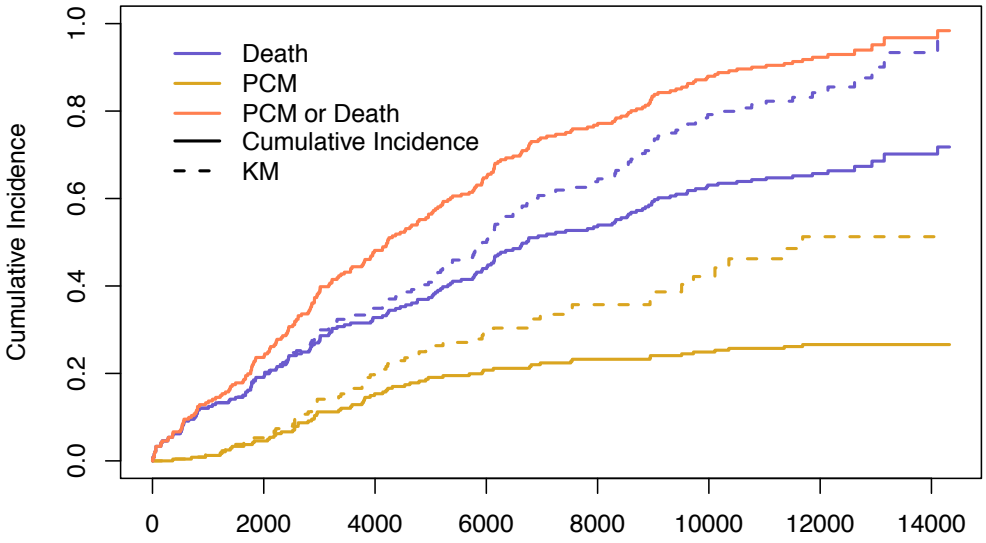
1. Redefine the event of interest to be either the original event of interest or death

$$\delta_i = \begin{cases} 1 & \text{event of interest or death from any cause} \\ 0 & \text{censored} \end{cases}$$

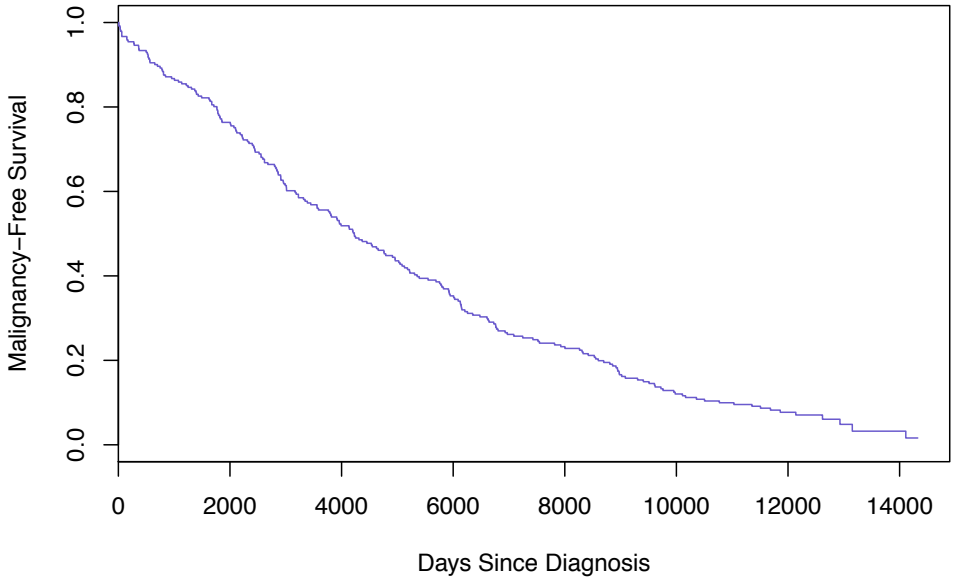
$$T_i = \text{ time to event of interest, death or censoring}$$

2. Compute the KM estimate of $S(t)$ in the usual way with (T_i, δ_i) data.

CUMULATIVE INCIDENCE



MALIGNANCY-FREE SURVIVAL



EXAMPLE

Miller K, Wang M, Gralow J, Dickler M, Cobleigh M, Perez EA, Shenkier T, Cella D, Davidson NE. Paclitaxel plus bevacizumab versus paclitaxel alone for metastatic breast cancer. New England Journal of Medicine. 2007;357(26):2666–2676.

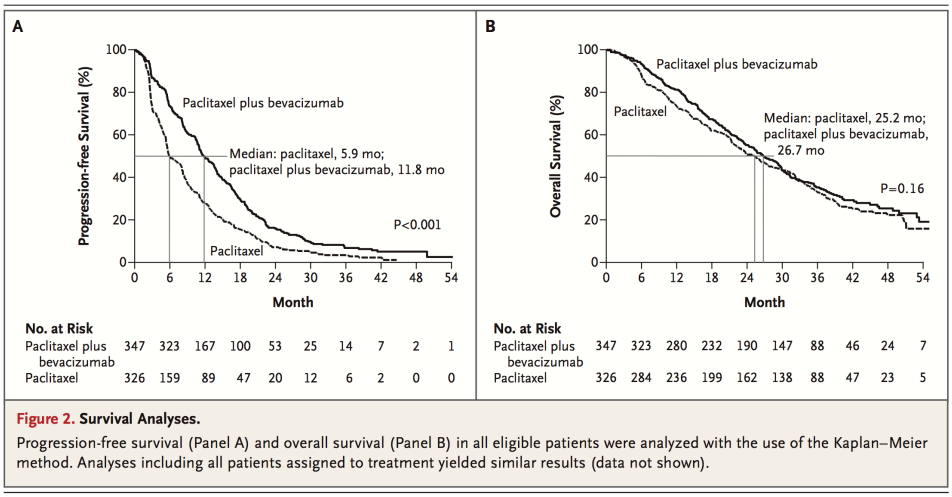


Figure 2. Survival Analyses.
Progression-free survival (Panel A) and overall survival (Panel B) in all eligible patients were analyzed with the use of the Kaplan–Meier method. Analyses including all patients assigned to treatment yielded similar results (data not shown).

SOME SUBTLETIES

- Cumulative incidence: the probability that an event of type *k* has occurred by time *t:*
  - Makes sense without requiring that a time to the k^{th} type of event be defined for all subjects
  - Depends on the portion of the population still at risk at each time, so its value will depend not only on the risk of the event of interest, but also on the risk of all the other causes of failure.
  - Is a population-specific quantity that depends on what other risks are operating in the population and how they are related to the risk of the event of interest.

OUTLINE

- Nonparametric estimation of hazard functions
- Competing risks:
  - Definition: when there is more than one cause of death/failure
  - Cumulative functions:
    - Event-free survival
    - Cumulative Incidence estimator
  - **Cause-specific and sub-distribution hazards**
  - **Fine-Gray and Cox regression models**
  - Interpretation subtleties

REGRESSION MODELS

- There are two types of regression models for competing risks outcome data that can have useful interpretations
  - Fine-Gray models, that model the association between independent variables and the **hazard** function associated with the cumulative incidence sub-distribution function
  - Cox regression model, that models the association between independent variables and the cause-specific **hazard** function

FINE-GRAY HAZARD

$$\lambda^{FG(k)}(t) = \Pr[T \in [t, t+\Delta t), c = k | T \geq t \text{ or both } T < t \text{ and } c \neq k]/\Delta t$$

Interpretations:

- the risk of failure of type k among those event free at t and those who have experienced all events other than a type k event. (Note if type k is not death, this would include subjects who had already died.)
- the hazard function associated with the sub-distribution given by the cumulative incidence function for a type k failure

FINE-GRAY MODEL

Fine-Gray Hazard:

$$\lambda^{FG(k)}(t) = \Pr[T \in [t, t+\Delta t), c = k | T \geq t \text{ or both } T < t \text{ and } c \neq k]/\Delta t$$

Model:

$$\lambda^{FG(k)}(t|\mathbf{x}) = \lambda^{FG(k)}(t|\mathbf{0})e^{\beta \mathbf{x}}$$

INTERPRETATION

Before, we noted that the interpretation of the Fine-Gray hazard is the risk of failure of type k among those event free at t **and** those who have experienced all events other than a type k event.

Suppose a type k failure is not death, and high values of x are associated with higher risk of death. Then if x has no causal association with type k failure, it will still be negatively associated with the Fine-Gray hazard of type k failure, since subjects with high values of x will be less likely to live long enough to experience a type k failure.

- Cumulative incidence of type k events will be lower for high values of x.
- Appropriate model when concern is about population burden, cost of type k events or estimating probabilities of the different possible outcomes for a patient.
- Not appropriate for causal modeling.

COX CAUSE-SPECIFIC HAZARD

$$\lambda^{(k)}(t) = \Pr[T \in [t, t+\Delta t), c = k | T \geq t]/\Delta t$$

Interpretation:

- Risk of type k event among subjects who have not experienced any of the types of events.
- Not influenced by the number of event-free subjects BUT can be influenced by who the event-free subjects are (Tsiatis, 1975).
- No real interpretation in terms of cumulative functions. The analogue would be the KM-survival curve, which is not appropriate.
- Appropriate for causal modeling with interpretation restricted to the population from which data are drawn (Prentice et al., 1978)

PROPERTIES

T = time to first "failure" of any type

$$\lambda^{(k)}(t) = \lim_{\Delta t \to 0} \Pr[T \in [t, t+\Delta t), c = k | T \geq t] / \Delta t$$

- The different events defined by c must be mutually exclusive
- The different events defined by c must be exhaustive
- The hazard function for the distribution of T is given by :

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

ESTIMATION

Cause-specific Hazard Functions:

- For estimation, we can treat failures from other causes as censored, and estimate the cause-specific hazard $\lambda^{(k)}(t)$ and cumulative cause-specific hazard $\Lambda^{(k)}(t)$ in the usual way.

Q: Why does this work?

A:

Q: How are failures from other causes conceptually different from the censoring we have talked about earlier in this course?

A:

ESTIMATION

Cumulative Incidence:

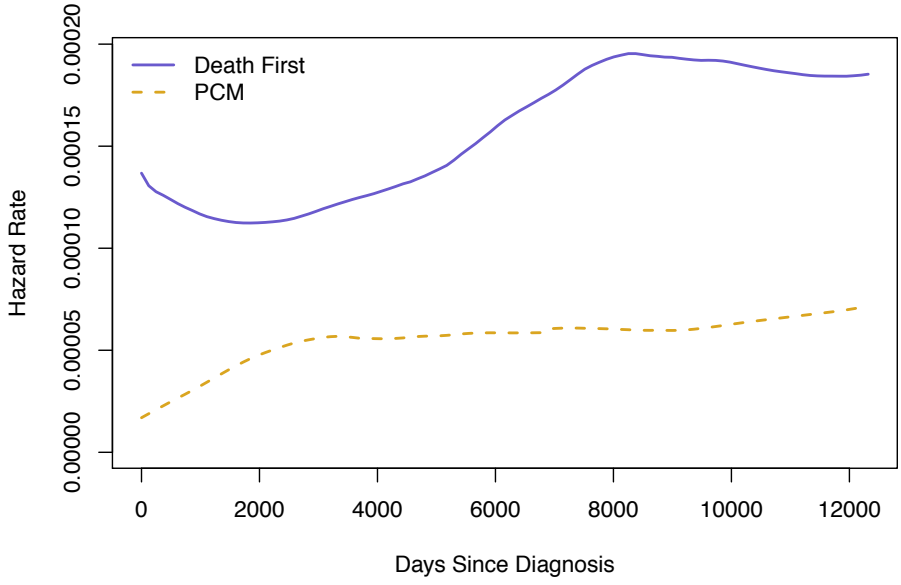
- We can treat failures from other causes as censored, and estimate the cause-specific hazard $\lambda^{(k)}(t)$ and cumulative cause-specific hazard $\Lambda^{(k)}(t)$ in the usual way.

- If we do the same thing for the cumulative incidence, the Kaplan-Meier $1 - \hat{S}^{(k)}(t)$ is a biased estimate of $F^{(k)}(t)$.

Q: Why?

A:

MGUS DATA



OUTLINE

- Nonparametric estimation of hazard functions
- Competing risks:
  - Definition: when there is more than one cause of death/failure
  - Cumulative functions:
    - Event-free survival
    - Cumulative Incidence estimator
  - Cause-specific and sub-distribution hazards
  - Fine-Gray and Cox regression models
  - **Interpretation subtleties**

IDENTIFIABILITY AND INTERPRETATION

- Tsiatis (1975) showed that we cannot identify from (T, $c = k$) data whether subjects who fail from one cause would have been more or less susceptible later to failure from another cause, had they survived.
  - Cannot tell whether those who die from heart disease would have been more or less likely to develop cancer later.
- Prentice et al (1978) argued that the cause-specific hazard function (Cox model) was the best basis for causal inference in the population as it is constituted, but cannot extend interpretation to another population where competing risks are not operating.
  - Cannot say how x might be related to cancer risk in a population where there are no deaths from MI

INTERPRETATION SUBTLETY

Cannot estimate what the hazard $\lambda^{(k)}$ (or survival function $S^{(k)}(t)$) would be if competing causes of failure were removed.

Reason: $\lambda^{(k)}(t)$ = risk of type k failure at t among those still at risk at t.

$\neq$ risk of type k failure at t among those still at risk at t if other causes were removed.

For these to be equal, population at risk at T would need to have the same risk of event k whether or not other causes of failure were removed. This is a strong and unverifiable assumption (Tsiatis, 1975).

INTERPRETATION SUBTLETY

- Cause-specific hazard is nonetheless, less dependent on the risk of other types of failure in the following sense
  - If some intervention increases the risk of another type of failure, but does not change how susceptible survivors of that failure type are to the type of interest, the intervention will not influence the cause-specific hazard function
  - If some intervention increases the risk of another type of failure, there will be fewer subjects at risk in the population at later times, and the cumulative incidence of the event of interest will be lower.

COX MODEL

Cause-specific Hazard:

$$\lambda^{(k)}(t) = \Pr[T \in [t, t+\Delta t), c = k | T \geq t]/\Delta t$$

Cox Model:

$$\lambda^{(k)}(t|\mathbf{x}) = \lambda^{(k)}(t|\mathbf{0})e^{\beta \mathbf{x}}$$

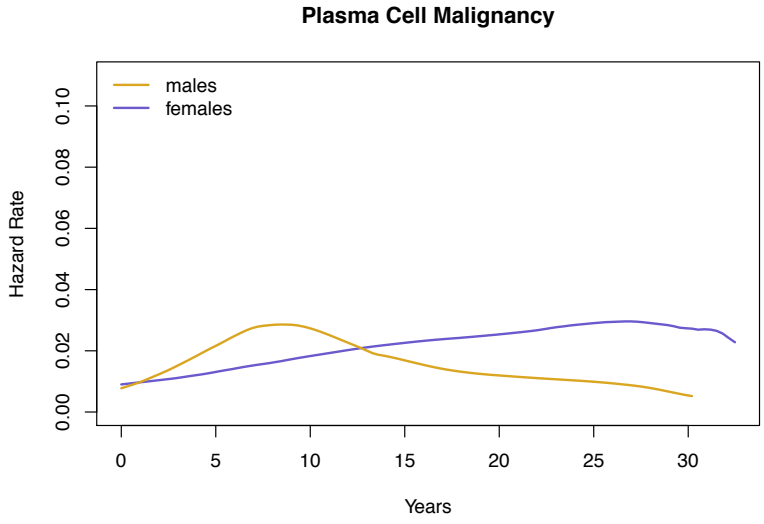
CHOOSING A COMPETING RISKS MODEL

- Cox model with cause-specific hazard good when interested in what causes the event of interest in the population as it is constituted.
- Fine-Gray good when interested in predicting patient prognosis or population disease/cost burden.
- Both are proportional hazards models, but for different hazard functions

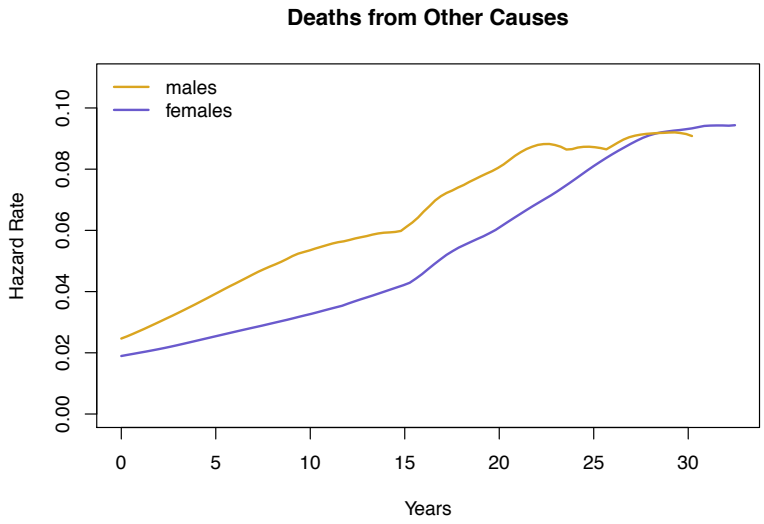
EXAMPLE

- Cox and Fine-Gray models for the association of sex with PCM and Death before PCM in the Monoclonal Gammopathy data.
- Will show
  - Cause-specific hazard functions by sex and cause
  - Cumulative incidence functions by sex and cause
  - Estimated Hazard ratios (male to female) by cause under both models

CAUSE-SPECIFIC HAZARD ESTIMATES



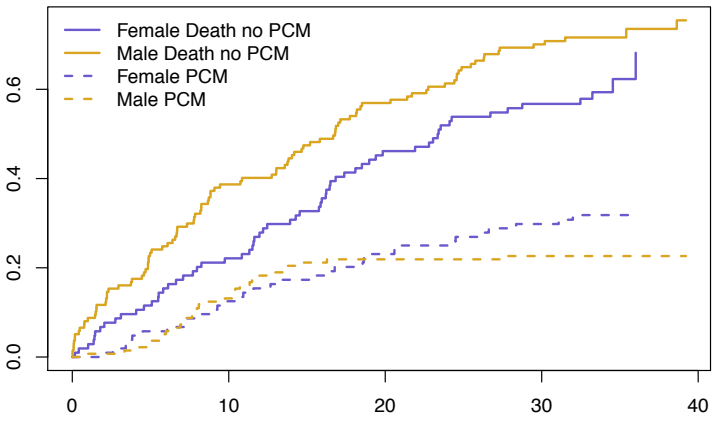
CAUSE-SPECIFIC HAZARD ESTIMATES



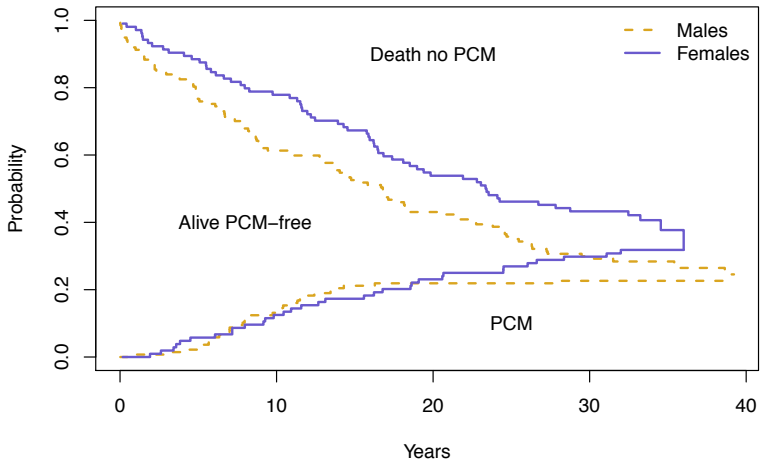
COX MODELS

Outcome Type	M/F Hazard Ratio	95% CI	P-value
Plasma Cell Malignancy	0.95	(0.58, 1.56)	0.8441
Death from Other Causes	1.55	(1.13, 2.14)	0.0064

CUMULATIVE INCIDENCE



CUMULATIVE INCIDENCE



FINE-GRAY MODELS

Outcome Type	M/F Hazard Ratio	95% CI	P-value
Plasma Cell Malignancy	0.71	(0.44, 1.16)	0.17
Death from Other Causes	1.45	(1.06, 1.97)	0.02

PCM hazard ratio farther from one here because men are more likely to die from other causes and not survive to develop PCM.

COMMON MISINTERPRETATION

- The Fine-Gray model is not the model that "accounts for competing risks"
- Both the Cox model with cause-specific hazard functions and the Fine-Gray model account for competing risks, but as we have seen their targets of inference are different.
  - Hazard among those still at risk of the event (Cox)
  - Hazard among those who have not yet experienced the event of interest, but could have experienced others, including death, already (F-G)

EXAMPLE

- Ashburner et al (2017) studied a cohort of 13,559 subjects diagnosed with atrial fibrillation (AF) at Kaiser Northern California
  - 1092 thromboembolism events (1017 ischemic strokes)
  - 4414 experienced death without thromboembolism event
  - Thromboemolism-free Death rate was 5.5/100 PY among warfarin takers and 8.1/100 PY among non-takers
  - Non-takers were older had higher stroke-risk scores
- They compared Cox and F-G regression with time-dependent current warfarin use as the exposure

EXAMPLE

Event	Model	Adjusted Hazard Ratio	95% CI
Thromboembolism	Cox	0.57	(0.50, 0.65)
	Fine-Gray	0.87	(0.77, 0.99)

- They concluded that the Fine-Gray model that "accounted for" competing risks gave a better "real-world" assessment of the benefit of warfarin.
- What are your thoughts?

Ashburner JM, Go AS, Chang Y, Fang MC, Fredman L, Applebaum KM, Singer DE. J Am Geriatr Soc. 2017 Jan 1;65(1):35–41.

SOME COMPETING RISKS REFERENCES

- Austin PC, Lee DS, Fine JP. Introduction to the Analysis of Survival Data in the Presence of Competing Risks. Circulation. 2016 Feb 9;133(6):601–609.
- Fine JP, Gray RJ. A Proportional Hazards Model for the Subdistribution of a Competing Risk. Journal of the American Statistical Association. 1999 Jun 1;94(446):496–509.
- Lau B, Cole SR, Gange SJ. Competing risk regression models for epidemiologic data. American journal of Epidemiology. 2009;170(2):244–256.
- Noordzij M, Leffondré K, Stralen V, J K, Zoccali C, Dekker FW, Jager KJ. When do we need competing risks methods for survival analysis in nephrology? Nephrol Dial Transplant. 2013 Nov 1;28(11):2670–2677.
- Prentice RL, Kalbfleisch JD, Peterson Jr AV, Flournoy N, Farewell VT, Breslow NE. The analysis of failure times in the presence of competing risks. Biometrics. 1978;541–554.
- Tsiatis A. A nonidentifiability aspect of the problem of competing risks. Proceedings of the National Academy of Sciences. 1975;72(1):20–22.

TO WATCH OUT FOR

- Interpretation in the presence of competing risks can be subtle and requires care.
  - S(t) defined in terms of the probability distribution of T_k does not make sense
  - Cannot interpret functions of the cause-specific hazard as applying in a population without competing risks present.
  - 1 – KM estimator can give upward biased estimate of cumulative incidence.
  - Cannot interpret cumulative incidence as applying in a population without competing risks present
  - Fine-Gray model is not THE way to account for competing risks.