Competing risks analysis: practical considerations and recent developments

Andrea Bellavia

TIMI Thursday meeting - 6/23/2022

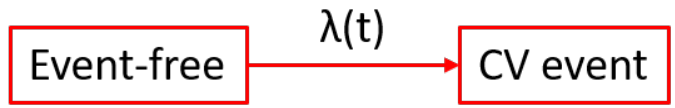
Outline

1. Competing risks overview
2. Consequences of ignoring competing risks
3. Literature example
4. Practical considerations
5. Recent developments (semi-competing risks)

1. Competing risks overview

Survival analysis framework

In survival analysis we are interested in the risk/rate of experiencing a given endpoint of interest and how these are affected by covariates



At any time point, any individual has a given probability (the hazard) of experiencing the event.

Cox model is commonly used to estimate the hazard ratios as a function of individual characteristics

$$\lambda(t|X) = \lambda_0(t)\exp(\beta_j X)$$

$$HR = \exp(\beta)$$

Individuals who leave the study without experiencing the event during follow-up are treated as censored information

From hazard to event probability

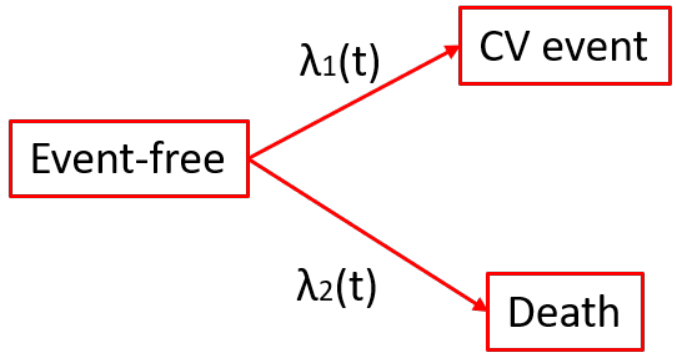
If the interest is in the actual **risk** of the event, we estimate cumulative incidence functions (CIF) with 1-KM or use Cox models results to predict risk using the following relationship.

$$1 - CIF(t) = (1 - CIF_0(t))^{\exp(\beta)}$$

The existence of this mathematical relationship is critical for interpretation: for example, even though the HR is not a RR we know that a HR>1 will always imply that RR>1

Competing risks

If the occurrence of another event (e.g. death) prevents occurrence of the event of interest (formally censoring the individual) we refer to this second event as competing.



At any time point, individuals will have specific probabilities (hazards) of experiencing either event.

In the presence of strong competing events it is usually not possible to make assumptions of **independent censoring** (i.e. individuals who experience event 2 and individuals who do not, have the same risk of event 1)

A quick note on composite endpoints

Creating a composite endpoint that incorporates competing events is a simple and correct way of accounting for competing risks.

Needless to say, in many settings this will prevent assessing the actual effects on the endpoint of interest, thus requiring additional approaches.

Competing risks framework

In a competing risks setting we get to observe, for individual i:

- $Y_i = min(T_{i1}, T_{i2}, C_i)$
- $\delta_{i1} = (0, 1)$
- $\delta_{i2} = (0, 1)$

In words, we get to observe whatever occurs first between the time to competing event 1 (T_{i1}), the time to competing event 2 (T_{i2}), or the time to censoring without having experienced neither event (C_i).

Note: everything that follows could be generalized to n competing events

Cause-specific approach

A common way to treat a competing event is to account for individual experiencing the competing event as censored individuals (i.e. practically ignoring them)

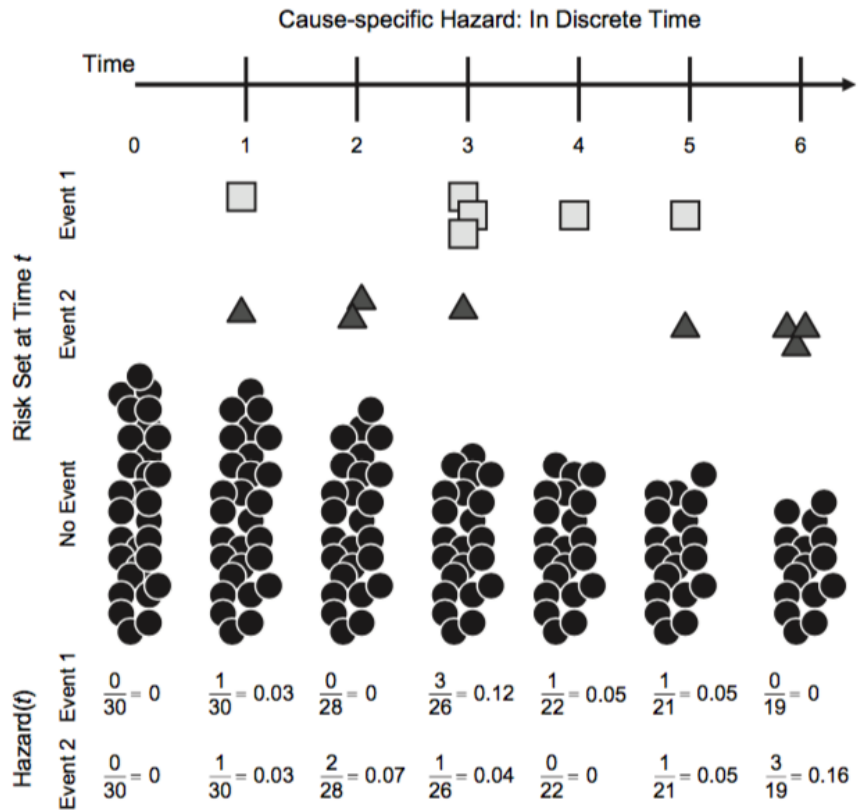


Figure 1. Cause-specific hazard schematic. The risk set starts with 30 individuals (solid circles). Over time, individuals have either event 1 (square) or event 2 (triangle). As individuals have either event, they are removed from the remaining risk sets. The calculation for the cause-specific hazard is given at the bottom of the figure.

Figure 1: Cause-specific hazard as presented in Lau et al. 2009, AJE

Regression models for cause-specific hazards

A proportional hazard model can be constructed for the cause-specific hazard:

$$\lambda_j(t|Z) = \lambda_{0j}(t)\exp(\beta_j Z)$$

where λ_{0j} is the arbitrary baseline cause-specific hazard.

Estimates are used to derive cause-specific hazard ratios, which correspond to results from the standard Cox model

Are these estimates correct?

The cause-specific HR is interpretable as the relative change in the cause-specific hazard.

No assumptions of the relation between the competing outcomes are needed for estimation, and for this reason the impact of competing events on HRs estimates is often negligible (Kalbfleisch & Prentice, 1980)

Things are different if we attempt to estimate the risk of the event.

Cause-specific estimation of the cumulative incidence function (CIF)

A naive estimation of the CIF would be to use 1-KM assuming the competing event as non-informative censoring.

- Define $F_1(t)$ the probability (cumulative incidence) of experiencing competing event 1 before time t
- Define further the cause-specific hazard function for event 1 as $\lambda_1(t)$. This represents the probability of experiencing event 1 in a small interval of time, say $t + \Delta$, conditional of being in the study and free of event 1 at time t
- However, the individual is also at risk of experiencing event 2 in the same interval ($\lambda_2(t)$)
- Therefore, $F_1(t)$ not only depends on $\lambda_1(t)$, but also on $\lambda_2(t)$

In a competing risks setting there is no longer a one-to-one correspondence between the (cause-specific) hazard $\lambda_1(t)$ and the probability of the event (cumulative incidence) $F_1(t)$

Specifically, these new relationships can be derived:

$$S(t) = e^{-\Lambda_1(t)-\Lambda_2(t)}$$

$$F_1(t) = \int_0^t S(u)\lambda_1(u)du$$

Estimators that do not account for $S_2(t)$ (like KM) will be biased. Several alternatives based on the equations above have been derived (the most common being Aalen-Johansen)

2. Consequences of ignoring competing risks (cause-specific approach)

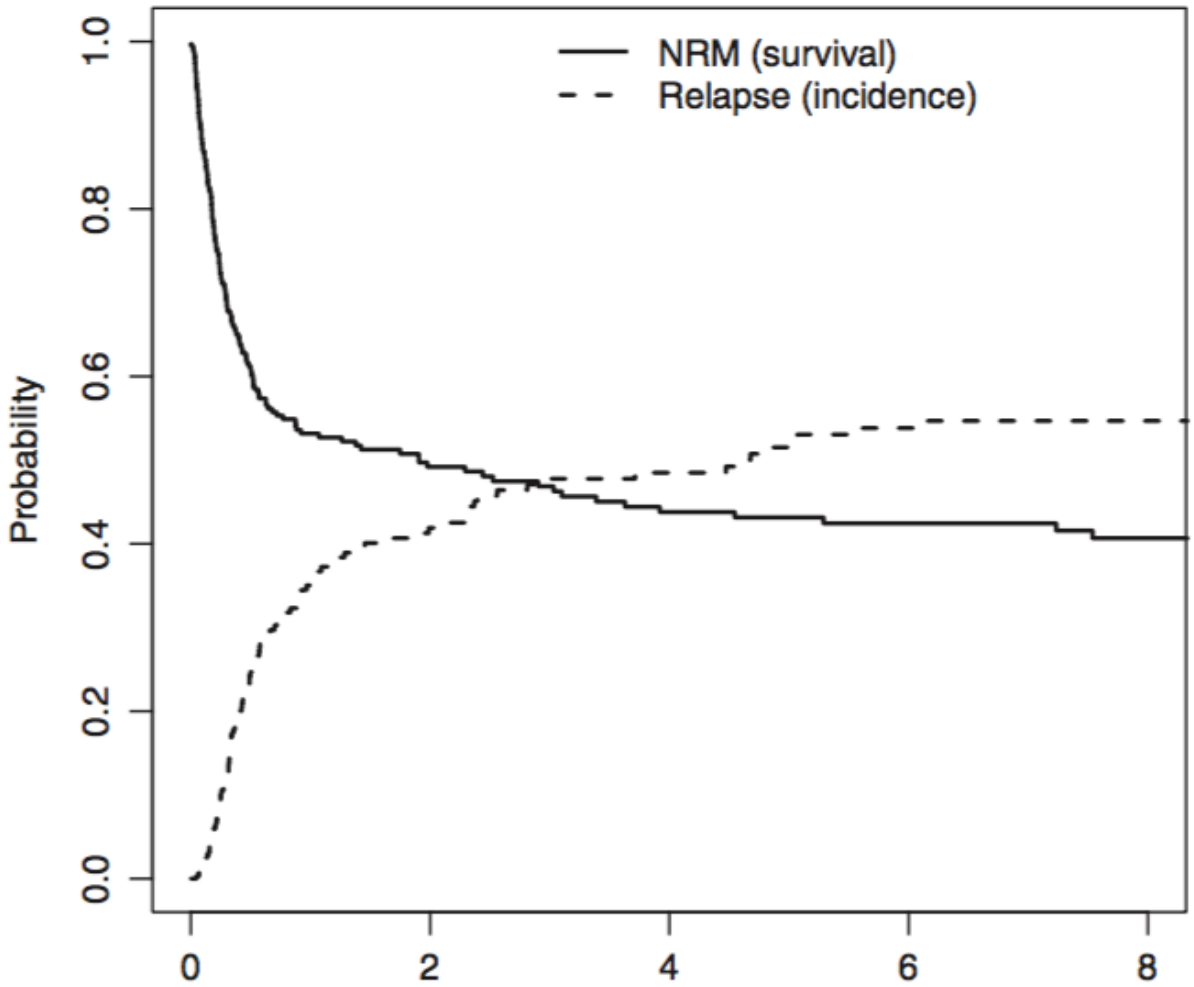
Problems with cause-specific approach

This key feature of competing risks (no longer having a one-to-one correspondence between cumulative incidence and cause-specific hazard) has two important consequences:

- The estimator of the cumulative incidence $F_1(t)$ calculated by $1 - \hat{S(t)}$, where $\hat{S(t)}$ is estimated via KM assuming event 2 as independent censoring, will be biased.
- The way $F_1(t)$ is associated with covariates will not necessarily coincide with the way in which $\lambda_1(t)$ is associated with covariates, but will also depend on the associations with $\lambda_2(t)$

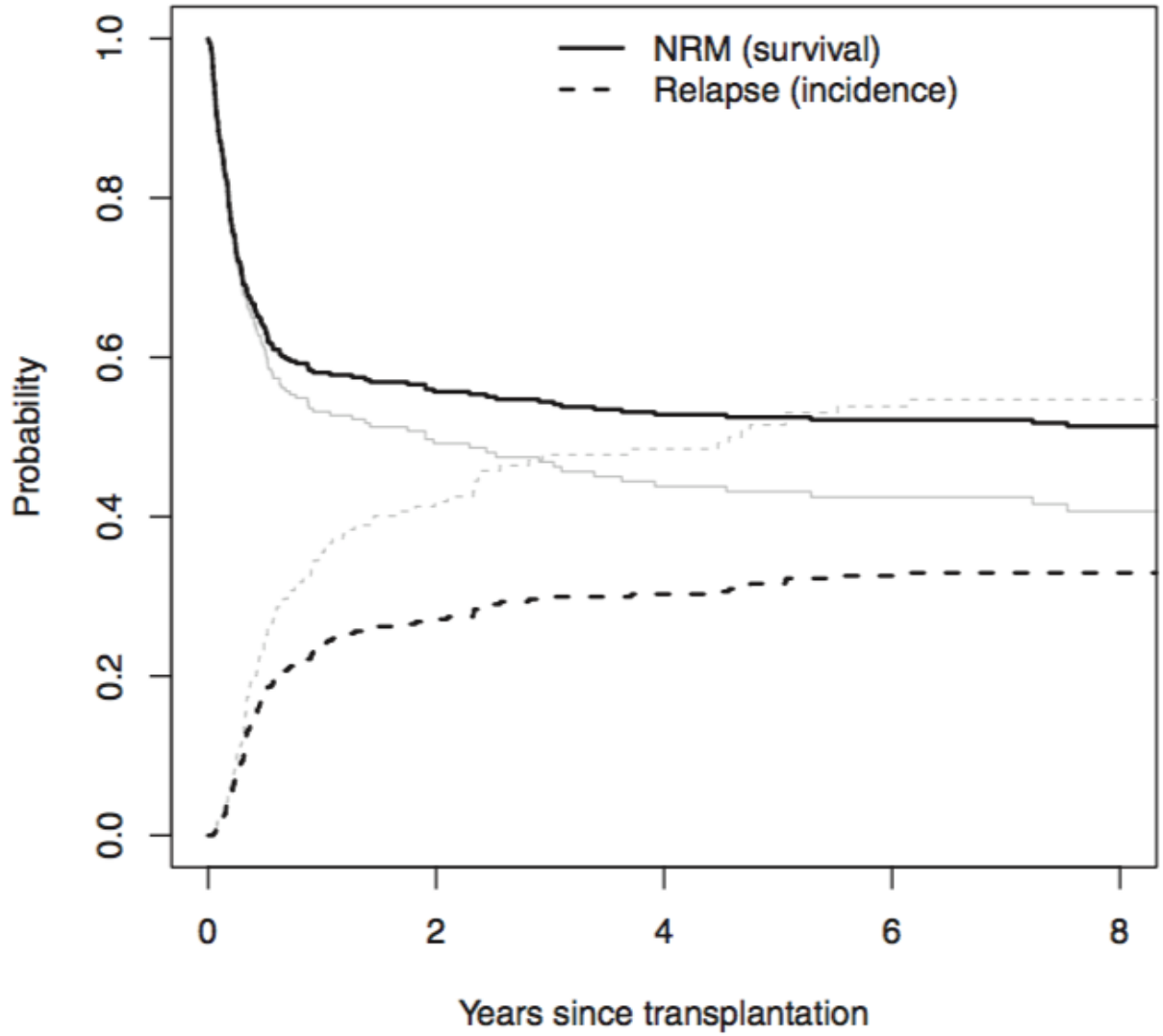
Extreme situation

The KM bias is clearly observed in this figure, presenting the CIF and survival curve for 2 competing events (relapse and non-relapse mortality in a oncology study), estimated with 1-KM and KM, respectively



The 5-year probability of relapse is 0.515, and the 5-year probability of NRM is 0.569. However, these could never sum-up to more than 1, since events are mutually exclusive.

Here are the same curves estimated with Aalen-Johansen



Adjusted risk estimation

- To predict events probability adjusting for covariates we use HRs estimated from adjusted Cox model and make predictions of the risk of the event.
- While no assumption on competing events is required to estimate and interpret cause-specific HRs, the assumption of **independent competing events** is strongly needed to underpin the inference that the cause-specific hazard and corresponding cumulative incidence functions quantify the risk
- For these reasons alternative techniques based on the idea of subdistribution hazards have been developed.

How to re-establish the hazard/risk relationship (subdistribution hazard)

The subdistribution hazard is calculated by including, in the risk set for event 1, both individuals who have not experienced event 1 and individuals who have experienced the competing event.

This may seem counter-intuitive. However:

- It is a simple way of avoiding the independence assumption.
- It can be shown that a one-to-one relationship between risk and rate is re-established

A recent paper has addressed most of the issues related to the validity of subdistribution HRs' interpretation (Putter et al, 2020, Biom J)

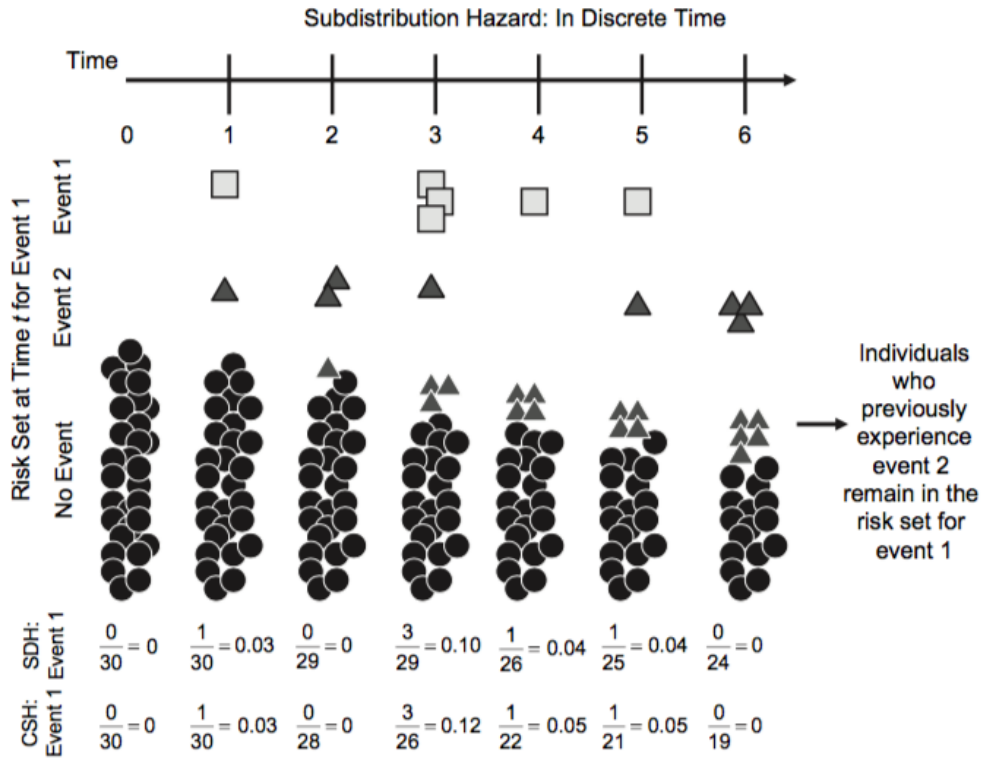


Figure 2. Subdistribution hazard schematic. The risk set starts with 30 individuals (solid circles). Over time, individuals have either event 1 (square) or event 2 (triangle). As individuals have the competing event (event 2, triangle), they are maintained in the risk set as triangles. Thus, over time, a greater proportion of the risk set becomes full of triangles that are individuals who have had the competing event prior to that time. The subdistribution hazard (SDH) for event 1 is given near the bottom of the figure along with the cause-specific hazard (CSH) for event 1 for comparison. Note that, because individuals are maintained in the risk set, the SDH tends to be lower than the CSH.

Figure 2: Subdistribution hazard as presented in Lau et al. 2009, AJE

Proportional hazards models for subdistribution hazards

Based on the subdistribution hazards we can define an alternative proportional hazards model

$$\lambda_j(t|Z) = \lambda_{0j}(t)\exp(\alpha_j Z)$$

where λ_{0j} is the arbitrary baseline subdistribution hazard.

Subdistribution hazards and related PH models were introduced in 1999 by Fine and Grey in a seminal paper. For this reason this model is often referred to as the Fine and Grey model for competing risks.

As the relationship between hazard and incidence is re-established, an estimator of the adjusted survival curve can be derived.

3. Literature example

Example from literature (Austin PC et al. Circulation. 2016)

- Enhanced Feedback for Effective Cardiac Treatment (EFFECT) Study
- Clinical data on 16,237 patients hospitalized with heart failure (HF) at 103 hospitals in Ontario, Canada
- 11 baseline covariates, which make up the EFFECT-HF mortality prediction model
- Interest in death for CV and non-CV causes
- 10,215 (63%) patients died during 5 years of follow-up. Of these, 5970 died of cardiovascular causes, and 4245 died of noncardiovascular causes.

KM estimates of CIF

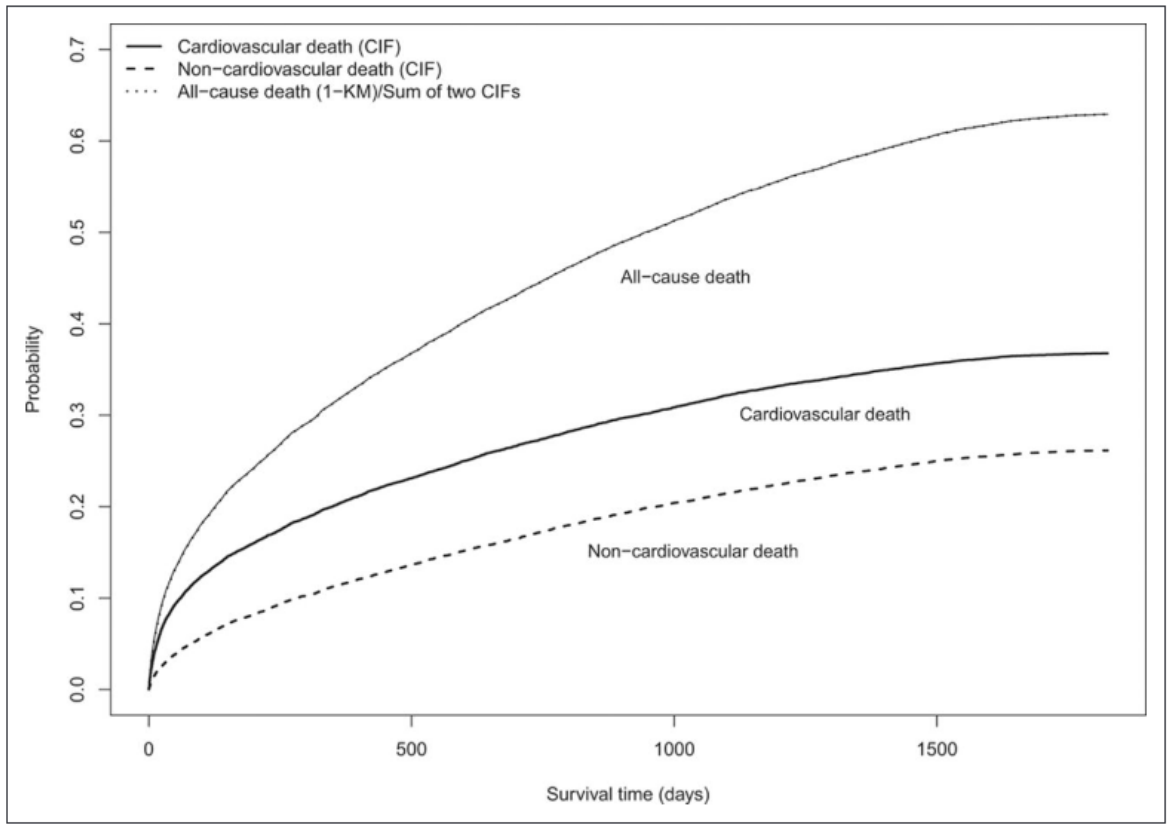


Figure 1. Cumulative incidence functions. CIF indicates cumulative incidence function; and KM, Kaplan–Meier.

HRs from cause-specific and subdistribution models

Table 2. Hazard Ratios (and 95% Confidence Intervals) From Cause-Specific and Subdistribution Hazard Models for Cardiac and Noncardiac Death (Table view)

	Subdistribution Hazard Model		Cause-Specific Hazard Model	
Variable	Cardiac Death	Noncardiac Death	Cardiac Death	Noncardiac Death
Age (per 10-y increase in age)	1.42 (1.38–1.46)	1.14 (1.10–1.17)	1.52 (1.48–1.56)	1.31 (1.27–1.35)
Respiratory rate (per 10 breaths/min increase)	1.12 (1.08–1.17)	1.07 (1.02–1.12)	1.18 (1.14–1.23)	1.15 (1.09–1.20)
Systolic blood pressure (per 10 mm Hg increase)	0.93 (0.92–0.94)	1.00 (0.99–1.01)	0.91 (0.90–0.92)	0.96 (0.95–0.97)
Urea nitrogen	1.02 (1.02–1.02)	1.01 (1.01–1.01)	1.03 (1.02–1.03)	1.02 (1.02–1.02)
Cancer	0.82 (0.75–0.89)	1.85 (1.71–2.01)	0.96 (0.89–1.04)	1.85 (1.71–2.00)
Cirrhosis	1.12 (0.82–1.54)	1.49 (1.10–2.02)	1.20 (0.89–1.62)	1.66 (1.24–2.22)
Cerebrovascular disease	1.26 (1.18–1.35)	1.01 (0.93–1.09)	1.30 (1.22–1.38)	1.13 (1.04–1.22)
Dementia	1.22 (1.11–1.33)	1.35 (1.22–1.50)	1.48 (1.37–1.60)	1.74 (1.59–1.92)
COPD	1.03 (0.97–1.1)	1.42 (1.33–1.52)	1.14 (1.08–1.21)	1.53 (1.44–1.64)
Low hemoglobin	0.87 (0.80–0.94)	1.49 (1.37–1.62)	0.95 (0.88–1.02)	1.50 (1.39–1.63)
Low sodium concentration	1.14 (1.07–1.22)	1.13 (1.05–1.22)	1.18 (1.11–1.25)	1.19 (1.11–1.28)

CIF estimate with CR approach

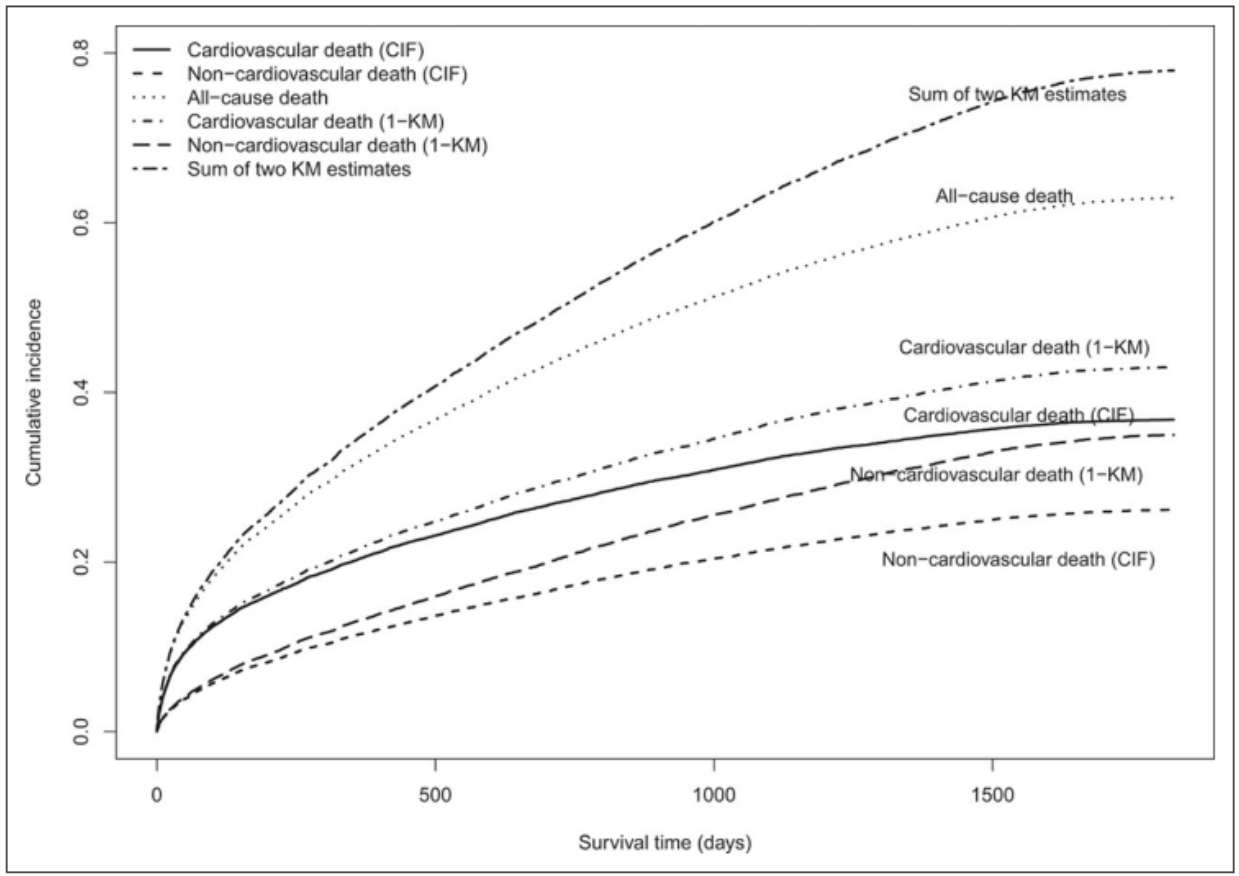


Figure 2. Cumulative incidence functions and Kaplan–Meier estimates. CIF indicates cumulative incidence function; and KM, Kaplan–Meier.

4. Practical considerations

Recap and discussion

- Competing risks break the relationship between risk and hazard.
- Both cause-specific and subdistribution models provide a valid estimate of the association between a covariate and the hazard ratio
- CIFs can be estimated directly from subdistribution hazards or with ad-hoc alternatives of the KM approach

- In practice, **HRs estimated with the 2 approaches will seldom diverge** (especially with low propotion of events)
- On the other hand, **CIFs estimated with 1-KM or predicted from Cox models be biased (over-estimated) in almost every situation**
- (Risks differences and ratios will however be consistent)

When should we bother about competing risks

Several researchers have argued that the choice of how CRs should be addressed should be based on the research question of interest. In particular:

- Disease etiology / epi studies / RCTs-> cause-specific hazard
- Prognostic studies / risk prediction -> subdistribution hazard

Ignoring competing risks in studies where the goal is to predict the individual probabilities of experiencing an event might severely overestimate the risks.

These 2 publications are great references for further reading and to support this argument

Competing Risk Regression Models for Epidemiologic Data FREE

Bryan Lau, Stephen R. Cole, Stephen J. Gange

American Journal of Epidemiology, Volume 170, Issue 2, 15 July 2009, Pages 244–256, https://doi.org/10.1093/aje/kwp107

Published: 03 June 2009 **Article history** ▾

Introduction to the Analysis of Survival Data in the Presence of Competing Risks

Peter C. Austin ✉, Douglas S. Lee and Jason P. Fine

Originally published 9 Feb 2016 | https://doi.org/10.1161/CIRCULATIONAHA.115.017719 | Circulation. 2016;133:601–609

In terms of study design and SAP preparation, it is important to incorporate competing risks analysis for prognostic studies. Other studies could simply include that as a sensitivity analysis (which will likely be requested by Reviewers anyway. . .)

This table from the Circulation paper nicely summarizes the ideal guidelines:

Table 3. Recommendations for Analyzing Competing Risk Survival Data

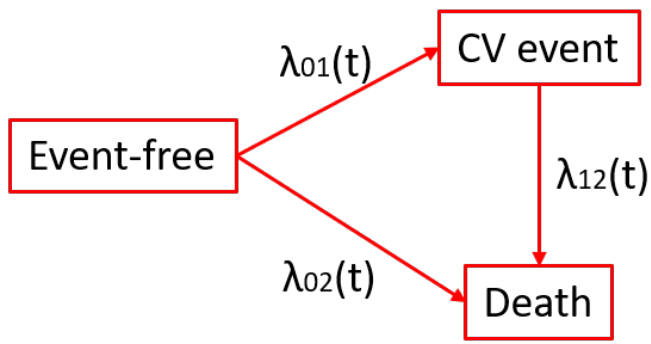
• Cumulative incidence functions (CIFs) should be used to estimate the incidence of each of the different types of competing risks. Do not use the Kaplan-Meier estimate of the survival function for this purpose.
• Researchers need to decide whether the research objective is on addressing etiologic questions or on estimating incidence or predicting prognosis.
• Use the Fine-Gray subdistribution hazard model when the focus is on estimating incidence or predicting prognosis in the presence of competing risks.
• Use the cause-specific hazard model when the focus is on addressing etiologic questions.
• In some settings, both types of regression models should be estimated for each of the competing risks to permit a full understanding of the effect of covariates on the incidence and the rate of occurrence of each outcome.

5. Recent developments

Illness-death models

Illness-death models allow to evaluate **semi-competing risks**, where the competing events are not mutually exclusive and one of the 2 states is a final absorbing state (usually death).

It is possible to estimate how covariates affect the hazard of transitioning between the different states, including the hazard of death after an event.



Methods have been developed and are available on software. Applications are still sporadic.