Published in final edited form as:

Biometrics. 2016 June ; 72(2): 525–534. doi:10.1111/biom.12427.

Semiparametric Methods to Contrast Gap Time Survival Functions: Application to Repeat Kidney Transplantation

Xu Shu^{1,*} and Douglas E. Schaubel^{1,**}

¹Department of Biostatistics, University of Michigan, Ann Arbor, Michigan, 48109-2029, U.S.A

Summary

Times between successive events (i.e., gap times) are of great importance in survival analysis. Although many methods exist for estimating covariate effects on gap times, very few existing methods allow for comparisons between gap times themselves. Motivated by the comparison of primary and repeat transplantation, our interest is specifically in contrasting the gap time survival functions and their integration (restricted mean gap time). Two major challenges in gap time analysis are non-identifiability of the marginal distributions and the existence of dependent censoring (for all but the first gap time). We use Cox regression to estimate the (conditional) survival distributions of each gap time (given the previous gap times). Combining fitted survival functions based on those models, along with multiple imputation applied to censored gap times, we then contrast the first and second gap times with respect to average survival and restricted mean lifetime. Large-sample properties are derived, with simulation studies carried out to evaluate finite-sample performance. We apply the proposed methods to kidney transplant data obtained from a national organ transplant registry. Mean 10-year graft survival of the primary transplant is significantly greater than that of the repeat transplant, by 3.9 months ($p = 0.023$), a result that may lack clinical importance.

Keywords

Gap time; conditional model; Cox regression; multiple imputation; restricted mean lifetime

1. Introduction

In epidemiologic studies, a sequence of serial events is often of interest. Examples include numbered hospitalizations, tumor recurrences and, in a more general sense, transitions between states visited in a fixed order. There are two ways to define the time scale for multiple event data. The first, *total time*, measures time from a fixed time origin to an event. The second, *gap time*, measures time between successive events. However, for serial events, gap times are often of more direct interest than total times. For instance, a patient discharged from the hospital may be concerned about time until readmission.

* shuxu@umich.edu

** deschau@umich.edu

Supplementary Materials: Web Appendix A, referenced in Section 3, is available with this paper at the *Biometrics* website on Wiley Online Library.

The analysis of gap times has several methodologic issues when pursuing marginal inference. Generally, the within-subject gap times are not independent. Even if total times are censored independently, the second and subsequent gap times will be subject to induced dependent censoring (Lin et al., 1999; Huang, 2000). For example, a longer first event time would typically indicate a greater probability of censoring for the second event. Thus, if the within-subject gap times are correlated, the second and subsequent gap times will depend on the censoring variables. The second issue is non-identifiability. Specifically, unless the gap times are independent, when the support of the first gap time is not contained within the support of the censoring distribution, the marginal distributions of the second and subsequent gap times cannot be identified without making parametric assumptions (Lin et al., 1999; Wang, 1999; Huang, 2002; Schaubel and Cai, 2004a). To work around these issues, we propose to use conditional modeling in pursuit of marginal inference.

Methods in this report are motivated by comparisons of graft survival between primary versus repeat kidney transplantation. This is a controversial research question of great interest to transplant surgeons and patients, which cannot be accurately addressed using existing gap time regression methods. The preferred therapy for patients with end-stage renal disease (ESRD) is kidney transplantation, due to increased survival and quality of life compared to the alternative, dialysis. Re-transplantation may be required if the original kidney transplant fails. In 2012, there were more than 104,000 patients on the waiting list for kidney transplantation, while the number of donor kidneys transplanted was approximately 13,000 (www.unos.org). Due to the relative scarcity of donor kidneys, it is meaningful to study whether patients with a repeat transplant have inferior outcomes compared to patients with a primary transplant. Such results would provide evidence to potentially serve as the basis for future organ allocation policy.

It has been frequently reported in the literature that graft survival is significantly lower for re-transplants relative to primary kidney transplants (Tejani and Sullivan, 1996; Pour-Reza-Gholi et al., 2005; Ahmed et al., 2008). However, there are also a few studies indicating that there was no significant difference (Gruber et al., 2009; Barba et al., 2011). Most of these articles used the Kaplan-Meier method (Kaplan and Meier, 1958) and hence were not covariate-adjusted. In addition, the issues of non-identifiability and induced dependent censoring were not taken into consideration. Given the complexities of the data structure, much more detailed and robust analysis is required.

Our chief objective is to compare the average graft survival curve (i.e., an identifiable version thereof) for repeat kidney transplant patients, to the analogously averaged first-transplant survival curve. Rather than carry out predictions, our interest is in comparing primary and repeat transplant survival with respect to average survival, with the averaging being across which patients (i.e., as indexed by the covariate vector) receive a repeat transplant and when (in terms of follow-up time since first transplant). Consider the survival function for a re-transplant patient. Is this survival function really lower than that which would apply if in fact the patient were instead receiving a primary kidney transplant? If graft survival is truly lower for re-transplants, this would call into question the current policy of essentially assigning equal priority to primary and repeat kidney transplant candidates. Note that patient-specific contrasts between first and second transplant survival are at most of

secondary interest; this is particularly true from a public health perspective, due to the impracticality of implementing patient-tailored organ allocation rules. In contrast, a global policy (applying to all patients) is feasible and is more likely to be perceived as fair by surgeons, patients and the public.

Many nonparametric gap time methods have been proposed, including Wang and Wells (1998), Lin, Sun and Ying (1999), Wang and Chang (1999), Peña, Strawderman and Hollander (2001), van der Laan, Hubbard and Robins (2002), Schaubel and Cai (2004a), and Andrei and Murray (2006). The majority of these works developed nonparametric methods to estimate the joint and/or conditional distribution of the gap times. Semiparametric regression models have been proposed; e.g., Prentice, Williams and Peterson (1981) which assumed proportional hazards (stratified on gap time) and assumed independent within-subject gap times. Huang (2002) proposed gap time regression methods based on the accelerated failure time model. Chen, Wang, and Huang (2004) proposed stratified proportional reverse-time hazards models to estimate a longitudinal pattern of gap times. Schaubel and Cai (2004b) developed regression methods for the gap time hazard functions. Strawderman (2005) extended the accelerated failure time model for gap times that are independent conditional on the covariate. The method was subsequently extended to accommodate correlated gap times through a multiplicative gamma frailty (Strawderman, 2006). Huang and Liu (2007) used a joint frailty model to analyze disease recurrences and survival. Clement and Strawderman (2009) adapted generalized estimating equations to estimate the parameters indexing the conditional means and variances of gap times.

Most existing regression methods target covariate effects within gap time, as opposed to contrasts between gap times themselves. One could append the first and second gap time data sets, then fit a marginal (common baseline) Cox (1972) model with an indicator for second gap time (first gap time then serving as the reference). This could be interpreted as a version of the Prentice et al. (PWP; 1981) method, or a form of the Wei, Lin and Weissfeld (1989) approach. Such a procedure (which is clearly not in line with the intended use of either PWP or WLW) would be biased due to failing to address either the previously mentioned identifiability or induced dependent censoring issues. One could not fit a frailty version of this model since there would be no repeated events within-individual (with covariate fixed; e.g., with an indicator for second-gap-time included in the covariate vector), and hence no information to estimate the frailty variance. The methods proposed by Chen, Wang, and Huang (2004) could be used to compare gap times, through an estimated longitudinal pattern parameter. The pattern parameter quantifies the increasing or decreasing trend in gap times, although a monotone trend is assumed. In addition, the subject-specific baseline hazard functions are sometimes unidentifiable. In summary, there are very few methods in the existing literature for comparing gap times, and the obvious extensions to existing methods have substantial limitations.

In this report, we develop methods to compare the survival functions and restricted mean lifetimes of the first (T_{1l}) and the second gap times (T_{2l}). In particular, we contrast the average survival function for T_{2l} (obtained through appropriate conditioning, such as to respect the afore-described identifiability issues) with the corresponding survival function for T_{1l} obtained through the same averaging. The method we propose does not require

inverse weighting or frailty modeling, and works around the issues of non-identifiability and dependent censoring through flexible assumptions regarding the association between T_{11} and T_{12} . Specifically, we contrast first and second gap time survival functions, as well as their integration over $[0, L]$ for pre-specified L . This difference in restricted mean survival times has been studied by many authors in various contexts (e.g., Zucker, 1998; Chen and Tsiatis, 2001; Zhang and Schaubel, 2011; Zhang and Schaubel, 2012).

The remainder of this report is organized as follows. In Section 2, we describe the measures proposed to compare gap times and their corresponding estimation procedures. In Section 3, the asymptotic properties of the proposed estimators are derived. A simulation study is described in Section 4. In Section 5, we apply the proposed methods to kidney transplant data obtained from a national registry. Some remarks and discussion are given in Section 6.

2. Proposed Methods

In this section, we begin by formalizing the data structure and issues described in Section 1. We then describe the contrast of interest. This is followed by a description of the assumed models and proposed estimation procedures.

2.1 Notation and data structure

We first introduce the requisite notation. Let T_{ij} denote the j th total time ($j = 1, 2$) for subject i ($i = 1, \dots, n$). To make our description more concrete, suppose that we are interested in comparing the first two gap times, T_{11} and $T_{12} = T_{22} - T_{11}$. The case of comparing three or more gap times will be discussed in Section 6. The censoring time of the i th subject is denoted as C_i . Hence, T_{11} is potentially censored by C_i and T_{12} is potentially censored by $C_{12} = C_i - T_{11}$. We let $\mathbf{Z}_i$ denote a vector of covariates for the i th subject, measured at baseline. Time-dependent covariates will be discussed in Section 6. We let $\tau_1 = \sup\{t: P(C_i > t) > 0\}$ denote the upper bound of the support of the first gap time's censoring distribution, and $\tau_2 = \sup\{t: P(C_{12} > t) > 0\}$. We define the counting process, $N_{11}(t) = I(T_{11} \leq t \wedge C_i)$. As well, it is convenient to define $N_{\bullet 1}(t) = \sum_{i=1}^n N_{i1}(t)$.

Since T_{11} and T_{12} are not likely to be independent in most biomedical examples, the two challenges described in Section 1, induced dependent censoring and non-identifiability, remain in the context of existing non- and semiparametric methods. In particular, T_{12} is censored by $C_{12} = C_i - T_{11}$. Therefore, even if C_i is independent of both T_{11} and T_{12} , T_{12} will still be censored by a variate with which it is correlated. Further, although $P(T_{11} > t)$ is identifiable nonparametrically on $[0, \tau_1]$, $P(T_{12} > t)$ is not identifiable without assumptions on the nature of the association between T_{11} and T_{12} . However, conditional survival functions for T_{12} can be identified. For example, $P(T_{12} > t | T_{11} = t_1)$ is identifiable for a pre-specified and fixed time point, t_1 ; but, only on $t \in [0, \tau_1 - t_1]$. This notion has been used by previous authors (Lin et al., 1999; Schaubel and Cai, 2004a), but restrictions on the inference are clear. Although it leads to a useful description of T_{12} nonparametrically, this construct is of limited value with respect to comparing gap times and was not motivated by such comparisons.

2.2 Contrasting first and second gap times

Recall that our objective is to compare the first and second gap times. As implied in Section 2.1, the subjects are not homogenous, each being indexed by a covariate vector, $\mathbf{Z}_i$. Although, correspondingly, it is possible that the contrast between first and second gap times interacts with $\mathbf{Z}_i$, our interest is primarily in the average contrast. Naturally, to be meaningful, the average taken across the second gap times must be consistent with that taken across the first gap times, such that confounding eliminated by incorporating $\mathbf{Z}_i$ is not reintroduced by the averaging. Since survival probability tends to be easily understood by clinical investigators, we choose to contrast the gap times through differences in the survival function, and the integration thereof (restricted mean gap times).

To further elaborate on our perspective, consider again the motivating example. We could take an appropriately defined average graft survival function for repeat kidney transplants. A specific covariate distribution was used in deriving this average, and the same distribution would be used to average over the covariate-specific graft survival function for first transplants. The difference could then be taken (in order to compute the difference in graft survival probability) and integrated (to obtain the difference in mean graft survival time, capped at 10 years).

We now formalize the concepts described above, starting with the second gap time, T_{i2} . As described in Section 1, we are unable to estimate $P(T_{i2} > t | \mathbf{Z}_i)$, but can estimate $S_{i2}(t | \mathbf{Z}_i, T_{i1}; \tau_1) \equiv P(T_{i2} > t | \mathbf{Z}_i, T_{i1}, T_{i1} \leq \tau_1)$ for $t \in [0, \tau_1]$. Taking the area under the curve, we can estimate $\mu_{i2}(L | \mathbf{Z}_i, T_{i1}; \tau_1) \equiv \int_0^L S_{i2}(t | \mathbf{Z}_i, T_{i1}; \tau_1) dt$, with $L = \tau_1$. We let $q_{i1}(u, \mathbf{Z}_i)$ denote the joint density of $(T_{i1}, \mathbf{Z}_i)$; we then define $\pi_{i1}(u, \mathbf{Z}_i; \tau_1)$ to be the corresponding joint density of across the observable region pertaining to the first gap time,

$$\pi_{i1}(u, \mathbf{Z}_i; \tau_1) = \frac{q_{i1}(u, \mathbf{Z}_i)}{\int_{\mathbf{Z}} \int_0^{\tau_1} q_{i1}(u, \mathbf{Z}_i) du d\mathbf{Z}_i},$$

where $\int_{\mathbf{Z}}$ represents an integral of dimension equal to that of $\mathbf{Z}_i$. Taking an average across $\{T_{i1}, \mathbf{Z}_i : T_{i1} \leq \tau_1\}$, we obtain

$$S_2(t; \tau_1) = \int_{\mathbf{Z}} \int_0^{\tau_1} S_{i2}(t | \mathbf{Z}_i, u; \tau_1) \pi_{i1}(u, \mathbf{Z}_i; \tau_1) du d\mathbf{Z}_i. \quad (1)$$

Note that $\int_{\mathbf{Z}} \int_0^{\tau_1} \pi_{i1}(u, \mathbf{Z}_i; \tau_1) du d\mathbf{Z}_i = 1$ is a valid joint distribution of $\{(T_{i1}, \mathbf{Z}_i) : T_{i1} \in (0, \tau_1]\}$. Having defined an appropriate survival function, we can then take

$\mu_2(L; \tau_1) = \int_0^L S_2(t; \tau_1) dt$. We compute the average survival for the first gap time by taking an average analogous to (1), which implies using

$$S_1(t; \tau_1) = \int_{\mathbf{Z}} \int_0^{\tau_1} S_{i1}(t | \mathbf{Z}_i) \pi_{i1}(u, \mathbf{Z}_i; \tau_1) du d\mathbf{Z}_i. \quad (2)$$

Note that, as defined in (2), $S_1(t; \tau_1) = P(T_{i1} > t | T_{i1} \leq \tau_1)$, which would not yield an appropriate comparison. The survival function $S_1(t; \tau_1)$ was derived specifically as an appropriate comparator to $S_2(t; \tau_1)$ and its utility is mostly tied to that purpose. In the context of the kidney transplant example, $S_2(t; \tau_1)$ represents that appropriately averaged survival function for second transplant patients. The quantity $S_1(t; \tau_1)$ reflects survival after first transplant, averaged across the $\mathbf{Z}_i$ component used in the calculation of $S_2(t; \tau_1)$. That is, we have forced the survival functions being compared, $S_2(t; \tau_1)$ and $S_1(t; \tau_1)$, to be averaged across the same covariate values, in order to avoid introducing confounding. Finally, difference in the average survival curves is denoted $\delta(t; \tau_1) = S_2(t; \tau_1) - S_1(t; \tau_1)$, with the area between the survival curves given by $\Delta(L; \tau_1) = \int_0^L \delta(t; \tau_1) dt$.

2.3 Assumed models and proposed estimators

We seek to compare the first and second gap times in a manner which allows us to use all of the observed event times and does not require inverse weighting, without imposing unrealistic or unverifiable modeling assumptions. We therefore keep the modeling within a framework where model checking and validation are well-established. Along those lines, we assume that the first gap time follows a proportional hazards model (Cox, 1972),

$$\lambda_{i1}(t | \mathbf{Z}_i) = \lambda_{01}(t) \exp\{\boldsymbol{\beta}'_1 \mathbf{Z}_i\}, \quad (3)$$

where $\lambda_{i1}(t | \mathbf{Z}_i) = \lim_{\delta \rightarrow 0} \delta^{-1} P(t - T_{i1} < t + \delta | T_{i1} = t, \mathbf{Z}_i)$. We chose a proportional hazards model because it is commonly used in censored data; it is flexible; and model checking procedures are widely available (e.g., Klein and Moeschberger, 2003). To address identifiability issues, we build a connection between the first and second gap time. We choose to work with the hazard function for the conditional variate, $\{T_{i2} | \mathbf{Z}_i, T_{i1}, T_{i1} \leq \tau_1\}$,

$$\lambda_{i2}(t | \mathbf{Z}_i, T_{i1}; \tau_1) = \lim_{\delta \rightarrow 0} \delta^{-1} P(t \leq \tilde{T}_{i2} < t + \delta | \tilde{T}_{i2} \geq t, \mathbf{Z}_i, T_{i1}, T_{i1} \leq \tau_1).$$

In particular, we assume that this quantity follows the proportional hazards model,

$$\lambda_{i2}(t | \mathbf{Z}_i, T_{i1}; \tau_1) = \lambda_{02}(t; \tau_1) \exp\{\boldsymbol{\beta}'_2 \mathbf{Z}_i + \boldsymbol{\phi}'_2 \mathbf{f}(T_{i1})\}, \quad (4)$$

where $\mathbf{f}(x)$ is a parametric possibly vector valued function of x . Model (3) and (4) together allow one to not only quantify the covariate effects on the hazards of first and second gap times, but also quantify the connection between the gap times. Moreover, the connection is parameterized in a very flexible way in model (4), because $\mathbf{f}$ can take a large number of possible forms (e.g., polynomial, spline, etc.). In order to decide what form $\mathbf{f}$ should take, one common strategy is to break continuous T_{i1} into a categorical variable through a set of functions. The model would then be fitted with the categorical version of T_{i1} in order to

determine an appropriate functional form for f . Denote $\theta_2 = (\beta_2', \phi_2')'$ and set

$$\mathbf{W}_i = (\mathbf{Z}_i', f(T_{i1}))'.$$

The parameters β_1 from model (3) and θ_2 from model (4) can be estimated through partial likelihood (Cox, 1975); while Breslow (1972) estimators are available for $\Lambda_{01}(t)$ and $\Lambda_{02}(t; \tau_1)$. After fitting models (3) and (4), the corresponding subject-specific survival functions can be estimated as follows

$$\begin{aligned}\hat{S}_{i1}(t|\mathbf{Z}_i) &= \exp\{-\hat{\Lambda}_{i1}(t|\mathbf{Z}_i)\} & \hat{\Lambda}_{i1}(t|\mathbf{Z}_i) &= \hat{\Lambda}_{01}(t) \exp\{\hat{\beta}_1' \mathbf{Z}_i\} \\ \hat{S}_{i2}(t|\mathbf{Z}_i, T_{i1}; \tau_1) &= \exp\{-\hat{\Lambda}_{i2}(t|\mathbf{Z}_i, T_{i1}; \tau_1)\} & \hat{\Lambda}_{i2}(t|\mathbf{Z}_i, T_{i1}; \tau_1) &= \hat{\Lambda}_{02}(t; \tau_1) \exp\{\hat{\theta}_2' \mathbf{W}_i\}.\end{aligned}$$

Ultimately, we will be averaging over $\{(\mathbf{Z}_i, T_{i1}) : T_{i1} < \tau_1\}$. The observed-data version of such averaging will depend on the censoring distribution, which of course is undesirable. As such, we multiply impute censored T_{i1} values. Specifically, a total of M imputations will be generated such that, in each imputation m , we set $T_{i1}^m = T_{i1}$ for subjects with $T_{i1} < C_i$; otherwise, we impute T_{i1}^m from the truncated distribution,

$$\hat{P}(T_{i1} > t | C_i, T_{i1} > C_i, \mathbf{Z}_i) \equiv \frac{\hat{S}_{i1}(t|\mathbf{Z}_i)I(t > C_i)}{\hat{S}_{i1}(C_i|\mathbf{Z}_i)}. \quad (5)$$

It is natural to use this distribution because the only information we have is that the imputed T_{i1}^m will be larger than the censoring time C_i . Owing to its nonparametric component, the survival function estimator is not defined after τ_1 . However, as will be described shortly, if $T_{i1}^m > \tau_1$, then subject i does not contribute to the computation of the average survival curve for either the first or second gap time. Note that we used an ‘improper’ imputation method, referred to as Type-B imputation by Wang and Robins (1998) and Robins and Wang (2000), which means the estimated parameters $\hat{\beta}_1$ and $\hat{\Lambda}_{01}(t)$ used in (5) are only estimated once and held as fixed in the imputation algorithm. This precludes the use of the familiar techniques for variance estimation in the presence of multiple imputation (e.g., Little and Rubin, 2002), as will be seen later.

After imputing T_{i1}^m when $T_{i1} > C_i$ and setting $T_{i1}^m = T_{i1}$ when $T_{i1} < C_i$, we can compute

$$\hat{S}_{i2}(t|\mathbf{Z}_i, T_{i1}^m; \tau_1) = \exp\{-\hat{\Lambda}_{i2}(t|\mathbf{Z}_i, T_{i1}^m; \tau_1)\},$$

for the subset $\{i: T_{i1}^m \leq \tau_1\}$, where $\hat{\Lambda}_{i2}(t|\mathbf{Z}_i, T_{i1}^m; \tau_1) = \hat{\Lambda}_{02}(t; \tau_1) \exp\{\hat{\theta}_2' \mathbf{W}_i^m\}$, with $\mathbf{W}_i^m = (\mathbf{Z}_i', f(T_{i1}^m))'$.

In evaluating $S_{i2}(t|\mathbf{Z}_i, T_{i1}; \tau_1)$, a natural comparison is with $S_{i1}(t|\mathbf{Z}_i)$, which suggests the contrast, $S_{i2}(t|\mathbf{Z}_i, T_{i1}^m; \tau_1) - S_{i1}(t|\mathbf{Z}_i)$. For instance, in the context of kidney transplantation, both $\mathbf{Z}_i$ and T_{i1} are known at the time of the second transplant. In advising a patient about to undergo retransplantation (second kidney transplant), the survival distribution for the second gap time is naturally important; but, also the patient would likely be interested in how their graft failure the second time around would be (given what is known at the time of re-transplant: $\mathbf{Z}_i, T_{i1}$) compared to the risk they faced before the first transplant (given $\mathbf{Z}_i$). Note that, for the first and second gap time, the conditioning is on all information known at the respective gap time origins.

This gives rise to two useful contrasts, namely,

$$\hat{\delta}_i^m(t|\mathbf{Z}_i, T_{i1}^m; \tau_1) = \hat{S}_{i2}(t|\mathbf{Z}_i, T_{i1}^m; \tau_1) - \hat{S}_{i1}(t|\mathbf{Z}_i) \quad (6)$$

$$\hat{\Delta}_i^m(L|\mathbf{Z}_i, T_{i1}; \tau_1) = \int_0^L \{\hat{S}_{i2}(t|\mathbf{Z}_i, T_{i1}^m; \tau_1) - \hat{S}_{i1}(t|\mathbf{Z}_i)\} dt = \int_0^L \hat{\delta}_i^m(t|\mathbf{Z}_i, T_{i1}^m; \tau_1) dt. \quad (7)$$

The contrast in (6) represents the estimated distance between the first and second gap time survival functions, while (7) reflects the area between the survival functions over $[0, L]$. The conditional distribution $\hat{S}_{i2}$ is an identifiable version of the survival function for the second gap time, while $\hat{S}_{i1}$ is the marginal distribution for the first gap time. The survival functions are evaluated at the same covariate values in order to avoid confounding. Note that, since $\hat{S}_{i2}$ represents a conditional quantity (and, in particular, since the conditioning involves the first gap time), care must be taken in deciding on the comparator. We return to this issue in Section 6.

Note that the contrasts are, now, specifically for each subject. Following (2) and (1), the average difference in gap time survival is given by

$$\hat{\delta}^m(t; \tau_1) = \int_{\mathbf{Z}} \int_0^{\tau_1} \hat{\pi}_{i1}^m(u, \mathbf{Z}_i; \tau_1) \hat{\delta}_i^m(t|\mathbf{Z}_i, u; \tau_1) du d\mathbf{Z}_i, \quad (8)$$

where $\hat{\pi}_{i1}^m(u, \mathbf{Z}_i; \tau_1) = N_{\bullet 1}^m(\tau_1)^{-1} \int_0^{\tau_1} dI\{T_{i1}^m \leq u\}$ with $N_{\bullet 1}^m(\tau_1) = \sum_{i=1}^n I\{T_{i1}^m \leq \tau_1\}$. The difference between the restricted mean lifetimes is then estimated by

$\hat{\Delta}^m(L; \tau_1) = \int_0^L \hat{\delta}^m(t; \tau_1) dt$. The final estimates are averages of the M estimators obtained through multiple imputation:

$$\hat{\delta}(t; \tau_1) = M^{-1} \sum_{m=1}^M \hat{\delta}^m(t; \tau_1) \quad \hat{\Delta}(L; \tau_1) = \int_0^L \hat{\delta}(t; \tau_1) dt.$$

Note that, based on the above developments, we have $\hat{\mathcal{X}}(t; \tau_1) = \hat{S}_2(t; \tau_1) - \hat{S}_1(t; \tau_1)$ and $\hat{\mu}(L; \tau_1) = \mu_2(L; \tau_1) - \mu_1(L; \tau_1)$, where

$$\hat{S}_2(t; \tau_1) = M^{-1} \sum_{m=1}^M N_{\bullet 1}^m(\tau_1)^{-1} \sum_{i=1}^n N_{i1}^m(\tau_1) \hat{S}_{i2}(t | \mathbf{Z}_i, T_{i1}^m; \tau_1)$$

$$\hat{S}_1(t; \tau_1) = M^{-1} \sum_{m=1}^M N_{\bullet 1}^m(\tau_1)^{-1} \sum_{i=1}^n N_{i1}^m(\tau_1) \hat{S}_{i1}(t | \mathbf{Z}_i),$$

with $\hat{\mu}_j(L; \tau_1) = \int_0^L \hat{S}_j(t; \tau_1) dt$ for $j = 1, 2$.

3. Asymptotic Properties

We begin by establishing counting processes corresponding to the observed gap times.

Recall (Section 2.1) that we defined $N_{i1}(t) = \mathbb{I}(T_{i1} \leq t \wedge C_i)$ and $N_{\bullet 1}(t) = \sum_{i=1}^n N_{i1}(t)$ corresponding to T_{i1} . With respect to the imputed T_{i1} , we also defined $N_{i1}^m(t) = \mathbb{I}(T_{i1}^m \leq t)$ and $N_{\bullet 1}^m(t) = \sum_{i=1}^n N_{i1}^m(t)$. For the second gap time, $\{T_{i2} | T_{i1} \leq \tau_1 \wedge C_i\}$, we now define $\tilde{N}_{i2}(t) = \mathbb{I}(T_{i2} \leq t \wedge C_i, T_{i1} \leq \tau_1 \wedge C_i)$. The at-risk processes are given by $Y_{i1}(t) = \mathbb{I}(T_{i1} \wedge C_i \leq t)$ and $\tilde{Y}_{i2}(t) = \mathbb{I}(T_{i2} \wedge C_i \leq t, T_{i1} \leq \tau_1 \wedge C_i)$, respectively. Then, the pertinent zero-mean processes are given by $M_{i1}(t) = N_{i1}(t) - \int_0^t \lambda_{i1}(u) Y_{i1}(u) du$, and $\tilde{M}_{i2}(t) = \tilde{N}_{i2}(t) - \int_0^t \lambda_{i2}(u; \tau_1) \tilde{Y}_{i2}(u) du$.

The essential asymptotic properties are summarized by the following two theorems. Note that the assumed regularity conditions are listed in the Supplementary Materials document.

Theorem 1

Under conditions (a) to (e), we have the following linear representations pertinent to the second gap time asymptotically,

$$n^{1/2} [\hat{S}_2(t; \tau_1) - S_2(t; \tau_1)] = n^{-1/2} \sum_{i=1}^n \varphi_{i1}(t) + o_p(1) \quad (9)$$

$$n^{1/2} [\hat{\mu}_2(L; \tau_1) - \mu_2(L; \tau_1)] = n^{-1/2} \sum_{i=1}^n \varphi_{i2}(L) + o_p(1), \quad (10)$$

where $\varphi_{i1}(t)$ and $\varphi_{i2}(L)$ ($i = 1, \dots, n$) are independent and identically distributed mean-zero random variables, such that $E\{\varphi_{i1}(t)^2\} < \infty$, $E\{\varphi_{i2}(L)^2\} < \infty$. Thus, (9) and (10) are

asymptotically normal with mean 0 and variances $E\{\phi_{i1}(t)^2\}$ and $E\{\phi_{i2}(L)^2\}$, respectively, with

$$\varphi_{i1}(t) = M^{-1} \sum_{m=1}^M \{P(T_{i1} \leq \tau_1 \wedge C_i)^{-1} N_{i1}(\tau_1) \phi_{i1}^m(t) + P(T_{i1} \leq \tau_1)^{-1} N_{i1}^m(\tau_1) [S_{i2}(t | \mathbf{W}_i^m; \tau_1) - S_2(t; \tau_1)]\},$$

$$\varphi_{i2}(L) = \int_0^L \varphi_{i1}(t) dt,$$

with $\phi_{i1}^m(t)$ defined in the Supplementary Materials, along with expressions for consistent variance estimators.

The asymptotic linear representations of (9) and (10) follow from the large-sample results of Andersen and Gill (1982), under the implicit assumption that the imputation model is correctly specified (such that T_{i1}^m has the same distribution as T_{i1}).

Theorem 2

Under conditions (a) to (e), we have the following linear representations pertinent to the first gap time asymptotically,

$$n^{1/2}[\hat{S}_1(t; \tau_1) - S_1(t; \tau_1)] = n^{-1/2} \sum_{i=1}^n \varphi_{i3}(t) + o_p(1) \quad (11)$$

$$n^{1/2}[\hat{\mu}_1(L; \tau_1) - \mu_1(L; \tau_1)] = n^{-1/2} \sum_{i=1}^n \varphi_{i4}(L) + o_p(1), \quad (12)$$

where $\phi_{i3}(t)$ and $\phi_{i4}(L)$ ($i = 1, \dots, n$) are independent and identically distributed mean-zero random variables, such that $E\{\phi_{i3}(t)^2\} < \infty$, $E\{\phi_{i4}(L)^2\} < \infty$. Thus, (11) and (12) are asymptotically normal with means 0 and variances $E\{\phi_{i3}(t)^2\}$ and $E\{\phi_{i4}(L)^2\}$, respectively. Expressions for $\phi_{i3}(t)$ and $\phi_{i4}(L)$ are provided in the Supplementary Materials, along with expressions for consistent estimators of the corresponding variances.

Combining the results in Theorem 1 and 2, the asymptotic properties for $n^{1/2}\{\hat{\mathcal{X}}(t; \tau_1) - \mathcal{X}(t; \tau_1)\}$ and $n^{1/2}\{\hat{\mathcal{X}}(L; \tau_1) - \mathcal{X}(L; \tau_1)\}$ can be readily summarized in the following theorem.

Theorem 3

Under conditions (a) to (e), $n^{1/2}\{\hat{\mathcal{X}}(t; \tau_1) - \mathcal{X}(t; \tau_1)\}$ and $n^{1/2}\{\hat{\mathcal{X}}(L; \tau_1) - \mathcal{X}(L; \tau_1)\}$ have linear representations asymptotically, i.e.,

$$n^{1/2}\{\hat{\delta}(t; \tau_1) - \delta(t; \tau_1)\} = n^{-1/2} \sum_{i=1}^n [\varphi_{i1}(t) - \varphi_{i3}(t)] + o_p(1);$$

$$n^{1/2}\{\hat{\Delta}(L; \tau_1) - \Delta(L; \tau_1)\} = n^{-1/2} \sum_{i=1}^n [\varphi_{i2}(L) - \varphi_{i4}(L)] + o_p(1),$$

where $\varphi_{i1}(t)$, $\varphi_{i2}(L)$, $\varphi_{i3}(t)$, $\varphi_{i4}(L)$, $i = 1, \dots, n$ are the same as above. Thus, $n^{1/2}\{\hat{\delta}(t; \tau_1) - \delta(t; \tau_1)\}$ and $n^{1/2}\{\hat{\Delta}(L; \tau_1) - \Delta(L; \tau_1)\}$ are asymptotically normal with means 0 and variances $E\{[\varphi_{i1}(t) - \varphi_{i3}(t)]^2\}$, $E\{[\varphi_{i2}(L) - \varphi_{i4}(L)]^2\}$, respectively. The variances can be estimated as $n^{-2} \sum_{i=1}^n [\hat{\varphi}_{i1}(t) - \hat{\varphi}_{i3}(t)]^2$ and $n^{-2} \sum_{i=1}^n [\hat{\varphi}_{i2}(L) - \hat{\varphi}_{i4}(L)]^2$.

4. Simulations

We first describe the settings used in our simulation study. Each subject had two binary covariates Z_{i1} and Z_{i2} with $\Pr\{Z_{i1} = 1\} \Pr\{Z_{i2} = 1\} = 0.5$. For each subject, two gap times T_{i1} and T_{i2} were generated from the following proportional hazards models:

$$\lambda_{i1}(t) = \lambda_{01}(t) \exp\{\beta_1 Z_{i1} + \beta_2 Z_{i2}\}$$

$$\lambda_{i2}(t) = \lambda_{02}(t) \exp\{\beta_3 Z_{i1} + \beta_4 Z_{i2} + \beta_5 T_{i1}\}.$$

Parameter values used in Settings 1-4 are listed at the bottom of Table 1. The censoring time, C_i , followed a Uniform (0, 12) distribution. We set $L = \tau_1 = 5$, and the number of multiple imputations to $M = 5$. The sample size was $n = 250$ for each data configuration, and we ran 1,000 replicates per configuration.

In Table 1, we present results from four parameter settings. In Settings 1-2, survival is much greater for T_{i1} than T_{i2} , while the opposite is true for Settings 3-4. In each setting, bias is very small for both $\hat{\delta}(L; \tau_1)$ and the estimated survival probabilities. The estimated standard deviations (ESDs) and asymptotic standard errors (ASEs) match quite well, indicating that our asymptotic variance estimators are fairly accurate in reasonable size samples. Empirical coverage probabilities (ECP) are all around 0.95. The bias and discrepancy between the ASE and ESD are much larger for $\hat{\delta}(L; \tau_1)$ than those of the estimated survival functions, since restricted mean lifetimes can be viewed as an accumulation of survival probability, such that bias essentially propagates as t increases. Another thing to note is that the estimated survival probabilities at later time points are often more biased compared to those at earlier time points, which is intuitive because data are more sparse towards the tail of the observation time distribution.

Additional data configurations are shown in the Supplementary Materials. Overall, the proposed methods are demonstrated to work well under the scenarios considered.

5. Application to kidney transplant data

We applied the proposed methods to kidney transplant data obtained from the Scientific Registry of Transplant Recipients (SRTR). The SRTR data system includes data on all donors, wait-listed candidates, and transplant recipients in the United States; these data are submitted by the members of OPTN and have been described elsewhere. The Health Resources and Services Administration (US Department of Health and Human Services) provides oversight for the activities of the OPTN and SRTR contractors. The survival time of interest is time between kidney transplantation and graft failure, where graft failure is said to occur when the patient dies or the transplanted kidney ceases to function. A patient can have multiple kidney transplants if graft failure of the previous transplant(s) occurred. Our objective is to contrast the first and second transplants with respect to graft survival and restricted mean graft survival time. We included adult patients (age ≥ 18) who had their first transplant between January 1, 1998 and December 31, 2011. The observation period concluded on December 31, 2011. In our analysis we only include transplants from deceased donors.

Recipient-specific covariates include age at transplant, gender, race, diabetes status, body mass index (BMI), time waited for a transplant, calendar year of the transplant, and panel reactive antibodies (PRA). Covariates based on the donor include donor age, BMI, serum creatinine, whether death was caused by stroke, hypertension, and diabetes status and duration. The covariate vector for each subject is recorded at each transplant and, hence, is transplant-specific. We estimated the functional form of $f(T_1)$ by first modeling a categorized version of T_1 , in order to investigate the association pattern (discussed in Section 2). An appropriate function was determined to be $f(T_1) = (T_1 - 4)I(T_1 > 4)$, with time given in years. A table summarizing the covariate distributions at each transplant is provided in the Supplementary Materials.

Proportional hazards models were fitted for each of the two gap times. There are $n = 113,621$ subjects in total and the Cox model for the first gap time T_1 was fitted using $n_1 = 113,246$ subjects with no missing covariates at first transplant. Among the n_1 subjects, 39% are female; and 48% are white. The mean age at transplant is 52, with a standard deviation of 13. There are 39,817 graft failures or deaths after the first transplant.

We set $L = \tau_1 = 10$ years, which uses most of the available data while using values of τ_1 and L that would be meaningful to nephrologists and patients. There are 2,765 subjects with a second transplant that occurred within 10 years from the first transplant. Among those, $n_2 = 2,630$ subjects do not have any missing covariates at second transplant and were used to fit a Cox model for the second gap time T_2 . Of the n_2 subjects, 38% are female; and 52% are white. The mean age at transplant is 49, with a standard deviation of 13. There are 793 graft failures or deaths after the second transplant. As per the proposed methods, we average the primary- and repeat-transplant survival functions using the same covariate distribution. In cases where T_{i1}^m was imputed, the covariate from the primary transplant was used. Note that this would not affect the model fitting for re-transplanted patients, as it comes into play after model (4) has already been fitted. We used $M = 5$ multiple imputations.

Estimated average survival curves for first and second transplants are shown in Figure 1. The estimated survival function of the second gap time is below that of the first gap time for the first 8 years after transplant. However, the curves seem to get close and overlap after 8 years. The $\hat{S}_1(t)$ curve is more smooth than $\hat{S}_2(t)$, especially in the tail region, because the sample size is so much larger for the first gap time.

Estimated average 10-year mean graft survival times are contrasted in Table 2. Estimated restricted mean lifetime for the first gap time was $\hat{\mu}_1(10; 10) = 6.916$ years with an estimated standard deviation of 0.014 years. The estimated restricted mean lifetime for the second gap time was $\hat{\mu}_2(10; 10) = 6.588$ years with an estimated standard deviation of 0.144 years. The occurrence of this relatively large standard deviation is due to the respectively small sample size with respect to T_2 . The difference between $\hat{\mu}_2(10; 10)$ and $\hat{\mu}_1(10; 10)$, $\hat{\mu}_1(10; 10)$, was -0.328 years (i.e., approximately 3.9 months), with an estimated standard deviation of 0.144 years. Thus, there is a statistically significant difference between first and second kidney transplants with respect to mean 10-year graft survival. However, the clinical importance of a difference of 0.144 additional years is debatable.

We carried out various model diagnostics familiar to Cox regression (e.g., see Klein and Moeschberger, 2003). As an assessment of overall fit, the Cox-Snell residuals are plotted in Figure 2 for the first and second transplants (top left and top right panels, respectively). That the lines in each plot are approximately straight (except for the very end of follow-up) indicates no evidence of lack of fit in a general sense. Another concern was that the contrast in average graft survival would obscure important relationships at the patient-level. For example, it is possible that various covariates have important but opposite effects on the first and second transplant survival. This is assessed in Figure 2 (bottom left panel) where we plot scaled versions of the Z -scores for coefficients (scaled, to account for n_1 relative to n_2) from the T_{11} and T_{12} Cox models. Most points in this plot are in the upper right or lower left quadrant, indicating that the direction of the covariate effect is usually the same for the first and second transplant survival. Moreover, most points are close to the 45 degree line, indicating that the magnitude of the effect is typically quite similar for first and second transplants. A complete listing of parameter estimates and SEs for the first and second models is available from the Supplementary Materials. We also plotted a histogram (bottom right panel) of the patient-specific $\hat{\lambda}(10; 10)$ values use to compute the average effect (listed in Table 2). The distribution is bell-shaped with light tails and centered at ≈ 0 . One would be concerned about using the mean if it appeared (from such a histogram) to be heavily influenced by the tail. However, in our application, the mean does appear to represent the center of the data, with the majority of the $\hat{\lambda}(10; 10)$ values being within ± 1 year. Hence, the mean is a reasonable summary measure in this application.

A naive way to contrast the first and second transplant would be to stack all data together, then fit a Cox regression with an indicator of re-transplant as a covariate. Doing so would be ignoring the induced dependent censoring and identifiability issues described in Section 1. Results from such an analysis of the model $\lambda_{12}(t) = \lambda_0(t) \exp\{\beta'Z_i + \theta\}$ are summarized in Table 3. The estimated hazard ratio for re-transplant was 1.24 ($p < 0.0001$). As such, the post-second-transplant graft failure hazard would be interpreted as 24% higher than that of the post-first-transplant, which, in addition to being statistically significant, would be

regarded as clinically important by most transplant surgeons and patients. The analysis in Table 3 is probably what would be carried out by the majority of data analysts unfamiliar with the statistical issues inherent to the gap time data structure.

6. Discussion

In this report, we propose semiparametric methods to compare the first and second gap times with respect to survival probability and restricted mean lifetime. Separate Cox models are assumed for the first and second gap times, with the first gap time used as a predictor of the hazard function for the second gap time. Multiple imputation of the first gap time is applied to identify subjects to be averaged in computing the survival probabilities and restricted mean lifetimes. Large-sample properties of the estimators are derived and demonstrated to work well in finite samples based on simulation studies.

We applied the proposed methods to compare the mean graft functioning lifetimes following first versus second kidney transplant, based on a 10-year time horizon. Our results imply that there is a significant difference between the two, which agrees with most existing studies (Tejani and Sullivan, 1996; Pour-Reza-Gholi et al., 2005; Ahmed et al., 2008) but contradicts with some recent analyses (Gruber et al., 2009; Barba et al., 2011). In contrast to results based on the proposed methods, results based on a model which simply used a second-transplant indicator show a significant 24% increase (Table 3) in the graft failure hazard associated with second kidney transplants, which would likely be viewed as clinically significant. Although statistically significant, the approximately 1/3 year difference in 10-year graft survival between primary and repeat kidney transplants is unlikely to be viewed as clinically important. The noteworthy difference in the results generated by the proposed method illustrates the importance of the method to the evaluation of re-transplantation.

One complicating issue in comparing graft failure between primary and repeat transplants is that patients with a repeat transplant necessarily did not die in experiencing their first graft failure; i.e., they had to have experienced transplant failure. To address this issue, one could eliminate death as a component of the graft failure definition. However, such an approach (referred to as “death-censored graft failure”) is clouded by the fact that, in reality, death can occur before transplant failure is detected, and as a result of the transplant failing. In such cases, censoring at death would introduce dependent censoring.

The proposed methods entail conditional inference on $T_{22}^{\sim}$ given $T_{11}^{\sim} = \tau_1$. It is then required to pre-specify τ_1 , which would typically be done before the analysis based on available follow-up and perhaps the general pattern of observed first and second gap time events. Under the proposed models and assumptions, one generally cannot identify the marginal distribution of $T_{22}^{\sim}$; the conditional distribution of the second gap time is identifiable for $T_{11}^{\sim} = \tau_1$. Regarding the choice of τ_1 , if one views marginal inference on $T_{22}^{\sim}$ as the gold standard, then the lower the choice of τ_1 , the more restricted is the resulting inference regarding the conditional distribution, $P(T_{22}^{\sim} > t | T_{11}^{\sim} = \tau_1)$. In principle, τ_1 could be any value at or below the largest observed censoring time. We would prefer τ_1 to be as large as possible so that we can include all available data into the analysis. From this angle, setting τ_1 to the largest censoring time makes sense. In practice, it would often make more sense to choose a ‘round’

number close to the maximum censoring time, which can be readily grasped by the intended audience. For example, we chose $\tau_1 = 10$ years in Section 5.

One advantage of our methods is the degree to which the observed data are utilized. In particular, various related procedures which do not model T_{i2}^* conditionally (given T_{i1}) are essentially forced into additional re-censoring; this is because the ‘useable’ range of T_{i2}^* depends on what range of T_{i1} is used. For example, in Schaubel and Cai (2004a), $P(T_{i2}^* > t | T_{i1} = t_1)$ is estimated nonparametrically and can only be identified for (t, t_1) such that $t + t_1 \leq \tau_1$.

The proposed contrast in (6) involves using the marginal survival function for $(T_{i1} | \mathbf{Z}_i)$ as the comparator to the conditional survival for $(T_{i2}^* | \mathbf{Z}_i, T_{i1}, T_{i1} = \tau_1)$. Assuming that T_{i1} and T_{i2}^* are positively correlated, it would be possible for the marginal survival functions of $T_{i1} | \mathbf{Z}_i$ and $T_{i2}^* | \mathbf{Z}_i$ to be equal, but with $\delta_X(t | \mathbf{Z}_i, T_{i1}; \tau_1) < 0$ from (6), due to the use of a T_{i2}^* survival which conditions on worse prognosis. Another choice for the comparator would be to preserve the conditionality on $T_{i1} = \tau_1$, but this would amount to contrasting the conditional survival for T_{i2}^* with a survival function for T_{i1} that is forced to drop to 0, which would be undesirable and lack meaningful interpretation.

The proposed methods are cast in terms of a baseline covariate, $\mathbf{Z}_i$. In settings where the covariates depend on follow up time, $\mathbf{Z}_i(t)$, one could model $\lambda_{i2}(t)$ as a function of $\{\mathbf{Z}_i(t), T_{i1}\}$, which would be like a partly conditional model (Zheng and Heagerty, 2005; Gong and Schaubel, 2013). A complication would be that $\mathbf{Z}_i(T_{i1}^m)$ is not known in cases where $T_{i1} > C_i$ such that an imputed $\mathbf{Z}_i^m(T_{i1}^m)$ would be required. One possibility would be to apply longitudinal models of the time-dependent elements of $\mathbf{Z}_i(t)$.

For simplicity of exposition, in this report we focused on comparing the first two gap times. However, there can be situations where comparing three or more gap times is of interest; including the clinical setting that motivated our work, repeat kidney transplantation. One could generally follow the same framework we proposed.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

This work was supported in part by National Institutes of Health Grant R01 DK070869 and by an M-Cubed Grant from the University of Michigan. The authors thank two anonymous Reviewers for their constructive comments which strengthened the manuscript. The data reported here have been supplied by the Minneapolis Medical Research Foundation (MMRF) as the contractor for the Scientific Registry of Transplant Recipients (SRTR). The interpretation and reporting of these data are the responsibility of the authors and in no way should be seen as an official policy of or interpretation by the SRTR or the U.S. Government.

References

- Ahmed K, Ahmad N, Khan MS, Koffman G, Calder F, Taylor J, et al. Influence of number of retransplants on renal graft outcome. *Transplant Proceedings*. 2008; 40:1349–1352.
- Andersen PK, Gill RD. Cox's regression model for counting processes: A large sample study. *The Annals of Statistics*. 1982; 10:1100–1120.

- Andrei AC, Murray S. Estimating the quality-of-life-adjusted gap time distribution of successive events subject to censoring. *Biometrika*. 2006; 93:343–355.
- Barba Abad J, Robles García JE, Saiz Sansi A, Tolosa Eizaguirre E, Romero Vargas L, Algarra Navarro R, et al. Impact of renal retransplantation on graft and recipient survival. *Arch Esp Urol*. 2011; 64:363–370. [PubMed: 21610281]
- Breslow NE. Contribution to the discussion of paper by D.R. Cox. *Journal of the Royal Statistical Society Series B (Methodological)*. 1972; 34:216–217.
- Chen PY, Tsiatis AA. Causal inference on the difference of the restricted mean lifetime between two groups. *Biometrics*. 2001; 57:1030–1038. [PubMed: 11764241]
- Chen YQ, Wang MC, Huang Y. Semiparametric regression analysis on longitudinal pattern of recurrent gap times. *Biostatistics*. 2004; 5:277–290. [PubMed: 15054031]
- Clement DY, Strawderman RL. Conditional GEE for recurrent event gap times. *Biostatistics*. 2009; 10:451–467. [PubMed: 19297655]
- Cox DR. Regression models and life-tables. *Journal of the Royal Statistical Society Series B (Methodological)*. 1972; 34:187–220.
- Gong Q, Schaubel DE. Partly conditional estimation of the effect of a time-dependent factor in the presence of dependent censoring. *Biometrics*. 2013; 69:338–347. [PubMed: 23635094]
- Gruber SA, Brown KL, El-Amm JM, Singh A, Mehta K, Morawski K, et al. Equivalent outcomes with primary and retransplantation in African-American deceased-donor renal allograft recipients. *Surgery*. 2009; 146:646–652. discussion 652–653. [PubMed: 19789023]
- Huang X, Liu L. A joint frailty model for survival time and gap times between recurrent events. *Biometrics*. 2007; 63:389–397. [PubMed: 17688491]
- Huang Y. Two-sample multistage accelerated sojourn times model. *Journal of the American Statistical Association*. 2000; 95:619–627.
- Huang Y. Calibration regression of censored lifetime medical cost. *Journal of the American Statistical Association*. 2002; 97:318–327.
- Jirka J, Reneltová I, Rossmann P, Skibová J, Macurová H, Chadimová M, et al. Repeat kidney transplantation. *Cas Lek Cesk*. 1998; 137:686–689. [PubMed: 9929935]
- Kaplan EL, Meier P. Nonparametric estimation from incomplete observations. *Journal of the American Statistical Association*. 1958; 53:457–481.
- Lin DY, Sun W, Ying Z. Nonparametric estimation of the gap time distributions for serial events with censored data. *Biometrika*. 1999; 86:59–70.
- Little, RJA.; Rubin, DB. *Statistical Analysis with Missing Data*. 2nd. New York: John Wiley; 2002.
- United Network for Organ Sharing (UNOS). Data. UNOS Web site. Updated August 30, 2013. Available at <http://www.unos.org/donation/index.php?topic=data>
- Klein, JP.; Moeschberger, ML. *Survival analysis: techniques for censored and truncated data*. 2nd. New York, New York: Springer; 2003.
- Peña EA, Strawderman RL, Hollander M. Nonparametric estimation with recurrent event data. *Journal of the American Statistical Association*. 2001; 96:1299–1315.
- Pour-Reza-Gholi F, Nafar M, Saeedinia A, Farrokhi F, Firouzan A, Simforoosh N, et al. Kidney retransplantation in comparison with first kidney transplantation. *Transplant Proceedings*. 2005; 37:2962–2964.
- Prentice RL, Williams BJ, Peterson AV. On the regression analysis of multivariate failure time data. *Biometrika*. 1981; 68:373–379.
- Robins JM, Wang N. Inference for imputation estimators. *Biometrika*. 2000; 87:113–124.
- Schaubel DE, Cai J. Non-parametric estimation of gap time survival functions for ordered multivariate failure time data. *Statistics in Medicine*. 2004a; 23:1885–1900. [PubMed: 15195322]
- Schaubel DE, Cai J. Regression methods for gap time hazard functions for sequentially ordered multivariate failure time data. *Biometrika*. 2004b; 91:291–303.
- Strawderman RL. The accelerated gap times model. *Biometrika*. 2005; 92:647–666.
- Strawderman RL. A regression model for dependent gap times. *The International Journal of Biostatistics*. 2006; 2:1557–4679. ISSN (Online).

- Tejani A, Sullivan EK. Factors that impact on the outcome of second renal transplants in children. *Transplantation*. 1996; 62:606–611. [PubMed: 8830823]
- van der Laan MJ, Hubbard AE, Robins JM. Locally efficient estimation of a multivariate survival function in longitudinal studies. *Journal of the American Statistical Association*. 2002; 97:494–507.
- Wang MC. Gap time bias in incident and prevalent cohorts. *Statistica Sinica*. 1999; 9:999–1010.
- Wang MC, Chang SH. Nonparametric estimation of a recurrent survival function. *Journal of the American Statistical Association*. 1999; 94:146–153. [PubMed: 24244058]
- Wang N, Robins JM. Large-sample theory for parametric multiple imputation estimators. *Biometrika*. 1998; 85:935–948.
- Wang W, Wells MT. Nonparametric estimation of successive duration times under dependent censoring. *Biometrika*. 1998; 85:561–572.
- Zucker DM. Restricted mean life with covariates: Modification and extension of a useful survival analysis method. *Journal of the American Statistical Association*. 1998; 93:702–709.
- Zhang M, Schaubel DE. Estimating differences in restricted mean lifetime using observational data subject to dependent censoring. *Biometrics*. 2011; 67:740–749. [PubMed: 21039400]
- Zhang M, Schaubel DE. Double-robust semiparametric estimator for differences in restricted mean lifetimes in observational studies. *Biometrics*. 2012; 68:999–1009. [PubMed: 22471876]

Figure 1.

Analysis of SRTR data: Comparison of graft survival for first and second kidney transplants.

The solid line is $\hat{S}_2(t; 10)$; the dashed line is $\hat{S}_1(t; 10)$, with t measured in years.

Figure 2.

Additional analysis of SRTR data: Cox-Snell residuals for first transplant model (top left) and second transplant model (top right); plot of coefficients for first vs second transplant model (bottom left); histogram of patient-specific fitted difference in 10-year mean post-transplant survival time (bottom right).

Table 1
Simulation results based on $n = 250$, $M = 5$ and 1000 replications per setting

Parameter	Setting 1					Setting 2				
	True	BIAS	ESD	ASE	ECP	True	BIAS	ESD	ASE	ECP
$\mu_1(L; \tau_1)$	3.08	-0.015	0.120	0.122	0.95	2.80	-0.007	0.123	0.128	0.95
$\mu_2(L; \tau_1)$	1.95	-0.007	0.154	0.146	0.94	1.77	-0.008	0.149	0.143	0.93
$(L; \tau_1)$	-1.13	0.008	0.189	0.195	0.95	-1.02	-0.001	0.175	0.199	0.97
$S_1(1; \tau_1)$	0.81	0.000	0.026	0.026	0.95	0.75	0.004	0.030	0.030	0.94
$S_2(1; \tau_1)$	0.62	0.004	0.044	0.043	0.94	0.56	0.001	0.044	0.043	0.95
$\mathcal{R}_1(1; \tau_1)$	-0.19	0.004	0.050	0.051	0.94	-0.20	-0.004	0.051	0.051	0.95
$S_1(3; \tau_1)$	0.53	-0.001	0.034	0.034	0.95	0.46	-0.000	0.034	0.034	0.94
$S_2(3; \tau_1)$	0.25	0.003	0.044	0.041	0.93	0.22	0.004	0.039	0.039	0.94
$\mathcal{R}_3(3; \tau_1)$	-0.28	0.004	0.054	0.054	0.95	-0.24	0.004	0.048	0.049	0.95
$S_1(5; \tau_1)$	0.35	0.001	0.034	0.034	0.95	0.30	0.001	0.030	0.030	0.96
$S_2(5; \tau_1)$	0.11	0.006	0.034	0.032	0.93	0.11	0.007	0.032	0.031	0.94
$\mathcal{R}_5(5; \tau_1)$	-0.24	0.005	0.046	0.047	0.95	-0.19	0.006	0.041	0.041	0.95

Parameter	Setting 3					Setting 4				
	True	BIAS	ESD	ASE	ECP	True	BIAS	ESD	ASE	ECP
$\mu_1(L; \tau_1)$	2.13	-0.006	0.106	0.109	0.96	2.03	-0.008	0.101	0.111	0.96
$\mu_2(L; \tau_1)$	3.21	-0.008	0.146	0.141	0.94	3.01	-0.006	0.150	0.143	0.94
$(L; \tau_1)$	1.08	-0.002	0.168	0.178	0.96	0.98	0.002	0.162	0.185	0.97
$S_1(1; \tau_1)$	0.66	-0.001	0.030	0.031	0.95	0.61	-0.000	0.033	0.032	0.94
$S_2(1; \tau_1)$	0.82	-0.000	0.023	0.029	0.94	0.78	0.003	0.032	0.031	0.94
$\mathcal{R}_1(1; \tau_1)$	0.16	0.001	0.040	0.041	0.96	0.17	0.003	0.032	0.031	0.94
$S_1(3; \tau_1)$	0.29	0.002	0.030	0.030	0.95	0.28	0.003	0.027	0.028	0.95
$S_2(3; \tau_1)$	0.56	0.003	0.039	0.034	0.95	0.51	0.002	0.041	0.040	0.94
$\mathcal{R}_3(3; \tau_1)$	0.27	0.001	0.047	0.049	0.96	0.23	-0.001	0.046	0.046	0.94
$S_1(5; \tau_1)$	0.14	0.002	0.024	0.024	0.95	0.15	0.001	0.021	0.022	0.95

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Parameter	Setting 1						Setting 2					
	True	BIAS	ESD	ASE	ECP	True	BIAS	ESD	ASE	ECP	True	BIAS
$\mathcal{S}_2(5; \tau_1)$	0.39	0.004	0.044	0.043	0.94	0.36	0.003	0.043	0.041	0.94	0.36	0.003
$\mathcal{S}_2(5; \tau_2)$	0.25	0.003	0.049	0.049	0.95	0.21	0.002	0.046	0.046	0.96	0.21	0.002
Setting 1: $\lambda_{01}(t)=0.2, \lambda_{02}(t)=0.4, \beta_1 = \beta_3 = \log(1.5), \beta_2 = \beta_4 = -\log(1.5), \beta_5 = \log(1.05)$												
Setting 2: $\lambda_{01}(t)=0.2, \lambda_{02}(t)=0.4, \beta_1 = \beta_3 = \log(2.5), \beta_2 = \beta_4 = -\log(2.5), \beta_5 = \log(1.05)$												
Setting 3: $\lambda_{01}(t)=0.4, \lambda_{02}(t)=0.2, \beta_1 = \beta_3 = \log(1.5), \beta_2 = \beta_4 = -\log(1.5), \beta_5 = -\log(1.05)$												
Setting 4: $\lambda_{01}(t)=0.4, \lambda_{02}(t)=0.2, \beta_1 = \beta_3 = \log(2.5), \beta_2 = \beta_4 = -\log(2.5), \beta_5 = -\log(1.05)$												

Table 2
Analysis of SRTR data using proposed methods: Estimated 10-year mean graft survival time for first and second kidney transplants; t measured in years

Quantity	Estimate	Std. Error	<i>p</i>
$\mu_1(10; 10)$	6.916	0.014	–
$\mu_2(10; 10)$	6.588	0.144	–
(10; 10)	-0.328	0.144	0.023

Table 3
Naive analysis of SRTR data based on model: $\lambda_{i2}(t) = \lambda_0(t) \exp\{\beta'Z_i + \theta\}$

Gap time	$\hat{\theta}$	$SE(\hat{\theta})$	$\exp\{\hat{\theta}\}$	p
1(ref.)	0	0	1	-
2	0.213	0.036	1.237	< 0.0001