Setting

A **competing risk** is an event whose occurrence precludes observation of the event of interest. More generally, competing risks are present when a subject may experience one of K mutually exclusive types of event. Define $T = \min(T_1, \cdots, T_K)$, where T_k is the time to the kth competing risk. Let $D \in \{1, \cdots, K\}$ denote the type of event. Suppose the time to event is subject to random right censoring by C, and let $U = \min(T, C)$. Define the *status* as:

$$\delta = \begin{cases} 0, & C < T, \\ k, & T \leq C \text{ and } T = T_k. \end{cases}$$

Censoring is assumed *non-informative*, meaning the joint distribution of the time to event and the event type (T, D) is independent of the time to censoring C.

Cause-specific Hazard

Definition 2.0.1. The **cause-specific hazard** [4] for competing risk k is:

$$\lambda_k^{\text{cs}}(t) \equiv \lim_{\Delta \to 0} \frac{1}{\Delta} \mathbb{P}\big(t \leq T < t + \Delta, D = k | T \geq t\big). \tag{2.0.1}$$

Note that the risk set includes only those patients who have not experienced any event by time t. The corresponding proportional hazards model is:

$$\lambda_k^{\text{cs}}(t|\boldsymbol{z}) = \lambda_{k0}^{\text{cs}}(t) e^{\boldsymbol{z}'\boldsymbol{\beta}_k},$$

where $\lambda_{k0}^{\text{cs}}(t)$ is an unspecified baseline hazard for specific-cause k, $\boldsymbol{z}$ is a set of covariates, and $\boldsymbol{\beta}_k$ is a vector of cause-specific log hazard ratios. ■

Discussion 2.0.1. The cause-specific hazard $\lambda_k(t)$ is estimated by regarding the occurrence of competing events (e.g. death) as independent censoring. It is interpretable as the risk of experiencing a type k event in the next short time interval dt given the patient has not experienced any type of event by time t. The independent censoring assumption is can be problematic because it assumes the rate of event k is the same among those whose event times are in fact unknown and those who experience any one of the competing events k'. Moreover, even if the events are independent, the candidate survival function:

$$S_k^{\text{cs}}(t) = \exp\left\{ - \int_0^t \lambda_k^{\text{cs}}(u) du \right\},$$

lacks a practical interpretation; it represents the probability of being event-free by time t in a setting where the competing events cannot occur. Thus, when death is among the

competing events, $S_k(t)$ represents the probability of being event-free in the hypothetical scenario that death cannot occur before an event of type k [1]. ♠

Discussion 2.0.2. The cause-specific hazard ratio (HR_{cs}) assess whether a covariate is associated with the rate of a specific event. A $\text{HR}_{\text{cs}} > 1$ indicates the covariate is associated with an increased rate of the specific event, whereas a $\text{HR}_{\text{cs}} < 1$ indicates the covariate is associated with a decreased rate. ♠

Cumulative Incidence Function

Discussion 3.0.1. In the absence of competing events, the survival function $S(t)$ is typically estimated via the Kaplan-Meier (KM) product-limit estimator $\hat{S}(t)$, and $1 - \hat{S}(t)$ estimate provides a consistent estimate of the cumulative incidence function $F(t) = 1 - S(t)$. Treating subjects who experience competing events as censored and directly applying this approach results in underestimation of the survival function and overestimation of the cumulative incidence function. ♠

Definition 3.0.1. In the presence of competing risks, the **cumulative incidence function** (CIF) [3] is defined as:

$$\begin{aligned} F_k(t) &= \mathbb{P}(T \leq t, D = k) \\ &= \int_0^t S(u)\lambda_k(u)du = \int_0^t S(u)d\Lambda_k(u). \end{aligned}$$

where $S(t) = \mathbb{P}(T \geq t)$ is the overall survival function, and is expressible as:

$$S(t) = \exp\left\{-\sum_{k=1}^{K} \Lambda_k(u)\right\}.$$

■

Discussion 3.0.2. CIFs are additive. Thus, if $F_1(t)$ is the CIF for type 1 events and $F_2(t)$ is the CIF for type 2 events, $F_1(t) + F_2(t)$ is the CIF for the composite outcome of type 1 or 2 events. Graphically, the individual CIFs may be stacked such that the overall height represents the cumulative incidence of the composite event. ♠

Discussion 3.0.3. A covariate can have different associations with the cause-specific hazard and the CIF. For example, suppose a treatment reduces the mortality rate, but has no effect on the rate of event k. Then although the treatment has no association with the cause-specific hazard, the cumulative incidence of the event may increase since patients on treatment will survive longer, spending more time at risk for the event [5]. Conversely, a treatment may have an effect on the cause-specific hazard without having

an effect on the cumulative incidence [2]. Defining a composite endpoint between the event of interest and mortality may be more appropriate since this approach adjusts for mortality rate differences between the two arms. ♠

3.1 Estimation

To estimate the CIF, define the sample counting process for type k events:

$$N_k(t) = \sum_{i=1}^{n} \mathbb{I}(T_i \leq t, \delta_i = k),$$

and the sample at-risk process:

$$Y_k(t) = \sum_{i=1}^{n} \mathbb{I}(T_i \geq t).$$

The **Nelson-Aalen** estimator of the cumulative hazard for type k events is:

$$\hat{\Lambda}_k(t) = \int_0^t \frac{dN_k(u)}{Y(u)}$$

Let $\hat{S}(t)$ denote the standard KM estimate for the survival function of the random variable $T = \min(T_1, \cdots, T_K)$, then the CIF is consistently estimated by:

$$\hat{F}_k(t) = \int_0^t \hat{S}(u-) d\hat{\Lambda}_k(u).$$

3.2 Subdistribution Hazard

Definition 3.2.1. The **subdistribution hazard** of Fine and Gray [2] is defined as:

$$\lambda_k^{\text{sd}}(t) = \lim_{\Delta \to 0} \frac{1}{\Delta} \mathbb{P}\big\{t < T \leq t + \Delta, D = k | (T \geq t) \cup (T < t \cap D \neq k)\big\}.$$

Note here that the risk set consists of those who have not yet experienced any event $(T \geq t)$ and those who have experienced a competing event $(T < t \cap D \neq k)$. The corresponding proportional hazards model is:

$$\lambda_k^{\text{sd}}(t) = \lambda_{k0}^{\text{sd}}(t) e^{\boldsymbol{z}' \boldsymbol{\beta}_k},$$

where $\lambda_{k0}^{\text{sd}}(t)$ is an unspecified baseline subdistribution hazard, $\boldsymbol{z}$ is a set of covariates, and $\boldsymbol{\beta}_k$ is a vector of subdistribution hazard ratios. ■

Proposition 3.2.1. The subdistribution hazard is connected with the CIF via:

$$\lambda_k^{\text{sd}}(t) = -\frac{d \ln\big\{1 - F_k(t)\big\}}{dt}.$$

♦

Discussion 3.2.1. The subdistribution hazard ratio (HR_{sd}) connects covariates with the cumulative incidence function. A $HR_{sd} > 1$ indicates the covariate is associated with increased incidence of the event, whereas a $HR_{sd} < 1$ indicates the covariate is associated with decreased incidence. However, the HR_{sd} itself lacks a clinical interpretation. ♠

References

[1] Austin, PC and Lee, DS and Fine, JP. "Introduction to the Analysis of Survival Data in the Presence of Competing Risks". In: *Circulation* 133.6 (2016), pp. 601–609.

[2] Fine, JP and Gray, RJ. "A Proportional Hazards Model for the Subdistribution of a Competing Risk". In: *Journal of the American Statistical Association* 94.446 (1999), pp. 496–509.

[3] Lin DY. "Non-parametric inference for cumulative incidence functions in competing risks studies". In: *Statistics in Medicine* 16.3 (1997), pp. 901–910.

[4] Prentice, RL and Kalbfleisch, JD and Peterson, AV and others. "The Analysis of Failure Times in the Presence of Competing Risks". In: *Biometrics* 34.4 (1978), pp. 541–554.

[5] Wolbers, M and Koller, MT and Stel, VS and others. "Competing risks analyses: objectives and approaches". In: *European Heart Journal* 35 (2014), pp. 2936–2941.