

Competing Risks

“Competing risks” refers to the study of mortality patterns in a population of individuals, all subject to the same $k \geq 2$ competing risks or **causes of death**. Specifically, the objective is to isolate the effect of a given risk, or a subset of risks, acting on a population. The use of competing risks dates back to 1760 and evolved out of a controversy over smallpox inoculation.

According to Karn [22] and Todhunter [30], smallpox inoculation in the 1700s was administered by applying leeches to the body, a practice that could lead to acute illness and death. Physicians argued whether the benefits of inoculation outweighed the initial risk of death. Daniel **Bernoulli** [9], in a 1760 memoir entitled “Essai d’une nouvelle analyse de la mortalité causée par le petite vérole; et des avantages de l’inoculation pour le prévenir”, tried to estimate the expected increase in lifespan (*see* **Life Expectancy**) if smallpox were eliminated. This calculation could then be used to weigh the pros and cons of smallpox inoculation.

Similarly, in the modern treatment of competing risks we are interested in isolating the effect of individual risks. For example, suppose we wish to assess a new treatment for heart disease. In a long-term study of this treatment on a sample of individuals, some will die of causes other than heart disease. The appropriate analysis of this problem must account for the competing effects of death from other causes.

Various methods have been proposed to study the problem of competing risks. For example, Makeham [24] formulated the law of composition of decremental forces and applied it to competing risks theory. A multiple decrement model is a time-continuous **Markov model** with one transient state and k absorbing states. An excellent account of the use of multiple decrement theory to explain competing risks may be found in Chiang [12].

Another approach to modeling competing risks is through the use of latent failure times. This method was first advocated by Sampford [28] who proposed an “accidental death model”. In this approach each individual has latent failure times T_1 and T_2 , where T_1 corresponds to time of natural death and T_2 to time to accidental death. Sampford assumed that T_1 and T_2 are independent and **normally distributed** and death occurred at time X equal to the minimum

of T_1 and T_2 . Berkson & Elveback [8] considered a similar model to study the effect of smoking on lung cancer assuming that the latent failure times were independent **exponentially distributed random variables**. Moeschberger & David [25] generalized these ideas to k causes of death with general **survival distributions**. Excellent reviews of the theory of competing risks are given by Gail [18, 19], David & Moeschberger [15], and Birnbaum [10].

In this article, latent failure times are used to describe competing risks models. We assume that all individuals in a population are subject to k competing causes of death, D_1, \dots, D_k . For each possible cause of death, D_i , there corresponds a latent failure time, T_i , a positive random variable representing the age at death in the hypothetical situation in which D_i is the only possible cause of death. The joint distribution of the latent failure times is given by the **multivariate survival distribution**

$$H^C(t_1, \dots, t_k) = P(T_1 > t_1, \dots, T_k > t_k), \quad (1)$$

defined for all nonnegative values t_1, \dots, t_k . We use a superscript C to highlight that this is the joint distribution of the complete set of risks acting on the population. The latent failure times are mostly unobservable and serve only as a theoretical construct. In contrast, the observable random variables for each member of a population of individuals are the actual times to death, denoted by the positive random variable X , and the cause of death, Δ , which may take one of the integer values $1, \dots, k$. The observed time of death, X , is taken to be the minimum of T_1, \dots, T_k , and Δ indexes this cause of death, i.e. $\Delta = i$ if $X = T_i$. For simplicity we assume the joint distribution is absolutely continuous so that Δ is uniquely defined.

The study of competing risks considers the interrelationship of three types of probabilities of death from specific causes. These are:

1. *The crude probability*: the probability of death from a specific cause in the presence of all other risks acting on the population. This is also referred to as **absolute risk**. An example of a crude probability is the answer to the question: What is the chance that a woman will die of breast cancer between ages 40 and 60?
2. *The net probability*: the probability of death if a specific risk is the only risk acting on a population, or conversely, the probability of death if a

2 Competing Risks

specific cause is eliminated from the population. For example, what is the chance of surviving to age 60 if cancer were the only cause of death?

3. *The partial crude probability*: the probability of death from a specific cause when some risks are eliminated from the population. For example, what is the chance that a woman would die from breast cancer between ages 40 and 60 if smallpox were eliminated?

In the next section we define notation and give some fundamental relationships between the three different types of probabilities. Then we consider the issue of **identifiability** of these probabilities and discuss some philosophical issues regarding the study of competing risks in light of nonidentifiability. Finally, we address statistical issues of **estimation** and **hypotheses testing** based on a sample of observable data.

Notation and Relationships

Crude Probability

Crude probability is a way of describing the probability distribution for a specific cause of death in the presence of all causes. Crude probability refers to quantities derived from the probability distribution of the observable random variables, X and Δ , where X is time to death, and $\Delta = 1, \dots, k$ is cause of death. Two approaches have been used to describe the distribution of X and Δ . The first is through subdistribution functions:

$$F_i^C(x) = \Pr(X \leq x, \Delta = i), \quad i = 1, \dots, k.$$

The function $F_i^C(x)$ denotes the proportion of all individuals who are observed to die from cause D_i at or before time x in the presence of all causes of death. We use the superscript C to denote all causes of death, i.e. $C = \{1, \dots, k\}$. For example, if D_1 represents death from breast cancer, then the chance that a woman dies from breast cancer between ages 40 and 60 would be equal to $[F_1^C(60) - F_1^C(40)]$. Note that $F_i^C(\infty)$ is the proportion of individuals who will be observed to die from cause D_i , and $\sum_{i=1}^k F_i^C(x) = F^C(x)$ defines the distribution function for death from any cause, i.e. $F^C(x) = \Pr(X \leq x)$. We denote the overall survival distribution as $S^C(x) = 1 - F^C(x)$.

Another way to define the distribution of X and Δ is through the use of k cause specific **hazard rate** functions given by

$$\lambda_i^C(x) = \lim_{h \rightarrow 0} \left[\frac{\Pr(x \leq X < x + h, \Delta = i | X \geq x)}{h} \right],$$

$$i = 1, \dots, k.$$

The i th cause-specific hazard is the rate of death at time x from cause i among individuals who are still alive at time x . Calculus yields the following relationships:

$$\lambda_i^C(x) = \frac{dF_i^C(x)}{dx} / S^C(x),$$

$$\lambda^C(x) = \sum_{i=1}^k \lambda_i^C(x) = \frac{dF^C(x)}{dx} / S^C(x), \quad (2)$$

$$S^C(x) = \exp[-\Lambda^C(x)]; \quad \Lambda^C(x) = \int_0^x \lambda^C(u) du,$$

$$F_i^C(x) = \int_0^x \exp[-\Lambda^C(u)] \lambda_i^C(u) du.$$

Note that $\Lambda^C(x)$ is defined as the **cumulative hazard** function of death from any cause and is the sum of the individual cause-specific integrated hazards. The relationship given in (2) illustrates that there is a one-to-one relationship between subdistribution functions and cause-specific hazard functions.

The crude probability distributions may be derived from the joint distribution of the latent failure times as follows. Because $X = \min(T_1, \dots, T_k)$, it follows that $S^C(x) = H^C(x, \dots, x)$; hence, it is straightforward to show that

$$\frac{dF_i^C(x)}{dx} = - \frac{\partial H^C(t_1, \dots, t_k)}{\partial t_i} \Big|_{t_1 = \dots = t_k = x}.$$

Using (2), the cause-specific hazard function is given by

$$\lambda_i^C(x) = \frac{- \frac{\partial H^C(t_1, \dots, t_k)}{\partial t_i} \Big|_{t_1 = \dots = t_k = x}}{H^C(x, \dots, x)}. \quad (3)$$

This relationship was derived by Gail [18] and Tsiatis [31].

Cause-specific hazard functions and cause-specific subdistribution functions may also be defined for a subset of risks. We use italicized capital letters to

index a subset of the risks $1, \dots, k$; for example, J may be used to denote such a subset of risks. The complement of J is equal to $C - J$ and is denoted by \bar{J} . The subdistribution function for failing from any of the causes in J is given by

$$F_J^C(x) = \Pr(X \leq x, \Delta \in J) = \sum_{i \in J} F_i^C(x),$$

and the cause-specific hazard of failing from any of the causes in J is

$$\begin{aligned} \lambda_J^C(x) &= \lim_{h \rightarrow 0} \left[\frac{\Pr(x \leq X < x + h, \Delta \in J | X \geq x)}{h} \right] \\ &= \sum_{i \in J} \lambda_i^C(x). \end{aligned}$$

The Net Probability

The net probability is the probability distribution of time to death if only one cause of death acted on a population. If we are interested in the net probability distribution from cause D_i , then this would be the **marginal probability** distribution of the latent failure time, T_i , given by

$$\begin{aligned} S_i^j(x) &= \Pr(T_i > x) = H^C(t_1, \dots, t_k) |_{t_i = x,} \\ & \quad t_j = 0, j \neq i. \end{aligned}$$

We use superscript i to highlight the fact that we consider only the case where D_i is acting on a population. For example, if D_1 denotes death from cancer, then the chance of surviving to age 60 if cancer were the only cause of death would be given by $S_1^1(60)$.

The net distribution may be defined through the net or marginal hazard function for T_i , that is,

$$\lambda_i^j(x) = \lim_{h \rightarrow 0} \left[\frac{\Pr(x \leq T_i < x + h | T_i \geq x)}{h} \right].$$

The net hazard function and net survival distribution are related to each other as follows:

$$\begin{aligned} \lambda_i^j(x) &= -\frac{dS_i^j(x)}{dx} / S_i^j(x), \\ S_i^j(x) &= \exp[-\Lambda_i^j(x)], \end{aligned} \tag{4}$$

where $\Lambda_i^j(x) = \int_0^x \lambda_i^j(u) du$.

One of the key results in competing risks theory is for the case where the latent failure times are assumed

to be statistically independent, i.e.

$$H^C(t_1, \dots, t_k) = \prod_{i=1}^k S_i^i(t_i).$$

From (1) it is a simple exercise to show that the i th cause-specific hazard function, $\lambda_i^C(x)$, is equal to the i th net-specific hazard function, $\lambda_i^i(x)$. This important fact allows one to use the crude probability distribution of the observables to obtain net probabilities. Specifically, formulas (1) and (2) may be used to show that the net survival distribution is related to the crude subdistribution functions by

$$H_i^i(x) = \exp \left[- \int_0^x \frac{dF_i^C(u)}{S^C(u)} \right]. \tag{5}$$

Because $F_i^C(u)$ and $S^C(u)$ may be estimated from a sample of observable data, (5) suggests obvious methods for estimating net survival probabilities when the latent failure times are assumed independent, which are described in detail later. Although the crude cause-specific hazard is equal to the net-specific hazard when the latent failure times are independent, the converse is not true. Examples where non-independent latent failure times have cause-specific hazards equal to the net-specific hazards, although mathematically possible, are generally artificial constructs and not important from an applied perspective.

For many applications it may not be reasonable to assume that the latent failure times are independent. In such cases the relationship between net and crude probabilities becomes more complicated. Without additional assumptions, there is a problem of nonidentifiability discussed in greater detail later.

Partial Crude Probability

We now show how to characterize the distribution of probability of death from a subset of causes acting on a population in the hypothetical situation where all other causes of death are eliminated. Similar to crude probabilities, partial crude probabilities may be expressed through partial crude subdistribution functions or partial crude cause-specific hazard functions. Define X^J and Δ^J respectively as the time of death and cause of death in the hypothetical case where individuals are only subject to the causes of death in J , i.e. the causes \bar{J} are eliminated. In terms of latent failure times, $X^J = \min(T_i, i \in J)$ and $\Delta^J = i$, if

4 Competing Risks

$X^J = T_i$, $i \in J$. The partial crude subdistribution function is given by

$$F_i^J(x) = \Pr(X^J \leq x, \Delta^J = i), \quad i \in J,$$

and the partial crude cause-specific hazard is given by

$$\begin{aligned} \lambda_i^J(x) &= \lim_{h \rightarrow 0} \left(\frac{\Pr(x \leq X^J < x + h, \Delta^J = i | X^J \geq x)}{h} \right), \\ & \quad i \in J. \end{aligned}$$

These definitions may be extended in a natural way to subsets K of J , i.e.

$$F_K^J(x) = \sum_{i \in K} F_i^J(x)$$

and

$$\lambda_K^J(x) = \sum_{i \in K} \lambda_i^J(x).$$

If $J = C$, then partial crude probabilities are the same as crude probabilities, and if $J = i$, so that there is only one cause of death, then partial crude probability is the same as net probability.

Using the same logic as for crude probabilities, we can derive the partial crude cause-specific hazard function from the joint distribution of the latent failure times in a manner similar to that for (3). The partial crude cause-specific hazard is given by

$$\lambda_i^J(x) = \frac{\left. \frac{\partial H^C(t_1, \dots, t_k)}{\partial t_i} \right|_{t_j=x, j \in J; t_j=0, j \in \bar{J}}}{\left. H^C(t_1, \dots, t_k) \right|_{t_j=x, j \in J; t_j=0, j \in \bar{J}}}, \quad (6)$$

and the partial crude subdistribution function may be expressed as

$$F_i^J(x) = \int_0^x \exp[-\Lambda_J^J(u)] \lambda_i^J(u) \, du, \quad i \in J, \quad (7)$$

where $\Lambda_J^J(u) = \int_0^u \lambda_J^J(v) \, dv$.

Of particular interest is the case when the latent failure times in the set J are independent of the latent failure times in \bar{J} . Comparing (6) with (3) we see that the i th partial crude cause-specific hazard function, $\lambda_i^J(x)$, is equal to the overall crude cause-specific hazard function, $\lambda_i^C(x)$, $i \in J$. This allows us to express the unobservable partial crude probabilities in terms of the observable crude probabilities. So, for example, the partial crude subdistribution function

may be expressed in terms of the observable crude subdistribution functions as follows:

$$F_i^J(x) = \int_0^x \exp[-\Lambda_J^C(u)] \lambda_i^C(u) \, du, \quad i \in J, \quad (8)$$

where

$$\lambda_i^C(u) = \frac{dF_i^C(u)}{du} / S^C(u).$$

The above relationships hold whenever the latent failure times in J and \bar{J} are independent. It is not necessary that the failure times within J or \bar{J} be independent.

Issues Regarding the Use and Interpretation of Competing Risks

A major aim in many competing risks studies is the estimation of net survival probabilities. The ability to isolate the effect of one risk acting on a population is intuitively attractive, especially if the focus of a study is to evaluate the effect of an intervention that is targeted at reducing mortality from that specific cause. Of course, net survival probabilities are hypothetical quantities and not directly observable in a population; therefore they must be computed from the available information on the distribution of observables, or what we refer to as *crude probabilities*. Previously, we derived the net survival distribution for a specific risk D_i as a function of the observable crude probabilities under the assumption that the different latent failure times were independent of each other. The independence assumption is critical, because in this case the crude cause-specific hazard function is equal to the net hazard function, which leads to the important relationship given by (5).

In some situations such an assumption of independence may be reasonable. For example, when studying cause of death from a specific disease, it may be reasonable to assume that death from accidental causes is independent of those causes associated with the disease. Of course, there are other scenarios for which the independence assumption is not plausible. It is therefore important to consider the relationship of net probabilities to crude probabilities in the case where the latent failure times are not independent.

As we showed in (3), given any joint distribution of latent failure times, there exists a corresponding set of crude cause-specific hazard functions, or equivalently a set of crude cause-specific subdistribution

functions. Unfortunately, the converse is not true, as there exist many joint distributions, $H^C(t_1, \dots, t_k)$, that would result in the same set of crude subdistribution functions, $F_i^C(x)$, $i = 1, \dots, k$. These different joint distributions of latent failure times, each resulting in the same set of subdistribution functions, would lead to different net survival probabilities. Consequently we cannot identify net survival probabilities from corresponding crude probabilities. Because crude survival distributions define the observable random variables, we cannot estimate the net survival probabilities from observable data without making additional assumptions that cannot be verified from the observable data. Independence of the latent failure times is one assumption that would resolve the problem of identifiability and permit estimation of net probabilities; however, this assumption can never be verified. This problem of nonidentifiability was pointed out by Cox [13] and Tsiatis [31].

To get a sense of the extent of the nonidentifiability problem, Peterson [26] computed sharp bounds for net survival probabilities as a function of crude subdistribution functions. Specifically, he showed that

$$S^C(x) \leq S_i^i(x) \leq 1 - F_i^C(x).$$

Heuristically, these inequalities may be explained as follows. First, consider the hypothetical case that the causes of death are so highly correlated that an individual dying at time x from any cause other than D_i would have died from cause D_i immediately thereafter. For such a scenario the net survival probability at time x , $S_i^i(x)$, would be equal to the probability of surviving until time x from any cause, $S^C(x) = \Pr(X > x)$. At the other extreme, consider the hypothetical case where an individual who would die from any cause other than D_i would never die from cause D_i . Here, $\Pr(T_i \leq x) = 1 - S_i^i(x)$ would be equal to $F_i^C(x) = \Pr(X \leq x, \Delta = i)$. The upper and lower bounds for net survival probabilities may be quite substantial, as shown by Tsiatis [32].

This creates a philosophical dilemma in competing risks theory. Knowledge of the distribution of observable causes of death does not suffice to determine net survival probabilities. Only if additional assumptions are made on the joint distribution of the latent failure times are we able to identify uniquely the net survival probabilities. Two points of view have been taken in the literature. One is to restrict attention to certain dependency structures on the latent failure times that

allow for identification or, at least, restrict to a class of joint distributions where the bounds for the net survival probability are much tighter than the Peterson bounds. This has been the focus of research by Slud & Rubinstein [29], Klein & Moeschberger [23], and Zheng & Klein [33].

Another perspective is as follows. Because non-identifiability problems can only be handled by making additional assumptions that cannot be verified from the data, perhaps we should only consider making **inference** on the distribution of the observable random variables. That is, the focus should be on estimating cause-specific hazard and subdistribution functions and the comparison of such quantities under a variety of conditions that have practical importance. For example, comparisons may be made among different treatments or varying environmental conditions. This pragmatic point of view suggests that there is no reason to consider hypothetical quantities, such as net survival probabilities, because in fact we will never be in a position to evaluate one cause of death acting in isolation on a population. This point of view was eloquently presented by Prentice et al. [27].

This idea may be modified slightly in the case where a subset of the causes of death that are not of primary interest, denoted by \bar{J} , are thought a priori to be independent of the other causes of death, J , that are of interest. For example, certain accidental causes of death may fall into this category when studying treatment of disease. For these problems, inference using partial crude probabilities may be appropriate. We showed before how partial crude probabilities can be defined in terms of the distribution of the observable crude probabilities when causes J are independent of \bar{J} .

The Statistical Analysis of Competing Risk Data

Often, the data available for the analysis of competing risks are incomplete or right **censored**. This may be due to the termination of the study before all individuals fail, or to individuals who drop out of the study and subsequently are lost to follow-up. To accommodate this situation we extend the definition of competing risks to include censoring, i.e. we include an additional random variable, T_0 , that denotes the latent time to censoring. With this extended definition of competing risks, the observable data are defined by

6 Competing Risks

X^* and Δ^* , where $X^* = \min(T_0, \dots, T_k)$ and $\Delta^* = i$ if $X^* = T_i, i = 0, \dots, k$. We note that $\Delta^* = 0$ means that the failure time was censored at time X^* .

In a typical competing risks study we observe a sample of data (X_j^*, Δ_j^*, Z_j) , $j = 1, \dots, n$, where for the j th individual, X_j^* denotes the time to failure or censoring, Δ_j^* corresponds to cause of death or censoring, and Z_j corresponds to **covariate(s)** which we may use for modeling the distribution of competing risks. Using this extended notation, the observable data include censoring as a competing risk. We use an asterisk to denote the competing risks model that includes censoring. Therefore, the complete set of observable risks will be denoted by $C^* = 0, \dots, k$, in contrast to the risks of interest, $C = 1, \dots, k$, or perhaps some subset, J . In the previous section we denoted the complement of the subset J by $\bar{J} = C - J$; in the extended definition of competing risks we denote the complement of J by $\bar{J}^* = C^* - J$. In what follows it will be assumed that censoring, or risk 0, is independent of the other risks C . Without this assumption, nonidentifiability problems would not allow for **estimation** of the competing risk probabilities of interest regarding causes C .

One Sample Problems

Here we consider the problem of estimating relevant competing risk probabilities from a single sample of data $(X_j^*, \Delta_j^*), j = 1, \dots, n$.

Estimating Cause-Specific Hazard Functions

We showed before that the partial crude cause-specific hazard function is equal to the observable crude cause-specific hazard function whenever the risks in J are independent of the risks in \bar{J}^* , i.e.

$$\lambda_i^J(x) = \lambda_i^{C^*}(x). \quad (9)$$

Because censoring, or risk 0, is always assumed independent of the other risks, (9) will follow as long as the risks in J are independent of \bar{J} . It is important to note that the crude cause-specific hazard functions discussed in the previous section, $\lambda_i^C(x)$, are actually partial crude cause-specific hazard functions when we include censoring as a competing risk. However, because of (9) applied to $J = C$, $\lambda_i^C(x)$ is equal to the observable $\lambda_i^{C^*}(x)$. In the case when cause of death

D_i is independent of the other risks, the net-specific hazard function, $\lambda_i^i(x)$, is equal to $\lambda_i^{C^*}(x)$.

For certain independence assumptions, the cause-specific hazard functions are related to the observable crude cause-specific hazard functions, which by (2) is equal to

$$\lambda_i^{C^*}(x) = \frac{dF_i^{C^*}(x)}{dx} / S^{C^*}(x),$$

where

$$F_i^{C^*}(x) = \Pr(X^* \leq x, \Delta^* = i)$$

and

$$S^{C^*}(x) = \Pr(X^* > x).$$

The natural estimate for the crude subdistribution function is the empirical subdistribution function, i.e.

$$\hat{F}_i^{C^*}(x) = n^{-1} \sum_{j=1}^n I(X_j^* \leq x, \Delta_j^* = i),$$

where $I(\cdot)$ denotes the indicator function. This estimate puts mass $1/n$ at each observed event time from cause i . Similarly,

$$\hat{S}^{C^*}(x) = n^{-1} \sum_{j=1}^n I(X_j^* > x)$$

puts mass $1/n$ at each event time.

Because crude cause-specific hazards are functions of the crude subdistribution probabilities, the obvious estimates are obtained by substituting the corresponding functions of the empirical subdistribution probabilities. For example, the estimate of the cumulative cause-specific hazard function is

$$\hat{\Lambda}_i^{C^*}(x) = \int_0^x \frac{d\hat{F}_i^{C^*}(u)}{\hat{S}^{C^*}(u)} = \sum_{j=1}^n \frac{I(X_j^* \leq x, \Delta_j^* = i)}{Y(X_j^*)},$$

where $Y(u) = \sum_{j=1}^n I(X_j^* > u)$ denotes the number of individuals in the sample who are at risk at time u , i.e. neither died nor were censored. This estimator is the so-called **Nelson–Aalen estimator**; see Aalen [1]. Aalen [2, 3] derived the theoretical **large-sample** properties, including consistency and asymptotic normality, using the theory of **counting processes**.

This estimator of the i th crude cause-specific cumulative hazard is the appropriate estimator for the

partial crude cause-specific cumulative hazard whenever the causes in J are independent of the causes in \bar{J}^* , i.e.

$$\hat{\Lambda}_i^{J^*}(x) = \hat{\Lambda}_i^{C^*}(x), \quad i \in J.$$

In the special case where cause i is assumed independent of all other causes, the i th net-specific cumulative hazard function, $\Lambda_i^i(x)$, is estimated by $\hat{\Lambda}_i^{C^*}(x)$. The i th net survival distribution, $S_i^i(x)$, is equal to $\exp[-\Lambda_i^i(x)]$. Therefore, a natural estimator is the exponentiated negative of the Nelson–Aalen estimator. This estimator is

$$\hat{S}_i^i(x) = \exp - \left[\sum_{j=1}^n \frac{I(X_j^* \leq x, \Delta_j^* = i)}{Y(X_j^*)} \right].$$

Noting that this is equal to

$$\prod_{j=1}^n \exp \left[\frac{-I(X_j^* \leq x, \Delta_j^* = i)}{Y(X_j^*)} \right]$$

and that

$$\exp \left[\frac{-1}{Y(u)} \right] \approx \left[\frac{1-1}{Y(u)} \right],$$

yields the approximation

$$\hat{S}_i^i(x) \approx \prod_{j=1}^n \left[\frac{1-1}{Y(X_j^*)} \right]^{I(X_j^* \leq x, \Delta_j^* = i)}.$$

This is the well known **Kaplan–Meier** [21], or product-limit, estimator. The asymptotic equivalence of the exponentiated Nelson–Aalen estimator and the Kaplan–Meier estimator, and the large-sample properties of these estimators, are given by Breslow & Crowley [11].

It is important to note that the Kaplan–Meier estimator, by construction, is a **consistent estimator** of the exponentiated cumulative crude cause-specific hazard function. That this corresponds to an estimator of the net survival distribution follows only when the net hazard function is equal to the crude cause-specific hazard, i.e. when cause i is independent of all the other causes, including censoring. Without this assumption, the Kaplan–Meier estimator of the i th net-specific survival distribution does not estimate any interesting or relevant probability.

If we consider death from any cause, i.e. $\Delta \in C$, then the estimate of the corresponding survival distribution, $S^C(x)$, from a sample of potentially

censored data (X_j^*, Δ_j^*) , $j = 1, \dots, n$, follows from applying the same logic:

$$\hat{S}^C(x) = \prod_{j=1}^n \left[\frac{1-1}{Y(X_j^*)} \right]^{I(X_j^* \leq x, \Delta_j^* \in C)}.$$

This estimator for the survival distribution from any cause of death in the presence of censoring is the Kaplan–Meier estimator as originally presented in the seminal paper [21] in 1958. Failure is considered a death from any cause, and an incomplete observation is a censored observation. The estimator of the i th net survival function given above is also referred to as a Kaplan–Meier estimator, since it may be derived via the same formula, letting failure be death from cause i and an incomplete observation be death from any cause other than i or censoring.

Estimating Subdistribution Functions

We may use the above results to derive **nonparametric** estimators for crude and partial crude subdistribution functions. Using (2), the i th crude subdistribution function may be expressed as

$$F_i^C(x) = \int_0^x S^C(u) \lambda_i^C(u) du.$$

Because censoring is independent of the other causes of death, $\lambda_i^C(u) = \lambda_i^{C^*}(u)$. Therefore a natural estimator for the i th subdistribution function is given by

$$\hat{F}_i^C(x) = \int_0^x \hat{S}^C(u) \frac{d\hat{F}_i^{C^*}(u)}{\hat{S}^{C^*}(u)},$$

where $\hat{S}^C(u)$ is the Kaplan–Meier estimator for the survival distribution of time to death from any cause.

The large-sample statistical properties of this estimator may be derived using the theory of counting processes. Details may be found in Aalen [2, 3], Fleming [16, 17], Benichou & Gail [6], and Andersen et al. [5] when using **cohort** data, and in Benichou & Gail [7] when using **population-based case-control data**.

The Relationship of Competing Risks to Covariates

Often, we are interested in studying the relationship of time to death from one or many causes to other

covariates. For example, we may be interested in the effect of different treatments on reducing the risk of death from specific causes, or we may wish to model the relationship of competing risk probabilities to other **prognostic factors**. These problems are generally posed in terms of hypothesis testing or estimation of **regression** parameters. There is a wide literature on inferential techniques for hypothesis testing and regression modeling for survival problems with censored data. Because of the close relationship between censoring and competing risks, many of the methods developed for analyzing censored survival data may also be applied to competing risks data (*see Survival Analysis, Overview*).

Hypothesis Testing

The most widely used methods for testing the **null hypothesis** of no treatment effect among K treatments with censored survival data are the **logrank** or weighted logrank tests. These tests were designed to test the equality of the hazard functions for death among K treatments when the censoring time is independent of time to death within each treatment group. If we study these tests carefully, then we realize that they actually compare the observable cause-specific hazard functions among the different treatment groups. Therefore, we can immediately apply these methods for testing equality of cause-specific hazard functions among different treatments. To be more precise, we denote by $\lambda_{il}^{C^*}(x)$, $l = 1, \dots, K$, the i th cause-specific hazard function within treatment group l . The weighted logrank tests may then be used to test the null hypothesis that

$$\lambda_{i1}^{C^*}(x) = \dots = \lambda_{iK}^{C^*}(x), \quad x > 0.$$

The theoretical development for these tests is given by Andersen et al. [4]. This is carried out by letting failure correspond to death from cause i and an incomplete observation to correspond to death from any cause other than i or censoring ($\Delta^* = 0$).

We reiterate the interpretation of this null hypothesis and the results of the logrank test. If we are willing to assume that time to death from cause i is independent of the times to death from other causes as well as time to censoring, within each treatment group $l = 1, \dots, K$, then the cause-specific hazard function, $\lambda_{il}^{C^*}(x)$, is equal to the net-specific, or marginal, hazard function, $\lambda_{il}^l(x)$. Equality of net-specific hazard functions implies equality of net-specific survival

probabilities. Therefore, with the assumption of independence, the logrank test is a test of the null hypothesis that the K net-specific survival distributions are equal. This is often the hypothesis of interest.

To illustrate, consider a **clinical trial** of several treatments to reduce breast cancer mortality. Because breast cancer clinical trials generally occur over many years, some patients may die from causes other than breast cancer. Because the treatments were targeted to reduce breast cancer mortality, the investigators are not interested in the effect that treatment may have on other causes of death; rather, they are mainly interested in the effect of the treatments on breast cancer mortality in the absence of causes of death other than breast cancer. This is the classic competing risks problem of comparing net survival distributions. When the logrank test is used, patients not dying from breast cancer are treated as censored observations. As previously discussed, this is an appropriate test for the equality of net survival probabilities when the time to death from other causes is independent of time to death from breast cancer within each treatment group. This assumption may not be true, and in fact cannot be verified with the data because of nonidentifiability problems alluded to above. If this independence assumption is not true, then it is not clear what we are testing when we use the logrank test.

One way around this philosophical dilemma is to consider only tests of observable population parameters. An important observable population parameter is the crude cause-specific hazard function, $\lambda_i^C(x)$. We again emphasize that the population cause-specific hazard function, $\lambda_i^C(x)$, is observable only if there is no additional censoring introduced. With the introduction of censoring, the observable parameter is $\lambda_i^{C^*}(x)$. However, by assumption, the censoring ($\Delta^* = 0$) is independent of the other causes of death, in which case $\lambda_i^C(x) = \lambda_i^{C^*}(x)$.

As we pointed out, the logrank test tests the equality of the cause-specific hazard functions, $\lambda_{il}^{C^*}(x)$, and, with independent censoring, the equality of $\lambda_{il}^C(x)$. Therefore, the logrank test would be a valid test of the equality of the breast cancer specific hazard functions among the K treatments. Although this cause-specific hazard function may not be directly related to net-specific breast cancer mortality, if independence does not hold, then it still may be an important comparison. This point of view is given by Prentice et al. [27].

Another observable quantity is the subdistribution function $F_{il}^C(x)$ for cause i within treatment group l . Very little work has been done on deriving tests for the equality of these K -sample subdistribution functions. One exception is a class of tests derived by Gray [20] to test the null hypothesis that

$$F_{il}^C(x) = \dots = F_{iK}^C(x).$$

Regression Modeling

The most popular framework for modeling the association of censored survival data to prognostic variables is with the **proportional hazards** model of Cox [14] (see **Cox Regression Model**). In this model the hazard for death is related to a vector of covariates by

$$\lambda(t|\mathbf{z}) = \lambda_0(t) \exp(\beta^T \mathbf{z}),$$

where \mathbf{z} represents a vector of covariates, and $\lambda(t|\mathbf{z})$ is the hazard rate of death at time t given covariates \mathbf{z} . In this model, censoring is assumed to be independent of the failure time, conditional on the covariates. A careful study of the inferential procedure for estimating parameters in the Cox model reveals that this is actually the observable cause-specific hazard of death in the presence of censoring. That this corresponds to the actual net hazard function of death holds only when we add the assumption of independence of censoring time and failure time.

Consequently, this model may also be applied to competing risks data; that is, we may use the same inferential procedures to estimate the parameter β when considering the model

$$\lambda_i^{C^*}(t|\mathbf{z}) = \lambda_{i0}^{C^*}(t) \exp(\beta^T \mathbf{z}).$$

To apply software for the Cox model (see **Survival Analysis, Software**), we must define a failure as death from cause i , and an incomplete observation as either death from a cause other than i or censoring. The interpretation of this model and the parameters is the same as discussed above. That is, if we are willing to assume that time to death from cause i is independent of the times to death from other causes and time to censoring, then the observable cause-specific hazard, $\lambda_i^{C^*}(t|\mathbf{z})$, is equal to the net-specific hazard, $\lambda_i^i(t|\mathbf{z})$.

Even if we are unwilling to make this nonidentifiable assumption, the relationship of the observable cause-specific hazard to covariates may be of interest.

By assumption, censoring is independent of all other causes of death. This implies that $\lambda_i^C(t|\mathbf{z}) = \lambda_i^{C^*}(t|\mathbf{z})$. Therefore, the results of the Cox regression analysis may be used to estimate the parameters in the model of the cause-specific hazard function, given by

$$\lambda_i^C(t|\mathbf{z}) = \lambda_{i0}^C(t) \exp(\beta^T \mathbf{z}).$$

Using cause-specific hazards thus allows useful interpretation of relevant observable quantities without an additional assumption of independence of the different causes of death.

References

- [1] Aalen, O. (1976). Nonparametric inference in connection with multiple decrement models, *Scandinavian Journal of Statistics* **3**, 15–27.
- [2] Aalen, O. (1978). Nonparametric estimation of partial transition probabilities in multiple decrement models, *Annals of Statistics* **6**, 534–545.
- [3] Aalen, O. (1978). Nonparametric inference for a family of counting processes, *Annals of Statistics* **6**, 701–726.
- [4] Andersen, P.K., Borgan, O., Gill, R.D. & Kieding, N. (1982). Linear nonparametric tests for comparison of counting processes, with applications to censored survival data, *International Statistical Review* **50**, 219–258.
- [5] Andersen, P.K., Borgan, O., Gill, R.D. & Kieding, N. (1992). *Statistical Models Based on Counting Processes*. Springer-Verlag, New York, pp. 299–301.
- [6] Benichou, J. & Gail, M.H. (1990). Estimates of absolute risk in cohort studies, *Biometrics* **46**, 813–826.
- [7] Benichou, J. & Gail, M.H. (1995). Methods of interferences for estimates of absolute risk derived from population-based case–control studies, *Biometrics* **51**, 182–194.
- [8] Berkson, J. & Elveback, L. (1960). Competing exponential risks with particular reference to the study of smoking and lung cancer, *Journal of the American Statistical Association* **55**, 415–428.
- [9] Bernoulli, D. (1760). Essai d'une nouvelle analyse de la mortalité causée par la petite vérole, et des avantages de l'inoculation pour le prévenir. *Histoire avec les Mémoires, Académie Royal des Sciences*. Paris, pp. 1–45.
- [10] Birnbaum, Z.W. (1979). *On the Mathematics of Competing Risks*. US Department of Health, Education and Welfare, Washington.
- [11] Breslow, N. & Crowley, J. (1974). A large sample study of the life table and product limit estimates under random censorship, *Annals of Statistics* **2**, 437–453.
- [12] Chiang, C.L. (1968). *Introduction to Stochastic Processes in Biostatistics*. Wiley, New York, Chapter 11.

- [13] Cox, D.R. (1959). The analysis of exponentially distributed life-times with two types of failure, *Journal of the Royal Statistical Society, Series B* **21**, 411–421.
- [14] Cox, D.R. (1972). Regression models and life tables, *Journal of the Royal Statistical Society, Series B* **34**, 187–220.
- [15] David, H.A. & Moeschberger, M.L. (1978). The Theory of Competing Risks, Griffin's Statistical Monograph No. 39. Macmillan, New York.
- [16] Fleming, T.R. (1978). Nonparametric estimation for nonhomogeneous Markov processes in the problem of competing risks, *Annals of Statistics* **6**, 1057–1070.
- [17] Fleming, T.R. (1978). Asymptotic distribution results in competing risks estimation, *Annals of Statistics* **6**, 1071–1079.
- [18] Gail, M. (1975). A review and critique of some models used in competing risk analysis, *Biometrics* **31**, 209–222.
- [19] Gail, M. (1982). Competing risks, in *Encyclopedia of Statistical Sciences*, Vol. 2, S. Kotz & N.L. Johnson, eds. Wiley, New York, pp. 75–81.
- [20] Gray, R.J. (1988). A class of k -sample tests for comparing the cumulative incidence of a competing risk, *Annals of Statistics* **16**, 1141–1154.
- [21] Kaplan, E.L. & Meier, P. (1958). Non-parametric estimation from incomplete observations, *Journal of the American Statistical Association* **53**, 457–481.
- [22] Karn, M.N. (1931). An inquiry into various death rates and the comparative influence of certain diseases on the duration of life, *Annals of Eugenics* **4**, 279–326.
- [23] Klein, J.P. & Moeschberger, M.L. (1988). Bounds on net survival probabilities for dependent competing risks, *Biometrics* **44**, 529–538.
- [24] Makeham, W.M. (1874). On the law of mortality, *Journal of the Institute of Actuaries* **18**, 317–322.
- [25] Moeschberger, M.L. & David, H.A. (1971). Life tests under competing causes of failure and the theory of competing risks, *Biometrics* **27**, 909–933.
- [26] Peterson, A.V. (1976). Bounds for a joint distribution function with fixed subdistribution functions: applications to competing risks, *Proceedings of the National Academy of Sciences* **73**, 11–13.
- [27] Prentice, R.L., Kalbfleisch, J.D., Peterson, A.V., Flournoy, N., Farewell, V.T. & Breslow, N.E. (1978). The analysis of failure time data in the presence of competing risks, *Biometrics* **34**, 541–554.
- [28] Sampford, M.R. (1952). The estimation of response time distributions. II. Multistimulus distributions, *Biometrics* **8**, 307–353.
- [29] Slud, E.V. & Rubinstein, L.V. (1983). Dependent competing risks and summary survival curves, *Biometrika* **70**, 643–650.
- [30] Todhunter, J. (1949). *A History of the Mathematical Theory of Probability*. Chelsea, New York.
- [31] Tsiatis, A.A. (1975). A nonidentifiability aspect of the problem of competing risks, *Proceedings of the National Academy of Sciences* **72**, 20–22.
- [32] Tsiatis, A.A. (1978). An example of nonidentifiability in competing risks, *Scandinavian Actuarial Journal* **1978**, 235–239.
- [33] Zheng, M. & Klein, J.P. (1994). A self-consistent estimator of marginal survival functions based on dependent competing risk data and the assumed copula, *Communications in Statistics – Theory and Methods* **23**, 2299–2311.

A.A. TSIATIS