Published in final edited form as:

Lifetime Data Anal. 2016 January ; 22(1): 17–37. doi:10.1007/s10985-015-9324-1.

Comparing center-specific cumulative incidence functions

Ludi Fan

Eli Lilly and Company, 893 S Delaware St., Indianapolis, IN, 46285, lfan@umich.edu

Abstract

The competing risks data structure arises frequently in clinical and epidemiologic studies. In such settings, the cumulative incidence function is often useful to describe the ultimate occurrence of a particular cause of interest. If the objective of the analysis is to compare subgroups of patients with respect to cumulative incidence, imbalance with respect to group-specific covariate distributions must generally be factored out, particularly in observational studies. This report proposes a measure to contrast center- (or, more generally group-) specific cumulative incidence functions (CIF). One such application involves evaluating organ procurement organizations (OPO) with respect to the cumulative incidence of kidney transplantation. In this case, the competing risks include (i) death on the wait-list and (ii) removal from the wait-list. The proposed method assumes proportional cause-specific hazards, which are estimated through Cox models stratified by center. The proposed center effect measure compares the average CIF for a given center to the average CIF that would have resulted if that particular center had covariate pattern-specific cumulative incidence equal to that of the national average. We apply the proposed methods to data obtained from a national organ transplant registry.

Keywords

Cox regression; center effect; competing risks; cumulative incidence function; kidney transplantation

1 Introduction

In many studies with time to event data, comparisons among groups are of key interest. These groups can be defined by country, state, region or treatment center, for example. A natural grouping that often occurs in medical data is defined by treatment center or hospital. It is often of interest to study differences in outcomes by location, since geographic and center-based disparities can be present in health outcomes, and would be of concern in settings where uniformity in quality of care is a priority. There are a number of existing methods for estimating group or center effects. Harrington and Fleming (1982) developed a class of non-parametric methods based on the rank test that accommodates comparisons of more than two groups. Dabrowska and Doksum (1988) developed estimation and inference methods for comparison of the generalized odds-rate model in the two-group setting. Most of these methods assume, either explicitly or implicitly, that all events are of the same cause.

Often in survival data, the terminating event is due to one of several possible causes. Experiencing a particular type of event precludes an individual from experiencing any other types of events. Such events are typically referred to as competing risks (Prentice et al., 1978). An example of the competing risks setting would be time until death, with cause of death categorized into cardiovascular disease, cancer, accidental, and all other causes. This data structure occurs frequently, and needs to be explicitly acknowledged for the analysis to be accurate. Analyses that ignore the competing risks aspect of the data by not differentiating between the types of events may lose information on the covariate effects and lead to inaccurate interpretation of results. Several methods have been proposed for analysis of competing risks data. Benichou and Gail (1990) developed a method for estimating the absolute risk of an event during a time interval for a particular covariate pattern. The authors investigated the results of using the exponential, piecewise exponential, and Cox models (Cox, 1972) for the hazard functions. Cheng et al. (1998) investigated the prediction of cumulative incidence functions (CIF) and simultaneous confidence bands. Fine and Gray (1999) developed regression methods for the subdistribution hazard function.

Evaluating treatment centers and providers have been the focus of several papers. For example, Logan et al. (2008) developed an approach based on pseudo-values to compare center-specific 1-year survival outcomes transplant center network average. DeLong et al. (1997) compared methods to evaluate provider performance based on eight different risk adjustment algorithms. Inpatient mortality was the outcome and their methods employed either external or internal risk adjustment. Ultimately, the method that the authors recommended consists of a logistic model with random effects for the provider with an external risk-adjustment algorithm.

There are few methods for comparing groups or centers in the presence of competing risks. Gray (1988) developed a method to compare CIFs among multiple groups, nonparametrically and unadjusted for covariates. Zhang and Fine (2008) then summarized differences for several transformations of the CIF for two groups. We propose methods that compare centers in the presence of competing risks. The methods are targeted at observational data and, therefore, account for differences in group-specific covariate distributions.

The methods proposed in this report were motivated by data arising from the end-stage renal disease (ESRD) setting. A patient who experiences ESRD onset (also referred to as 'renal failure') typically begins ESRD therapy by receiving dialysis. If medically suitable, the patient may subsequently be placed on a wait-list for deceased-donor kidney transplantation. Patients often have to wait to receive an organ transplant because there are not nearly enough donor kidneys to accommodate all wait-listed patients (www.usrds.org). For the purpose of solid organ transplantation, the United States is divided into 11 Regions, each subdivided into donation service areas. For each donation service area, an organ procurement organization (OPO) maintains a wait-list (i.e., each OPO has its own wait-list) and is responsible for allocating organs to patients within its corresponding donation service area. We use the general terms group and center to indicate any factor that subdivides the study population. In the context of organ transplantation, the unit of our interest is an OPO. This setting corresponds to a competing risks framework because a patient on a wait-list will

generally experience one of three mutually exclusive events: (1) receipt of a kidney transplant (2) death on the wait-list (3) removal from the wait-list. The competing risks for receipt of a kidney transplant are then death while on wait-list and removal from the wait-list, with removal usually occurring when a patient's general health has declined to the point where transplantation is considered futile. Thus both removal and wait-list death represent the patient not having received a transplant quickly enough.

To be wait-listed by an OPO, a patient often must be able to travel to the location at which the transplant will occur within a relatively short period of time after being notified that a particular organ is available. Thus, since most individuals do not have unlimited resources to travel long distances within short notice, where a given patient can be put on a wait-list is largely determined by where he/she resides. A natural question to ask is how each OPO compares relative to the national average in terms of a wait-listed patient's cumulative incidence of kidney transplantation (i.e., acknowledging that the patient may instead die while on, or be removed from, the wait-list). In this report, our goal is to propose a metric that will compare the experience of wait-listed patients at an OPO to the experience that would have been observed if that particular OPO transplanted patients at a rate equal to the national average. Data were obtained from the Scientific Registry of Transplant Recipients (SRTR), a national population-based organ transplant registry.

There are two frameworks for casting the event times in a competing risks setting. The first method is based on latent failure times (Gail, 1975; Crowder, 2001) in which there exists a latent event time for every cause, but only the minimum of the latent event times is observed (Cox, 1959; Moeschberger and David, 1971). For marginal quantities (such as the failure-type-specific survival functions) to be identifiable under the latent failure times set up, each of the latent failure times must act independently. The second framework assumes there is only one event time for each subject, with the event occurring from one of two or more causes. Under this framework, there only exists one failure time, due to one cause. For example, if a particular type of event occurs, then the event times due to other causes are undefined. The key functions that arise from this framework are the crude functions known as cause-specific hazard (CSH) and the cumulative incidence function (CIF) (Chiang, 1968; Anderson et al., 1993; Kalbfleisch and Prentice, 2002). These models do not require the assumption that the causes are independent (Tsiatis, 1975).

In this report, we propose methods that contrast centers with respect to cumulative incidence. We model the cause-specific hazards explicitly through Cox regression, with the CSHs then combined and integrated, then transformed to obtain the CIF. The effect measure, or, the effect of a particular OPO on a specific cause, is obtained through appropriately averaging fitted values. The event times are subject to right censoring, which is assumed to be independent of the event time given the covariates.

In Section 2, we introduce the notation and present the proposed method and estimation procedures. Section 3 describes the asymptotic properties of the proposed estimators. In Section 4 we evaluate the finite sample performance of the proposed estimators in simulation studies. We return to the motivating example in Section 5, applying the proposed

methods to data from the SRTR to compare OPOs across the United States. Section 6 concludes with a discussion.

2 Proposed Methods

In this section, we first set up the requisite notation. We then describe the proposed center (or more generally, group) effect measures, and corresponding estimation procedures.

2.1 Data structure and notation

Let T_i and C_i be the failure time and censoring time, respectively, for individual i ($i = 1, \dots, n$). The observation time is then defined as $X_i = T_i \wedge C_i$, where $a \wedge b = \min\{a, b\}$. Center and cause will be denoted by j ($j = 1, \dots, J$) and k ($k = 1, \dots, K$), respectively. There are n individuals in the entire sample, with n_j individuals in center j . For concreteness, we refer to the ‘group’ of interest as center, although in practice the grouping variate could be any categorical factor (or combination of factors) defining subgroups of individuals. Let δ_i be the cause of failure for subject i , with $\delta_i = 0$ if $T_i > C_i$. Let A_i represent the center to which subject i belongs ($A_i = 1, \dots, J$), and set $A_{ij} = I(A_i = j)$, where $I(\cdot)$ is an indicator function taking the value 1 if its argument is true and 0 otherwise. The observed data consist of $\{X_i, \delta_i, \mathbf{Z}_i, A_i\}$, where $\mathbf{Z}_i$ is the vector of covariates which is assumed to be time constant. The at-risk indicator is given by $Y_{ij}(t) = I(X_i \geq t, A_i = j)$ and the counting process for subject i in center j , cause k is denoted by $N_{ijk}(t) = I(X_i \geq t, A_i = j, \delta_i = k)$. Let the CIF for cause k for individual i , at center j be denoted by

$$F_{ijk}(t) \equiv P(T_i \leq t, \Delta_i = k | A_i = j, \mathbf{Z}_i), \quad (1)$$

interpreted as the probability (competing risks version thereof) that individual i in center j experiences an event of type k by time t .

2.2 Center effect measures: cumulative incidence

For the purposes of describing the proposed measures for contrasting center-specific cumulative incidence functions, it is useful to remove the subscript i , such that the expressions that follow refer to a hypothetical patient. Recall from (1) that j indexes center, while k denotes cause of chief interest. Consider an evaluation of center j with respect to the cumulative incidence function for cause k . goal is to contrast the average CIF for center j with an appropriate average across all centers. A patient with covariate $\mathbf{Z}$ is subject to cause k cumulative incidence $F_k(t|A = j, \mathbf{Z})$ if treated at center j . Now, consider the average cumulative incidence (across all centers) for that same patient, which can be represented by $E_A[F_k(t|A, \mathbf{Z})]$, where E_A denotes the expectation with respect to the marginal distribution of A . This amounts to taking a weighted average of $F_k(t|A = \ell, \mathbf{Z})$, with weight $p_\ell = P(A = \ell)$ for $\ell = 1, \dots, J$. To evaluate center j , we contrast $F_k(t|A = j, \mathbf{Z})$ and $E_A[F_k(t|A, \mathbf{Z})]$ after averaging with respect to the center j covariate distribution. Specifically, the effect of center j with respect to cause k cumulative incidence is defined as

$$\delta_{jk}(t) = E[F_k(t|A = j, \mathbf{Z}) | A = j] - E[E_A[F_k(t|A, \mathbf{Z})] | A = j], \quad (2)$$

where the outer expectation of each term in (2) is with respect to the conditional distribution of $\mathbf{Z}|A = j$. The effect measure $\delta_{jk}(t)$ compares, through the CIF, two scenarios: (i) the observed reality: all individuals assigned to center j are treated at center j (ii) all individuals assigned to center j are instead treated at a hypothetical center with cumulative incidence equal to that of the national average.

2.3 Estimating the center effects

One option for obtaining the CIFs is to use the cause-specific hazards. The cause k CIF for individual i at center j is

$$F_{ijk}(t) = \int_0^t S_{ij}(s) \lambda_{ijk}^{\#}(s) ds, \quad (3)$$

where $S_{ij}(t) = P(T_i > t | \mathbf{Z}_i, A_i = j)$ is the survival function for individual i at center j , and the cause-specific hazards are given by

$$\lambda_{ijk}^{\#}(t) = \lim_{\varepsilon \downarrow 0} P(t \leq T_i < t + \varepsilon, \Delta_i = k | T_i \geq t, \mathbf{Z}_i, A_i = j)$$

for $k = 1, \dots, K$. Note that $F_{ijk}(t)$ can be written entirely in terms of the cause-specific hazard, since $S_{ij}(t) = \exp\{-\sum_{k=1}^K \Lambda_{ijk}^{\#}(t)\}$, where the cumulative CSH for individual i at center j is $\Lambda_{ijk}^{\#}(t) = \int_0^t \lambda_{ijk}^{\#}(s) ds$. Also called the subdistribution function, $F_{ijk}(t)$ is the probability that subject i experiences an event of type j by time t , acknowledging that he/she could experience another event first, which would preclude event k from happening. From (3), it can be seen that the CIF is a function of all of the CSHs, so that the CIF acknowledges the presence of other types of events. As $t \rightarrow \infty$, $F_{ijk}(t) \rightarrow P(\Delta_i = k | \mathbf{Z}_i, A_i = j)$. Therefore, unlike the cumulative distribution function in the single cause setting, the CIF will generally not approach 1 if the competing causes have non-zero probabilities.

In this report, Cox regression stratified on center is assumed to relate the covariates to the cause-specific hazards,

$$\lambda_{ijk}^{\#}(t) = \lambda_{0jk}(t) \exp\{\boldsymbol{\beta}_k^T \mathbf{Z}_i\}. \quad (4)$$

The Cox model is selected due to its flexibility and popularity. Stratifying on center adjusts for center effects non-parametrically. As such, we are assuming proportionality with respect to the cause k hazard at time t among subjects alive at time t with respect to the adjustment covariates, $\mathbf{Z}_i$, although not the centers.

A few comments on model (4) are in order. First, the stratification by center would make it difficult to estimate the effect of center j on $\lambda_{ijk}^{\#}(t)$ based on model (4) alone. However, as defined previously, the cause-specific hazards are not of primary interest per se, and are useful only through their connection to $F_{ijk}(t)$. Second, the covariate vector is assumed to be constant. If not, accurate estimation of $F_{ijk}(t)$ would be substantially more complicated, and the methods proposed in this report would not be recommended. Third, covariate effects, $\boldsymbol{\beta}_k$,

are allowed to vary by cause but are assumed to be constant over time. This could be relaxed through additional stratification, or by parametrically modeling covariate effects that are time-dependent; e.g., through the Cox non-proportional hazards model; see Klein and Moeschberger (2003).

We estimate $\delta_{jk}(t)$ by using the finite sample estimators for the quantities involved. The covariate effects β_k are estimated through partial likelihood (Cox, 1975), while the cumulative baseline cause-specific hazards $\Lambda_{0jk}^\#(t)$ are estimated by the Breslow estimator (Breslow, 1972). Such methods are valid under independent censoring, which can be formally written as

$$\lim_{\varepsilon \rightarrow 0} \frac{1}{\varepsilon} P(t \leq T_i < t + \varepsilon, \Delta_i = k | T_i > t, C_i > t, \mathbf{Z}_i, A_i) = \lim_{\varepsilon \rightarrow 0} \frac{1}{\varepsilon} P(t \leq T_i < t + \varepsilon, \Delta_i = k | T_i > t, \mathbf{Z}_i, A_i)$$

for $k = 1, \dots, K$.

Referring back to (2), the quantity $E[F_k(t) | A = j, \mathbf{Z}]$ can be estimated by taking the average of $\hat{F}_{ijk}(t)$ with respect to the empirical distribution of $[\mathbf{Z} | A = j]$. The CIF for individual i at the hypothetical national average center is estimated by taking a weighted

average of the CIFs for individual i across all centers, expressed as $\sum_{\ell=1}^J \hat{F}_{i\ell k}(t) \hat{p}_\ell$, with $\hat{p}_\ell = n_\ell / n$ an estimator for $p_\ell = P(A_i = \ell)$. Combining these estimators, we then estimate the proposed center effect measure by

$$\hat{\delta}_{jk}(t) = \frac{1}{n_j} \sum_{i=1}^n I(A_i = j) \hat{F}_{ijk}(t) - \frac{1}{n_j} \sum_{i=1}^n I(A_i = j) \left\{ \sum_{\ell=1}^J \hat{F}_{i\ell k}(t) \hat{p}_\ell \right\}, \quad (5)$$

where $\hat{F}_{ijk}(t) = \int_0^t \exp\left\{-\sum_{k=1}^K \hat{\Lambda}_{ijk}^\#(s)\right\} d\hat{\Lambda}_{ijk}^\#(s)$ and $\hat{\Lambda}_{ijk}^\#(t) = \int_0^t \exp(\hat{\beta}_k^T \mathbf{Z}_i) d\hat{\Lambda}_{0jk}^\#(s)$. Note that the use of $\hat{p}_\ell$ in (5) is one of several reasonable choices for constructing the comparator group, a point we discuss further in Section 6.

We now describe the asymptotic properties of the proposed estimators.

3 Asymptotic Properties

We begin by listing the assumed regularity conditions for $i = 1, \dots, n$, $j = 1, \dots, J$, and $k = 1, \dots, K$.

- a. $\{X_i, \mathbf{Z}_i, A_i\}$ are independent and identically distributed.
- b. $P(Y_{ij}(\tau) = 1) > 0$.
- c. $P(A_i = j | \mathbf{Z}_i) > 0$.
- d. $|\mathbf{Z}_{iq}| \leq K$, where $\mathbf{Z}_{iq}$ is the q th element of $\mathbf{Z}_i$ and K is a constant.
- e. $\int_0^\tau \lambda_{0jk}^\#(t) dt < \infty$.
- f. Continuity of the following functions:

$$\mathbf{r}_{jk}^{(d)}(t; \boldsymbol{\beta}) = E \left[Y_{ij}(t) \mathbf{Z}_i^{\otimes d} \exp(\boldsymbol{\beta}_k^T \mathbf{Z}_i) \right], d=0, 1, 2.$$

with $r_{jk}^{(0)}(t; \boldsymbol{\beta})$ bounded away from 0 for $t \in (0, \tau]$.

g. Positive-definiteness of the following matrices:

$$\mathbf{I}_k(\boldsymbol{\beta}_k) = E \left[\sum_{j=1}^J \int_0^\tau \left\{ \frac{\mathbf{r}_j^{(2)}(t; \boldsymbol{\beta}_k)}{r_j^{(0)}(t; \boldsymbol{\beta}_k)} - \bar{\mathbf{z}}_j(t; \boldsymbol{\beta}_k)^{\otimes 2} \right\} dN_{ijk}(t) \right],$$

$$\text{where } \bar{\mathbf{z}}_j(t; \boldsymbol{\beta}) = \mathbf{r}_{jk}^{(1)}(t; \boldsymbol{\beta})^{-1} r_{jk}^{(0)}(t; \boldsymbol{\beta}).$$

Condition (a) permits fairly standard applications of the Central Limit Theorem. Conditions (b) and (c) are required for identifiability of $\delta_{jk}(t)$, with (c) being analogous to the positivity assumption familiar to causal inference methodology. Condition (e) ensures boundedness of many integrals arising in the asymptotic development. The asymptotic framework we work with essentially requires $n_j \rightarrow \infty$. With the number of centers J remaining constant, $n_j \rightarrow \infty$ occurs when $n \rightarrow \infty$ due to Condition (c).

Theorem 1

Under Conditions (a) through (g), $\hat{F}_{ijk}(t)$ converges almost surely to $F_{ijk}(t)$ and $\hat{\delta}_{jk}(t)$ converges almost surely to $\delta_{jk}(t)$.

The proof of Theorem 1 involves casting $\hat{F}_{ijk}(t)$ as a functional of consistent estimators $\hat{\boldsymbol{\beta}}_k$ and $\hat{\Lambda}_{0jk}^\#(t)$. Consistency then follows from successive applications of the Continuous Mapping Theorem.

Theorem 2

Under conditions (a) through (g), $n^{\frac{1}{2}} \{ \hat{\delta}_{jk}(t) - \delta_{jk}(t) \}$ converges asymptotically to a zero-mean Gaussian process with variance function $\sigma_{jk}(t) = E [\psi_{mjk}(t; \boldsymbol{\beta})^2]$ where

$$\psi_{mjk}(t; \boldsymbol{\beta}) = \sum_{d=1}^4 \left\{ \phi_{\ell_{jk}}^d(t; \boldsymbol{\beta}) - \phi_{\ell_k}^d(t; \boldsymbol{\beta}) \right\} + o_p(1), \text{ with } \boldsymbol{\beta}^T = [\boldsymbol{\beta}_1^T, \dots, \boldsymbol{\beta}_k^T] \text{ and}$$

$$\phi_{\ell_k}^d(t; \boldsymbol{\beta}) = E_A \left[\phi_{\ell_{jk}}^d(t; \boldsymbol{\beta}) \right] \text{ and where we define}$$

$$\phi_{\ell_{jk}}^1(t; \boldsymbol{\beta}) = -\frac{1}{p_j} E \left[\sum_{m=1}^K A_{ij} \int_0^t \mathbf{D}_{ijm}^T(s; \boldsymbol{\beta}_m) dF_{ijk}(s) \right] \mathbf{I}_m(\boldsymbol{\beta}_m)^{-1} \mathbf{U}_{\ell_m}(\boldsymbol{\beta}_m),$$

$$\phi_{\ell_{jk}}^2(t; \boldsymbol{\beta}) = -\sum_{m=1}^K \int_0^t \frac{1}{p_j} E \left[A_{ij} \{ F_{ijk}(t) - F_{ijk}(u) \} Y_{ij}(u) \exp(\boldsymbol{\beta}_m^T \mathbf{Z}_i) r_{jk}^{(0)}(u; \boldsymbol{\beta}_m)^{-1} \right] dM_{\ell_{jm}}(u; \boldsymbol{\beta}_m),$$

$$\phi_{\ell jk}^3(t; \boldsymbol{\beta}) = \frac{1}{p_j} E \left[A_{ij} \int_0^t [\mathbf{Z}_i - \bar{\mathbf{z}}_j(s; \boldsymbol{\beta})]^T dF_{ijk}(s) \right] \mathbf{I}_k(\boldsymbol{\beta}_k)^{-1} \mathbf{U}_{\ell k}(\boldsymbol{\beta}_k),$$

$$\phi_{\ell jk}^4(t; \boldsymbol{\beta}) = \int_0^t \frac{1}{p_j} E \left[A_{ij} S_{ij}(s; \boldsymbol{\beta}_1, \boldsymbol{\beta}_2) Y_{ij}(s) \exp(\boldsymbol{\beta}_k^T \mathbf{Z}_i) r_{jk}^{(0)}(s; \boldsymbol{\beta}_2)^{-1} \right] dM_{\ell jk}(s; \boldsymbol{\beta}_k),$$

where $\mathbf{I}_k(\boldsymbol{\beta}_k)$ and $\bar{\mathbf{z}}_j(t; \boldsymbol{\beta})$ are defined in Condition (g), $\mathbf{r}_{jk}^{(d)}$ defined in Condition (f), and

$$\mathbf{D}_{ijk}(t; \boldsymbol{\beta}_k) = \int_0^t \{\mathbf{Z}_i - \bar{\mathbf{z}}_j(s; \boldsymbol{\beta}_k)\} d\Lambda_{ijk}^\#(s; \boldsymbol{\beta}_k),$$

$$\mathbf{U}_{ik}(\boldsymbol{\beta}_k) = \sum_{j=1}^J \sum_{i=1}^n \int_0^\tau \{\mathbf{Z}_i - \bar{\mathbf{z}}_j(t; \boldsymbol{\beta}_k)\} dM_{ijk}(t),$$

$$dM_{ijk}(s; \boldsymbol{\beta}_k) = dN_{ijk}(s) - Y_{ij}(s) d\Lambda_{ijk}^\#(s).$$

Note that the martingales, $M_{ijk}(t; \boldsymbol{\beta}_k)$ (for $k = 1, \dots, K$) are defined with respect to the joint filtration $\mathcal{F}_{ij}(t) = \sigma \{N_{ij}(s), Y_{ij}(s), \mathbf{Z}_i; s \in (0, t)\}$, where $N_{ij}(s) = I(X_i \leq s, A_i = j)$

The asymptotic variance can be estimated by:

$$\hat{V} \left[n^{\frac{1}{2}} \{\hat{\delta}_{jk}(t) - \delta_{jk}(t)\} \right] = \frac{1}{n} \sum_{\ell=1}^n \left[\sum_{d=1}^4 \left\{ \hat{\phi}_{ijk}^d(t; \boldsymbol{\beta}) - \hat{\phi}_{lk}^d(t; \boldsymbol{\beta}) \right\} \right]^2,$$

where the $\hat{\phi}_{ijk}^d(t)$ are obtained by replacing limiting values in $\phi_{ijk}^d(t)$ with their empirical counterparts.

As implied by the formulas in Theorem 2, calculation of the asymptotic variance is cumbersome computationally. As a result, and since the point estimator can be computed very quickly, we propose the bootstrap method, which we evaluate through simulation in the next section.

4 Simulation Studies

We conducted simulation studies to evaluate the finite sample performance of the proposed estimator $\hat{\delta}_{jk}(t)$. In the first simulation study, there are $K = 2$ causes, with cause $k = 1$ being of primary interest. There are $J = 5$ centers, with varying center sizes of $n_j = 100, 125$, and 150 individuals. There are three hierarchically dependent binary covariates. The covariate Z_{i1} follows a Bernoulli distribution with mean θ_{1j} ; $Z_{i2}|Z_{i1}$ is distributed as a Bernoulli with

$P(Z_{i2} = 1|Z_{i1}) = Z_{i1}\theta_{21} + (1 - Z_{i1})\theta_{22}$, and $Z_{i3}|Z_{i2}$ follows a Bernoulli distribution with $P(Z_{i3} = 1|Z_{i2}) = Z_{i2}\theta_{31} + (1 - Z_{i2})\theta_{32}$. The covariate patterns varied for each center due to center-specific θ_{1j} . In these set of simulations, $\theta_{1j}^T = [0.55, 0.75, 0.6, 0.65, 0.5]$, $\theta_{21} = 0.55$, $\theta_{22} = 0.45$, $\theta_{31} = 0.45$, and $\theta_{32} = 0.65$.

The event time for cause one T_{ij1} and the event time for cause two T_{ij2} follow exponential distributions, with

$$\lambda_{ijk}^{\#}(t) = \lambda_{0jk} \exp\{\beta_{k1}Z_{i1} + \beta_{k2}Z_{i2} + \beta_{k3}Z_{i3}\}$$

for $k = 1, 2$. The baseline hazards $\lambda_{0jk}(t)$ vary across centers and causes. For both configurations, the covariate effect for cause 1 is $\beta_1^T = [0.4, 0.5, 0.6]$ and for cause 2 is $\beta_2^T = [-0.1, 0.3, -0.2]$. The baseline cause-specific hazards are shown in Table 1.

The censoring distribution C_i is also exponential and also depends on the covariates, with

$$\lambda_i^C(t) = \lambda_0^C \exp\{\alpha_1 Z_{i1} + \alpha_2 Z_{i2} + \alpha_3 Z_{i3}\}.$$

The distribution of C_i differs for each center due to the different covariate patterns. There is administrative censoring at $t = 10$ to reflect the property that, in practice, the observation period is finite. The censoring baseline hazard is $\lambda_0^C = 0.02$ and $\alpha^T = [0.3, 0.5, 0.6]$. The observation time, X_i , is defined as the minimum of T_{i1} , T_{i2} , and C_i and whichever is the minimum will determine the value of X_i . Note that this set-up, although reminiscent of the latent failure time approach, generates data consistent with the assumed cause-specific hazard models. The average percent censored over the 500 iterations for Configuration 1 was 20%, 16%, 14%, 11% and 4% for centers $j = 1, \dots, 5$, respectively. For Configuration 2, the corresponding percentages were 20%, 4%, 13%, 4% and 8%.

We used the bootstrapping method to estimate the variability of the effect, with standard errors calculated using 25 bootstrap samples. The results based on 500 replicates of Configuration 1 and 2 are shown in Tables 2 and 3, respectively. We calculated the estimated effect at the times $t = 1$, $t = 3$, and $t = 5$ after the start of follow-up. In Configuration 1, the true effect varied greatly from having a large negative effect, -0.243 to a large positive effect, 0.369 . Given the inherent bounds of 0 and 1 on cumulative incidence, and thus bounds of -1 and 1 on the difference of cumulative incidence, a 30 percent displacement represents a rather large effect. Centers $j = 1$ and $j = 2$ are smaller centers that perform worse than expected and center $j = 4$ is a larger center that performs better than expected.

The proposed estimators $\hat{\delta}_{jk}(t)$ have small bias, even when the size of the center is small. As expected, the larger the center, in general the smaller the bias of $\hat{\delta}_{jk}(t)$. Within a center, the bias tends to be smaller at the earlier times than at later times. This could be due to the fact that there are more individuals at risk at the earlier times. The bootstrap standard errors

(BSE) were close to the empirical standard deviations (ESD). Coverage probabilities (CP) are mostly just under 95 percent.

In Configuration 2, we investigated the behavior of the proposed estimator in the presence of smaller center sizes. In Table 3, center $j = 1$ is a small center that has negative true effects of high magnitude, both factors that may hinder the estimator by reducing the number of type 1 events. Having only $n_j = 50$ individuals while the other four centers have at least twice as many individuals, center $j = 1$ not only has a small center in absolute terms, but also in relative terms. The bias of center 1 is still quite small at each of the three studied time points and its coverage probability is close to the nominal value. Thus, even for smaller centers that perform worse than expected, the proposed estimator performs quite well.

5 Application to SRTR data

We applied the proposed methods to compare, by OPO, the probability of receiving a kidney transplant among patients wait-listed for kidney transplantation. Data were obtained from the Scientific Registry of Transplant Recipients (SRTR). We selected patients from OPOs that make up Region 10 and who were wait-listed between January 1, 2000 and December 31, 2000. The observation period ended on December 31, 2009. The resulting sample size was $n = 1726$ across the $J = 6$ OPOs, with OPO sizes $n_1 = 197$, $n_2 = 625$, $n_3 = 368$, $n_4 = 182$, $n_5 = 259$ and $n_6 = 95$. Each patient's follow-up started when the patient was put on the wait-list. Follow-up ended at the earliest of receiving a transplant, death on wait-list, removal from wait-list, or loss to follow-up. Since we were interested in evaluating the OPOs based on their ability to ensure as many patients as possible receive the preferred treatment of receiving a deceased-donor transplant, transplantation was the cause of interest, with deaths and removals treated as competing risks. We focused on three time points, years 1, 3, and 5. This reflects current practice since survival statistics are usually reported at chosen year intervals rather than on a daily or a monthly basis.

A plot of the estimated effects $\hat{\delta}_{jk}(t)$ for $k = 1$ at the three time points is shown in Figure 1. We would have expected to see that as follow-up time increased, the spread of the estimated effects increased as well. This is because we believe that earlier differences in performance are less likely to be attributable to the OPO than are later differences. In the earlier period of follow-up, other factors independent of the OPO, such as the patient's inherent overall health, may affect his/her chance of surviving until a transplant becomes available. It appears that even one year after the start of follow-up, there is differentiation between the OPOs that perform better than expected and ones that perform worse than expected. The differences between the OPOs widen by year three, but decrease by year five. The 95% confidence intervals for $\hat{\delta}_{jk}(t)$ are shown in Table 4.

6 Discussion

We have proposed a summary measure that quantifies the center effect in terms of the cumulative incidence function. By averaging over transformed fitted values obtained by Cox models, we compare the patient experience under two scenarios, one actual and one hypothetical. The proposed method would allow one to determine which groups of patients

are at greater or lesser probability of experiencing the event of interest. In the context of evaluating centers, a center could compare its actual performance to that if it were performing at the overall average. Simulation studies imply that the proposed estimators have negligible bias. Although calculating the asymptotic standard error is cumbersome, the bootstrap standard error appeared to be fairly accurate, as evidenced by comparison to the empirical standard errors and coverage probabilities. The bootstrap is an attractive alternative in our setting, since the point estimators can be obtained quite quickly.

The proposed methods were applied to national transplant registry data to evaluate OPOs with respect to average probability of receiving a kidney transplant. From the perspective of each individual OPO, an estimate of the cumulative incidence of transplantation is a meaningful metric of quality of delivered care. It answers the question, how would these same patients have done elsewhere, on average? It gives an OPO valuable information since the evaluation is done based on that particular OPO's case mix. Applied to the country, this method would allow us to see which parts of the country are under-served or well-served, taking into account that the profiles of wait-listed patients of each OPO can be different. Each effect estimate is specific to an OPO's demographics, so the effect estimates from different OPOs generally cannot be compared to each other meaningfully.

The OPO is responsible for allocating organs to patients on a wait-list for organ transplants. It would be considered optimal if the greatest number of patients eventually receive an organ before an event can occur that prevents a transplant. Thus, the performance of an OPO is crucial for the patients whose health depends on these time sensitive transplants.

Zhang and Zhang (2011) developed methods to estimate the cumulative incidence function for different treatment groups by adjusting for the patient population in each group. Zhang and Zhang (2011) used the proportional sub-distribution hazards model (Fine and Gray, 1999) to regress the cumulative incidence function on the covariates. There are two main differences between Zhang and Zhang (2011) and the methods we propose. First, the method of Zhang and Zhang is more convenient, in the sense that only one regression model need be fitted. However, such convenience is a trade-off with the potential for model misspecification (primarily in the form of non-proportionality). Second, Zhang and Zhang (2011) employs direct standardization, which is advantageous since it permits comparisons between center effects.

The proposed methods compare the cumulative incidence function for each center to the CIF from a hypothetical national average. As described in (5), the national average is weighted by center size. This is an arbitrary albeit reasonable choices; other reasonable choices exist. Among alternatives, a popular choice would likely be to assign equal weight to each center in computing the comparator group. An advantage of center-size weighting is interpretability. A center with a larger (smaller) number of patients contributes more (less) to the national average, which makes sense since the national average is influenced more by larger centers. If one selected patients at random from the study population (without regard to center), the average CIF would look more like the center size-weighted average than an equal-weighted average. An advantage of equal weighting may be power to detect differences in CIFs for larger centers. That is, larger centers may be, by construction, too

similar to their comparator center. In practice, the choice of the comparator would often be dictated by the research objectives of the analysis.

We constructed a measure that uses the CIF to quantify the center effect. An alternative to modeling the CIF is to base the effect measure on the cumulative cause-specific hazard. The CIF incorporates information from all causes to give a natural interpretation in the competing risks setting. In our motivating example, we want to take into account the entire patient experience, and not just his/her transplant experience, to determine the probability of receiving a transplant. For example, patients in a particular OPO may be getting transplanted at a faster rate while alive, but are also dying on the wait-list at a faster rate. We would want to use both pieces of information in the comparison. However, if interest lies in estimating the event rate of one specific cause without the input of the other causes which are not of direct interest, then the CSH is more apt. For example, the cumulative CSH would be more appropriate for comparing the rate of transplants among surviving patients.

Acknowledgements

This research was supported in part by National Institutes of Health grant 5R01-DK070869 (DES). The authors thank the Associate Editor and Reviewers for constructive suggestions which improved the manuscript. The authors also thank the Scientific Registry of Transplant Recipients (SRTR) for access to the organ transplant database. The SRTR is funded by a contract from the Health Resources and Services Administration (HRSA), U.S. Department of Health and Human Services. The views expressed in this report do not represent those of the U.S. Government.

APPENDIX

Proof of Theorem 2

The proof revolves around asymptotic expansions of the following quantities.

1. $n^{\frac{1}{2}}(\hat{\beta}_k - \beta_k)$
2. $n^{\frac{1}{2}} \left\{ \hat{\Lambda}_{0jk}^{\#}(t) - \Lambda_{0jk}^{\#}(t) \right\}$
3. $n^{\frac{1}{2}} \left\{ \hat{\Lambda}_{ijk}^{\#}(t) - \Lambda_{ijk}^{\#}(t) \right\}$
4. $n^{\frac{1}{2}} \left\{ \hat{S}_{ij}(t) - S_{ij}(t) \right\}$
5. $n^{\frac{1}{2}} \left\{ \hat{F}_{ijk}(t) - F_{ijk}(t) \right\}$
6. $n^{\frac{1}{2}} \left\{ \hat{\delta}_{jk}(t) - \delta_{jk}(t) \right\}$

[1.] $n^{\frac{1}{2}}(\hat{\beta}_k - \beta_k)$

By a Taylor expansion of $U_k(\beta)$ around β_k ,

$$n^{\frac{1}{2}}(\hat{\beta}_k - \beta_k) = I_k^{-1}(\beta_k) n^{-\frac{1}{2}} \sum_{i=1}^n U_{ik}(\beta_k) + o_p(1),$$

where $\mathbf{U}_{ik}(\boldsymbol{\beta})$ and $\mathbf{I}_k(\boldsymbol{\beta})$ are as defined in the theorem. The result then follows from standard Martingale theory; e.g., Andersen and Gill (1982); Fleming and Harrington (1991). Note that the processes $M_{ijk}(t; \boldsymbol{\beta}_1)$, for $k = 1, \dots, K$ are martingales with respect to the filtration

$$\mathcal{F}_{ij}(t) = \sigma\{Y_{ij}(s), \mathbf{Z}_i; s \in (0, t]\}.$$

$$[2.] n^{\frac{1}{2}} \left\{ \hat{\Lambda}_{0jk}^{\#}(t) - \Lambda_{0jk}^{\#}(t) \right\}$$

We decompose the quantity as follows,

$$n^{\frac{1}{2}} \left\{ \hat{\Lambda}_{0jk}^{\#}(t) - \Lambda_{0jk}^{\#}(t) \right\} = n^{\frac{1}{2}} \left\{ \hat{\Lambda}_{0jk}^{\#}(t; \hat{\boldsymbol{\beta}}_k) - \hat{\Lambda}_{0jk}^{\#}(t; \boldsymbol{\beta}_k) \right\} \quad (6)$$

$$+ n^{\frac{1}{2}} \left\{ \hat{\Lambda}_{0jk}^{\#}(t; \boldsymbol{\beta}_k) - \Lambda_{0jk}^{\#}(t; \boldsymbol{\beta}_k) \right\}. \quad (7)$$

Since $\hat{\Lambda}_{0jk}^{\#}(t)$ is the Breslow-Aalen analog of $\hat{\Lambda}_{0jk}^{\#}(t)$, we adapt results derived for the Breslow-Aalen estimator (Fleming and Harrington, 1991). From this perspective, we can write

$$(6) = \mathbf{h}_{jk}^T(t; \boldsymbol{\beta}_k) \mathbf{I}_k(\boldsymbol{\beta}_k)^{-1} n^{\frac{1}{2}} \sum_{i=1}^n \mathbf{U}_{ik}(\boldsymbol{\beta}_k) + o_p(1),$$

where the last equality holds by Slutsky's Theorem. With respect to (7), we can write

$$(7) = n^{-\frac{1}{2}} \sum_{i=1}^n \int_0^t r_{jk}^{(0)}(s; \boldsymbol{\beta}_k)^{-1} dM_{ijk}(s; \boldsymbol{\beta}_k) + o_p(1),$$

since $R_{jk}^{(0)}(t) \rightarrow r_{jk}^{(0)}(t)$ in probability. Putting (6) and (7) together, we get:

$$[2] = n^{-\frac{1}{2}} \sum_{i=1}^n \Phi_{ijk}(t; \boldsymbol{\beta}_k) + o_p(1),$$

where $\Phi_{ijk}(t; \boldsymbol{\beta}) = -\mathbf{h}_{jk}^T(t; \boldsymbol{\beta}) \mathbf{I}_k(\boldsymbol{\beta})^{-1} \mathbf{U}_{ik}(\boldsymbol{\beta}) + \int_0^t r_{jk}^{(0)}(s; \boldsymbol{\beta})^{-1} dM_{ijk}(s; \boldsymbol{\beta})$.

$$[3.] n^{\frac{1}{2}} \left\{ \hat{\Lambda}_{ijk}^{\#}(t) - \Lambda_{ijk}^{\#}(t) \right\}$$

We start with the following decomposition,

$$[3] = n^{\frac{1}{2}} \left\{ \int_0^t Y_{ij}(s) \exp(\hat{\beta}_k^T \mathbf{Z}_i) d\hat{\Lambda}_{0jk}^\#(s; \hat{\beta}_k) - \int_0^t Y_{ij}(s) \exp(\beta_k^T \mathbf{Z}_i) d\hat{\Lambda}_{0jk}^\#(s; \hat{\beta}_k) \right\} \quad (8)$$

$$+ n^{\frac{1}{2}} \left\{ \int_0^t Y_{ij}(s) \exp(\beta_k^T \mathbf{Z}_i) d\hat{\Lambda}_{0jk}^\#(s; \hat{\beta}_k) - \int_0^t Y_{ij}(s) \exp(\beta_k^T \mathbf{Z}_i) d\Lambda_{0jk}^\#(s; \beta_k) \right\}. \quad (9)$$

By a Taylor expansion, then applying Result [1], we can write

$$(8) = \int_0^t \mathbf{Z}_i^T Y_{ij}(s) d\Lambda_{ijk}^\#(s; \beta_k) \mathbf{I}_k(\beta_k)^{-1} n^{-\frac{1}{2}} \sum_{\ell=1}^n \mathbf{U}_{\ell k}(\beta_k) + o_p(1),$$

where the last equality holds by the convergence in probability of $\hat{\Lambda}_{ijk}^\#(t)$ to $\Lambda_{ijk}^\#(t)$. We re-express (9) as

$$(9) = \int_0^t Y_{ij}(s) \exp(\beta_k^T \mathbf{Z}_i) \left\{ n^{-\frac{1}{2}} \sum_{\ell=1}^n d\Phi_{\ell jk}(t; \beta_k) \right\} + o_p(1),$$

by incorporating Result [2], where we define

$$d\Phi_{ijk}(t; \beta_k) = [-\bar{z}_j(s; \beta_k) d\Lambda_{0jk}(s; \beta_k)]^T \mathbf{I}_k(\beta_k)^{-1} \mathbf{U}_{ik}(\beta_k) + r_{jk}^{(0)}(s; \beta_k)^{-1} dM_{ijk}(s; \beta_k).$$

Combining (8) and (9), then further reorganizing, we obtain

$$[3] = \mathbf{D}_{ijk}^T(t; \beta_k) \mathbf{I}_k(\beta_k)^{-1} n^{-\frac{1}{2}} \sum_{\ell=1}^n \mathbf{U}_{\ell k}(\beta_k) + n^{-\frac{1}{2}} \sum_{\ell=1}^n J_{i\ell jk}(t; \beta_k),$$

where we let

$$\mathbf{D}_{ijk}(t; \beta) = \int_0^t \{ \mathbf{Z}_i - \bar{z}_j(s; \beta) \} d\Lambda_{ijk}^\#(s; \beta)$$

$$J_{i\ell jk}(t; \beta) = \int_0^t Y_{ij}(s) \cdot \exp(\beta_k^T \mathbf{Z}_i) r_{jk}^{(0)}(s; \beta)^{-1} dM_{\ell jk}(s; \beta).$$

$$[4.] n^{\frac{1}{2}} \left\{ \hat{S}_{ij}(t) - S_{ij}(t) \right\}$$

We decompose [4] as follows,

$$[4] = - \sum_{m=1}^K S_{ij}(t) n^{-1/2} \left\{ \hat{\Lambda}_{ijm}^{\#}(t) - \Lambda_{ijm}^{\#}(t) \right\},$$

due to the Functional Delta Method, combined with the convergence,

$e^{-\hat{\Lambda}_{ijm}^{\#}(s; \hat{\beta}_m)} \rightarrow e^{-\Lambda_{ijm}^{\#}(s; \beta_m)}$, by continuity. Using Result [3], we then obtain

$$[4] = -S_{ij}(t) \sum_{m=1}^K \left\{ \mathbf{D}_{ijm}^T(t; \beta_m) \mathbf{I}_m(\beta_m)^{-1} n^{-\frac{1}{2}} \sum_{\ell=1}^n \mathbf{U}_{\ell m}(\beta_m) + n^{-\frac{1}{2}} \sum_{\ell=1}^n J_{i\ell jm}(t; \beta_m) \right\}.$$

$$[5.] n^{\frac{1}{2}} \left\{ \hat{F}_{ijk}(t) - F_{ijk}(t) \right\}$$

We use the following decomposition,

$$[5] = n^{\frac{1}{2}} \left\{ \int_0^t \hat{S}_{ij}(s; \hat{\beta}) d\hat{\Lambda}_{ijk}^{\#}(s; \hat{\beta}_k) - \int_0^t S_{ij}(s; \beta) d\hat{\Lambda}_{ijk}^{\#}(s; \hat{\beta}_k) \right\} \quad (10)$$

$$+ n^{\frac{1}{2}} \left\{ \int_0^t S_{ij}(s; \beta) d\hat{\Lambda}_{ijk}^{\#}(s; \hat{\beta}_k) - \int_0^t S_{ij}(s; \beta) d\Lambda_{ijk}^{\#}(s; \beta_k) \right\}, \quad (11)$$

where $\beta^T = [\beta_1^T, \dots, \beta_K^T]$. Note that equation (10) will eventually give rise to

$\phi_{i\ell jk}^1(t, \beta)$, $\phi_{i\ell jk}^2(t, \beta)$ as defined in Theorem 2, while (11) will give rise to $\phi_{i\ell jk}^3(t, \beta)$ and $\phi_{i\ell jk}^4(t, \beta)$. We can write (10) as

$$(10) = \int_0^t n^{\frac{1}{2}} \left[\hat{S}_{ij}(s; \hat{\beta}) - S_{ij}(s; \beta) \right] d\hat{\Lambda}_{ijk}^{\#}(s; \hat{\beta}_k) = - \sum_{m=1}^K \int_0^t S_{ij}(s) \mathbf{D}_{ijm}^T(s; \beta_m) d\hat{\Lambda}_{ijk}^{\#}(s; \hat{\beta}_k) \times \mathbf{I}_m(\beta_m)^{-1} n^{-\frac{1}{2}} \sum_{\ell=1}^n \mathbf{U}_{\ell m}(\beta_m) \quad (12)$$

$$- n^{-\frac{1}{2}} \sum_{m=1}^K \int_0^t S_{ij}(s) \sum_{\ell=1}^n J_{i\ell jm}(s; \beta_m) d\hat{\Lambda}_{ijk}^{\#}(s; \hat{\beta}_k), \quad (13)$$

where we have used Result [4]. Focusing on (12), we have,

$$(12) = - \sum_{m=1}^K \int_0^t \mathbf{D}_{ijm}^T(s; \beta_m) dF_{ijk}(s) \mathbf{I}_m(\beta_m)^{-1} n^{-\frac{1}{2}} \sum_{\ell=1}^n \mathbf{U}_{\ell m}(\beta_m) + o_p(1),$$

by the fact that $\hat{\Lambda}_{ijk}^{\#}(s; \beta_k) \rightarrow \Lambda_{ijk}^{\#}(s; \beta_k)$, and $\hat{\beta}_k \rightarrow \beta_k$ and so, by the CMT,

$\hat{\Lambda}_{ijk}^{\#}(s; \hat{\beta}_k) \rightarrow \Lambda_{ijk}^{\#}(s; \beta_k)$. Therefore, from (12), define

$$\phi_{i\ell jk}^1(t, \beta) = \sum_{m=1}^K \left\{ \int_0^t \{ \mathbf{Z}_i - \bar{\mathbf{z}}_j(u; \beta_k) \}^T F_{ijk}(u) - F_{ijk}(t) \int_0^t \{ \mathbf{Z}_i - \bar{\mathbf{z}}_j(u; \beta_k) \}^T d\Lambda_{ijk}^{\#}(u; \beta_k) \right\} \mathbf{I}_m(\beta_m)^{-1} \mathbf{U}_{\ell m}(\beta_m).$$

Using analogous arguments, we can write

$$(13) = -n^{-\frac{1}{2}} \sum_{m=1}^K \sum_{\ell=1}^n \left\{ F_{ijk}(t) \int_0^t Y_{ij}(u) \exp(\beta_m^T \mathbf{Z}_i) r_{jk}^{(0)}(u; \beta_m)^{-1} dM_{\ell jm}(u; \beta_m) \right. \\ \left. - \int_0^t F_{ijk}(u) Y_{ij}(u) \exp(\beta_m^T \mathbf{Z}_i) r_{jk}^{(0)}(u; \beta_m)^{-1} dM_{\ell jm}(u; \beta_m) \right\},$$

due to previously described properties of $\hat{\Lambda}_{ijk}^{\#}(s; \beta_k)$ and $\hat{\beta}_k$. Thus, from (13) define:

$$\phi_{ijk}^2(t, \beta_1) = - \sum_{m=1}^K \int_0^t \{F_{ijk}(t) - F_{ijk}(u)\} Y_{ij}(u) \exp(\beta_m^T \mathbf{Z}_i) r_{jk}^{(0)}(u; \beta_m)^{-1} dM_{\ell jm}(u; \beta_m).$$

Now shifting attention to (11), we obtain

$$(11) = \left\{ \int_0^t S_{ij}(s; \beta) d\mathbf{D}_{ijk}^T(s; \beta_k) \right\} \mathbf{I}_k(\beta_k)^{-1} n^{-\frac{1}{2}} \sum_{\ell=1}^n \mathbf{U}_{\ell k}(\beta_k) \quad (14)$$

$$+ n^{-\frac{1}{2}} \int_0^t S_{ij}(s; \beta) \sum_{\ell=1}^n dJ_{i\ell jk}(s; \beta_k), \quad (15)$$

through Result [3]. We can then write

$$(14) = \left\{ \int_0^t S_{ij}(s; \beta) [\mathbf{Z}_i - \bar{\mathbf{z}}_j(s; \beta_k)]^T d\Lambda_{ijk}^{\#}(s; \beta_k) \right\} \mathbf{I}_k(\beta_k)^{-1} n^{-\frac{1}{2}} \sum_{\ell=1}^n \mathbf{U}_{\ell k}(\beta_k)$$

and, correspondingly, we define

$$\phi_{ijk}^3(t, \beta) = \left\{ \int_0^t S_{ij}(s; \beta) [\mathbf{Z}_i - \bar{\mathbf{z}}_j(s; \beta_k)]^T d\Lambda_{ijk}^{\#}(s; \beta_k) \right\} \mathbf{I}_k(\beta_k)^{-1} \mathbf{U}_{\ell k}(\beta_k).$$

We can re-express (15) as follows,

$$(15) = n^{-\frac{1}{2}} \sum_{\ell=1}^n \int_0^t S_{ij}(s; \beta) Y_{ij}(s) \exp(\beta_k^T \mathbf{Z}_i) r_{jk}^{(0)}(s; \beta_k)^{-1} dM_{\ell jk}(s; \beta_k),$$

and, correspondingly redefine

$$\phi_{ijk}^4(t, \beta) = \int_0^t S_{ij}(s; \beta_1) Y_{ij}(s) \exp(\beta_k^T \mathbf{Z}_i) r_{jk}^{(0)}(s; \beta_k)^{-1} dM_{\ell jk}(s; \beta_k).$$

Combining results derived in this subsection, we obtain,

$$n^{\frac{1}{2}} \left\{ \hat{F}_{ijk}(t) - F_{ijk}(t) \right\} = n^{-\frac{1}{2}} \sum_{\ell=1}^n \left\{ \sum_{d=1}^4 \phi_{i\ell jk}^d(t, \boldsymbol{\beta}) \right\} + o_p(1),$$

where the $\sum_{d=1}^4 \phi_{i\ell jk}^d(t, \boldsymbol{\beta})$ are asymptotically independent and identically distributed variates with mean 0.

$$[6.] n^{\frac{1}{2}} \left\{ \hat{\delta}_{jk}(t) - \delta_{jk}(t) \right\}$$

We now complete the proof by averaging over $i = 1, \dots, n$ to obtain the limiting distribution of the proposed estimator. Setting $n_j = \sum_{i=1}^n A_{ij}$ and $p_j = E[A_{ij}]$, we have

$$\begin{aligned} [6] &= \frac{1}{n_j} \sum_{i=1}^n A_{ij} \left\{ n^{-\frac{1}{2}} \sum_{\ell=1}^n \left[\sum_{d=1}^4 \phi_{i\ell jk}^d(t, \boldsymbol{\beta}) - \frac{1}{n} \sum_{r=1}^J \left\{ \sum_{d=1}^4 \phi_{i\ell rk}^d(t, \boldsymbol{\beta}) \right\} n_r \right] \right\} \\ &= n^{-\frac{1}{2}} \sum_{\ell=1}^n \left\{ \sum_{d=1}^4 \left[\frac{1}{n_j} \sum_{i=1}^n A_{ij} \phi_{i\ell jk}^d(t, \boldsymbol{\beta}) \right. \right. \\ &\quad \left. \left. - \frac{1}{n_j} \sum_{i=1}^n A_{ij} \frac{1}{n} \sum_{r=1}^J \phi_{i\ell rk}^d(t, \boldsymbol{\beta}) n_r \right] \right\}. \end{aligned} \quad (16)$$

Focusing on each component in (16), we have the following for the expression involving

$$\phi_{i\ell jk}^1(t; \boldsymbol{\beta}),$$

$$\frac{1}{n_j} \sum_{i=1}^n A_{ij} \phi_{i\ell jk}^1(t; \boldsymbol{\beta}) = -\frac{1}{p_j} E \left[\sum_{m=1}^K A_{ij} \int_0^t \mathbf{D}_{ijm}^T(s; \boldsymbol{\beta}_m) dF_{ijk}(s; \boldsymbol{\beta}_k) \right] \mathbf{I}_m(\boldsymbol{\beta}_m)^{-1} \mathbf{U}_{\ell m}(\boldsymbol{\beta}_m) = \phi_{\ell jk}^1(t; \boldsymbol{\beta}),$$

since $n_j/n \rightarrow p_j$ by the Weak Law of Large Numbers (WLLN), continuity, and Slutsky's Theorem. The simplification for the term involving $\phi_{i\ell jk}^3(t, \boldsymbol{\beta})$ unfold in a similar way. The term involving $\phi_{i\ell jk}^2(t; \boldsymbol{\beta})$ can be written as the following:

$$\begin{aligned} &\frac{1}{n_j} \sum_{i=1}^n A_{ij} \phi_{i\ell jk}^2(t; \boldsymbol{\beta}) \\ &= - \sum_{m=1}^K \int_0^t \left[\frac{1}{p_j} \frac{1}{n} \sum_{i=1}^n A_{ij} \{ F_{ijk}(t) - F_{ijk}(u) \} Y_{ij}(u) \exp(\boldsymbol{\beta}_m^T \mathbf{Z}_i) r_{jk}^{(0)}(u; \boldsymbol{\beta}_m)^{-1} \right] dM_{\ell jm}(u; \boldsymbol{\beta}_m) \\ &= \phi_{\ell jk}^2(t; \boldsymbol{\beta}), \end{aligned}$$

by the WLLN, continuity, and Slutsky's Theorem. The term involving $\phi_{i\ell jk}^4(t; \boldsymbol{\beta})$ unfold in a similar way. The term involving $\phi_{i\ell rk}^1(t; \boldsymbol{\beta})$ can be written as the following:

$$\begin{aligned}
& \frac{1}{n_j} \sum_{i=1}^n \sum_{m=1}^K A_{ij} \left\{ \frac{1}{n} \sum_{r=1}^J \phi_{i\ell rk}^1(t; \boldsymbol{\beta}) n_r \right\} \\
&= -\frac{1}{p_j} \sum_{r=1}^J p_r E \left[\sum_{m=1}^K A_{ij} \int_0^t \mathbf{D}_{irm}^T(s; \boldsymbol{\beta}_m) dF_{irk}(s; \boldsymbol{\beta}_k) \right] \mathbf{I}_m(\boldsymbol{\beta}_m)^{-1} \mathbf{U}_{\ell m}(\boldsymbol{\beta}_m) \\
&= \phi_{\ell k}^1(t; \boldsymbol{\beta}).
\end{aligned}$$

The term involving $\phi_{i\ell rk}^3(t; \boldsymbol{\beta})$ can be expressed in a similar way. The term involving $\phi_{i\ell rk}^2(t; \boldsymbol{\beta})$ can be written as,

$$\begin{aligned}
& \frac{1}{n_j} \sum_{i=1}^n \sum_{m=1}^K A_{ij} \left\{ \frac{1}{n} \sum_{r=1}^J \phi_{i\ell rk}^3(t; \boldsymbol{\beta}) \cdot n_r \right\} \\
&= \sum_{m=1}^K \int_0^t \sum_{r=1}^J \frac{1}{p_j} p_r E \left[A_{ij} \{F_{irk}(t) - F_{irk}(u)\} Y_{ir}(u) \cdot \exp(\boldsymbol{\beta}_m^T \mathbf{Z}_i) r_{rk}^{(0)}(u; \boldsymbol{\beta}_m)^{-1} \right] dM_{\ell rm}(u; \boldsymbol{\beta}_m) \\
&= \phi_{\ell k}^2(t; \boldsymbol{\beta}).
\end{aligned}$$

The term involving $\phi_{i\ell rk}^4(t; \boldsymbol{\beta})$ can be expressed analogously. Therefore, we can write

$$[6] = n^{-\frac{1}{2}} \sum_{\ell=1}^n \left\{ \sum_{d=1}^4 \left(\phi_{\ell jk}^d(t; \boldsymbol{\beta}) + \phi_{\ell k}^d(t; \boldsymbol{\beta}) \right) \right\}. \quad (17)$$

All summands across l have mean 0 since the ϕ 's have mean 0. If we apply the Functional Central Limit Theorem to [6], where each component is independent across l , we have

$$V \left(n^{\frac{1}{2}} \left\{ \hat{\delta}_{jk}(t) - \delta_{jk}(t) \right\} \right) = E \left[\left\{ \sum_{d=1}^4 \left(\phi_{\ell jk}^d(t; \boldsymbol{\beta}) - \phi_{\ell k}^d(t; \boldsymbol{\beta}) \right) \right\}^2 \right]. \quad (18)$$

References

- Andersen PK, Gill RD. Cox's regression model for counting processes: A large sample study. The Annals of Statistics. 1982; 10:1100–1120.
- Prentice RL, Kalbfleisch JD, Peterson AV, Flournoy V, Farewell VT, Breslow NE. The analysis of failure times in the presence of competing risks. Biometrics. 1978; 34:541–554. [PubMed: 373811]
- Cox D. Regression models and life-tables (with discussion). Journal of the Royal Statistical Society Series B. 1972; 34:187–220.
- Cox D. Partial likelihood. Biometrika. 1975; 62:262–276.
- Fleming, TR.; Harrington, DP. Counting Processes and Survival Analysis. New York: Wiley; 1991.
- Kalbfleisch, JD.; Prentice, RL. The Statistical Analysis of Failure Time Data. Hoboken, New Jersey: Wiley; 2002.
- Cox D. Regression models and life-tables (with discussion). Journal of the Royal Statistical Society Series B. 1972; 34:187–220.
- Cox D. Partial likelihood. Biometrika. 1975; 62:262–276.
- Fleming, TR.; Harrington, DP. Counting Processes and Survival Analysis. New York: Wiley; 1991.

10. Dabrowska DM, Doksum KA. Estimates and testing in a two-sample generalized odds-rate model. *Journal of the American Statistical Association*. 1988; 83:744–749.
11. Benichou J, Gail MH. Estimates of absolute cause-specific risk in cohort studies. *Biometrics*. 1990; 46:813–826. [PubMed: 2242416]
12. Cheng SC, Fine JP, Wei LJ. Prediction of cumulative incidence function under the proportional hazards model. *Biometrics*. 1998; 54:219–228. [PubMed: 9544517]
13. Fine JP, Gray RJ. A proportional hazards model for the subdistribution of a competing risk. *Journal of the American Statistical Association*. 1999; 94:496–509.
14. Gray RJ. A class of K-sample tests for comparing the cumulative incidence of a competing risk. *The Annals of Statistics*. 1988; 16:1141–1154.
15. Zhang MJ, Fine JP. Summarizing differences in cumulative incidence functions. *Statistics in Medicine*. 2008; 27:4939–4949. [PubMed: 18563792]
16. Gail, MH. *Competing Risks*. London: 2001.
17. Gail MH. A review and critique of some models used in competing risk analysis. *Biometrics*. 1975; 31:209–222. [PubMed: 1164533]
18. Crowder, MJ. *Classical Competing Risks*. London: Chapman and Hall/CRC Press; 2001.
19. Cox DR. The analysis of exponentially distributed lifetimes with two types of failure. *Journal of the Royal Statistical Society Series B*. 1959; 21:411–421.
20. Moeschberger ML, David HA. Life tests under competing causes of failure and the theory of competing risks. *Biometrics*. 1971; 27:909–923.
21. Tsiatis AA. A nonidentifiability aspect of the problem of competing risks. *Proceedings of the National Academy of Sciences*. 1975; 72:20–22.
22. Chiang, CL. *Introduction to Stochastic Processes in Biostatistics*. New York: Wiley; 1968.
23. Anderson PK, Borgan A, Gill RD, Keiding N. Statistical Models Based on Counting Processes. *Biometrics*. 1993; 24:100–101.
24. Klein, JP.; Moeschberger, ML. *Survival Analysis*. New York: Springer; 2003.
25. Breslow N. Contribution to the discussion on the paper by D. R. Cox, regression and life tables. *Journal of the Royal Statistical Society Series B*. 1972; 34:216–217.
26. Logan BR, Nelson GO, Klein JP. Analyzing center specific outcomes in hematopoietic cell transplantation. *Lifetime Data Analysis*. 2008; 14(4):389–404. [PubMed: 18836830]
27. Kaplan EL, Meier P. Nonparametric estimation from incomplete observations. *Journal of the American Statistical Association*. 1958; 53(282):457–481.
28. DeLong ER, Peterson ED, DeLong DM, Muhlbaier LH, Hackett S, Mark DB. Comparing risk-adjustment methods for provider profiling. *Statistics in Medicine*. 1997; 16:2645–2664. [PubMed: 9421867]
29. Zhang X, Zhang MJ. SAS macros for estimation of direct adjusted cumulative incidence curves under proportional subdistribution hazards models. *Computer Methods and Programs in Biomedicine*. 2011; 101:87–93. [PubMed: 20724020]

OPO Effect on Cumulative Incidence of Transplant

Fig. 1. Analysis of SRTR data: $\hat{\delta}_{j1}(t)$ for 1, 3, and 5 years post wait-listing, for $j = 1, \dots, 6$.

Table 1

Baseline cause-specific hazards for the simulation study

	$j = 1$	$j = 2$	$j = 3$	$j = 4$	$j = 5$
Configuration 1					
λ_{0j1}	0.1	0.15	0.2	0.22	0.7
λ_{0j2}	0.12	0.1	0.08	0.09	0.08
Configuration 2					
λ_{0j1}	0.1	0.8	0.2	0.76	0.4
λ_{0j2}	0.12	0.1	0.08	0.09	0.08

Table 2

Performance of $\hat{\delta}_{j1}(t)$ based on 500 simulations of Configuration 1, with bias, empirical standard deviation (ESD), the bootstrap standard error (BSE), and the 95% confidence interval coverage probabilities (CP).

j	n	t	$\hat{\delta}_{j1}(t)$	BIAS	ESD	BSE	CP
1	100	1	-0.205	0.015	0.037	0.037	0.93
		3	-0.243	0.016	0.041	0.044	0.94
		5	-0.221	0.018	0.040	0.045	0.94
2	100	1	-0.101	-0.009	0.043	0.043	0.92
		3	-0.078	-0.015	0.049	0.050	0.93
		5	-0.056	-0.022	0.047	0.052	0.93
3	125	1	-0.040	0.007	0.036	0.037	0.95
		3	0.004	0.008	0.035	0.036	0.93
		5	0.021	0.009	0.031	0.033	0.94
4	150	1	-0.009	-0.000	0.032	0.033	0.95
		3	0.032	0.003	0.031	0.031	0.93
		5	0.038	0.004	0.028	0.029	0.95
5	100	1	0.369	-0.011	0.035	0.038	0.96
		3	0.268	-0.019	0.027	0.029	0.92
		5	0.194	-0.019	0.026	0.028	0.92

Table 3

Performance of $\hat{\delta}_{j1}(t)$ based on 500 simulations of Configuration 2, with bias, empirical standard deviation (ESD), the bootstrap standard error (BSE), and the 95% confidence interval coverage probabilities (CP).

j	n	t	$\hat{\delta}_{j1}(t)$	BIAS	ESD	BSE	CP
1	50	1	-0.379	0.014	0.051	0.053	0.95
		3	-0.379	0.008	0.063	0.064	0.94
		5	-0.323	0.003	0.059	0.066	0.97
2	100	1	0.231	-0.016	0.034	0.037	0.94
		3	0.137	-0.024	0.023	0.028	0.92
		5	0.093	-0.021	0.022	0.027	0.93
3	125	1	-0.215	0.012	0.036	0.035	0.93
		3	-0.133	0.014	0.032	0.034	0.92
		5	-0.080	0.013	0.027	0.031	0.94
4	100	1	0.207	-0.002	0.036	0.037	0.95
		3	0.134	-0.008	0.025	0.027	0.96
		5	0.093	-0.007	0.023	0.025	0.96
5	150	1	0.014	-0.001	0.032	0.031	0.94
		3	0.057	0.001	0.023	0.024	0.96
		5	0.050	0.001	0.021	0.022	0.95

Table 4

Analysis of SRTR data: $\hat{\delta}_{j1}(t)$ with 95% confidence limits for 1, 3, and 5 years post wait-listing, for $j = 1, \dots, 6$.

year	OPO	$\hat{\delta}_{j1}(t)$	upper 95% confidence limit	lower 95% confidence limit
1	1	0.12	0.09	0.15
3	1	0.16	0.12	0.20
5	1	0.06	0.02	0.11
1	2	0.12	0.09	0.14
3	2	0.21	0.18	0.24
5	2	0.09	0.05	0.12
1	3	0.05	0.03	0.07
3	3	0.15	0.12	0.17
5	3	0.07	0.04	0.09
1	4	0.13	0.11	0.15
3	4	0.15	0.13	0.18
5	4	0.13	0.10	0.15
1	5	-0.07	-0.08	-0.06
3	5	-0.14	-0.16	-0.12
5	5	-0.10	-0.12	-0.07
1	6	-0.06	-0.07	-0.06
3	6	-0.08	-0.10	-0.07
5	6	-0.04	-0.06	-0.03