

# Non-parametric estimation of gap time survival functions for ordered multivariate failure time data

Douglas E. Schaubel<sup>1,\*</sup>,† and Jianwen Cai<sup>2</sup>

<sup>1</sup>*Department of Biostatistics, University of Michigan, Ann Arbor, MI 48109-2029, U.S.A.*

<sup>2</sup>*Department of Biostatistics, University of North Carolina at Chapel Hill, Chapel Hill, NC 27599-7420, U.S.A.*

## SUMMARY

Times between sequentially ordered events (gap times) are often of interest in biomedical studies. For example, in a cancer study, the gap times from incidence-to-remission and remission-to-recurrence may be examined. Such data are usually subject to right censoring, and within-subject failure times are generally not independent. Statistical challenges in the analysis of the second and subsequent gap times include induced dependent censoring and non-identifiability of the marginal distributions. We propose a non-parametric method for constructing one-sample estimators of conditional gap-time specific survival functions. The estimators are uniformly consistent and, upon standardization, converge weakly to a zero-mean Gaussian process, with a covariance function which can be consistently estimated. Simulation studies reveal that the asymptotic approximations are appropriate for finite samples. Methods for confidence bands are provided. The proposed methods are illustrated on a renal failure data set, where the probabilities of transplant wait-listing and kidney transplantation are of interest. Copyright © 2004 John Wiley & Sons, Ltd.

**KEY WORDS:** confidence bands; cumulative hazard function; empirical processes; induced dependent censoring; inverse weighting; multivariate survival analysis

## 1. INTRODUCTION

In biomedical investigations wherein subjects can experience multiple events and where such events occur in a fixed order, times between consecutive events (hereafter referred to as gap times) are often of interest. The events of concern may be of the same nature (i.e. recurrent events); examples include infections among bone marrow transplant recipients, tumour metastases among cancer patients, and hospitalizations in a health service utilization study. The events may also represent different states of a process, an example being a study of HIV infection, AIDS onset and death. Such studies are often of a fixed length, resulting in right

---

\*Correspondence to: Douglas E. Schaubel, Department of Biostatistics, University of Michigan, Ann Arbor, MI 48109-2029, U.S.A.

†E-mail: deschau@umich.edu

Contract/grant sponsor: National Institute of Health; contract/grant number: R01 HL-57444

censoring. In the context of biomedical studies, it is not usually reasonable to assume that events for a given subject are independent. As discussed by several authors [1–4], lack of within-subject gap time independence results in a form of dependent censoring. For example, the longer a given individual's time until first event, the greater the probability that the time between their first and second event is censored. This phenomenon has been termed 'induced dependent censoring'.

The motivating example for the proposed methods relates to end-stage renal disease, commonly referred to as renal failure, an affliction of increasing concern in North America due to its mortality and health care cost. Generally, after commencing renal replacement therapy (RRT), patients who desire and are deemed medically suitable for transplantation are placed on a transplant waiting list (WL). After being wait-listed, patients are sequentially selected for kidney transplantation (KT) given the availability of a suitable donor organ. Hence, the event sequence of interest is  $RRTI \rightarrow WL \rightarrow KT$ , where RRTI denotes RRT-initiation. Time until wait-listing and time between wait-listing and transplantation are positively correlated. Hence, for studies of fixed length, the time between wait-listing and kidney transplantation is subject to induced dependent censoring. Moreover, the marginal distribution of the time between wait-listing and transplantation is not identifiable unless the study duration exceeds the support of the time until WL distribution. For example, in a five year study,  $WL \rightarrow KT$  time is not even partially observed unless time until wait-listing is less than 5 years. Thus, unless  $RRTI \rightarrow WL$  and  $WL \rightarrow KT$  times are independent, observed and partially observed  $WL \rightarrow KT$  times must be interpreted as being conditional on time until wait-listing being less than the study duration. Since observation of the second gap time ( $WL \rightarrow KT$ ) is conditional on the first gap time ( $RRTI \rightarrow WL$ ) occurring before some specific time, the survival function for the  $WL \rightarrow KT$  gap time is not identifiable (i.e. cannot be estimated using only the observed data). However, as discussed by Lin *et al.* [3], we can identify the survival function of the second gap time, conditional on the first gap time falling in some interval contained by the observation period.

Since the marginal distributions of the second and subsequent gap times are generally not identifiable in the sequentially ordered multivariate failure time setting, it is desirable to estimate meaningful and identifiable conditional distributions related to the gap times. As such, we propose an estimator of a conditional survival function for the second and subsequent gap times; the estimated distribution is that of the  $j$ th gap time, conditional on the  $(j - 1)$ th event occurring prior to some fixed point. For example, in a 5 year study examining the  $RRTI \rightarrow WL \rightarrow KT$  sequence, one could estimate the conditional distribution of wait-list to transplant times, conditional on being wait-listed within 2 years of RRTI; this particular conditional distribution is identifiable up to 3 years post-wait-listing.

The problem of estimating gap time distributions in the context of correlated ordered failure time data has received much attention recently. Visser [5] considered non-parametric estimation in the discrete time setting. Huang and Louis [6] developed general methods for marked survival data, which could be applied to the analysis of gap times. Wang [2] described the issues of induced dependent censoring and identifiability. Wang and Wells [1] proposed a product-limit type estimator for the second gap time, while Lin *et al.* [3] and very recently van der Laan *et al.* [7] developed non-parametric estimators of the multivariate gap time distribution function. Wang and Chang [8] and Peña *et al.* [9] consider estimation of gap time survival functions when the number of (recurrent) events is considered random.

Despite the corresponding theoretical developments, the problems of induced dependent censoring and, in particular, non-identifiability are not well known among practitioners. In certain instances, the identifiability issue has been avoided by imposing unrealistic and unnatural assumptions on the censoring mechanism. In other cases, the issue has been ignored or not described in sufficient detail for most practitioners to readily interpret. Published analyses which explicitly take account of the key issues in gap time modelling are currently quite sparse. To fill this apparent void, in this article, we describe the issues of induced dependent censoring and identifiability in detail; propose a computationally simple estimator, as an alternative to those in the existing literature; propose confidence bands useful for studying the time interval of interest simultaneously; and, provide an illustrative real-data example based on an important health care issue. The proposed estimator is based on an adjusted version of the Nelson-Aalen [10, 11] cumulative hazard estimator. Adjustment for induced dependent censoring is accomplished by an inverse weighting technique, similar in spirit to that proposed by Robins and Rotnitzky [12]. We also present a method for estimating confidence bands for the conditional survival function based on a technique analogous to that of Lin *et al.* [13].

The remainder of this article is organized as follows. In Section 2, we introduce requisite notation, formalize the statistical issues pertaining to the analysis of times between events, and present the proposed methods and asymptotic properties of the proposed estimators. Simulation studies are described in Section 3. In Section 4, the proposed methods are applied to the analysis of renal failure data. Concluding remarks are contained in Section 5.

## 2. METHODS

Suppose that there are  $n$  subjects under observation and  $J$  ordered failure times of interest. For subject  $i$ , let  $T_{i,j}$  denote the *total* time (i.e. time under observation) of the  $j$ th failure. Let  $C_i$  denote censoring time; we make the standard assumption that  $C_i$  is independent of the failure (*total*) times:  $\{T_{i,1}, \dots, T_{i,J}\}$ . We set  $\tau_c = \sup_t \{P(C_i \geq t) > 0\}$ . Let the failure indicators and observed total times be represented by  $\Delta_{ij} = I(T_{i,j} < C_i)$  and  $X_{i,j} = T_{i,j} \wedge C_i$ , respectively, where  $I(A) = 1$  when the event  $A$  occurs and 0 otherwise and  $a \wedge b = \min\{a, b\}$ . Times between consecutive events are represented by  $\tilde{T}_{ij} = T_{i,j} - T_{i,j-1}$ , where  $T_{i,0} \equiv 0$ , with corresponding observed gap times  $\tilde{X}_{ij} = \tilde{T}_{ij} \wedge \tilde{C}_{ij}$ , gap censoring times  $\tilde{C}_{ij} = C_i - X_{i,j-1}$  and risk set indicators  $Y_{ij}(s) = I(\tilde{X}_{ij} \geq s)$ . The underlying counting processes are given by:  $N_{ij}^*(t) = I(\tilde{T}_{ij} \leq t) = \int_0^t dN_{ij}^*(s)$ , with their observed counterparts  $N_{ij}(t) = I(\tilde{X}_{ij} \leq t, \Delta_{ij} = 1) = \int_0^t dN_{ij}(s)$ .

Of interest are the gap-time-specific survival functions,  $P(\tilde{T}_{ij} > t)$ , and the first concern is identifiability. For  $\tilde{T}_{i1} = T_{i,1}$ , we have the familiar constraint that we can only estimate  $P(\tilde{T}_{i1} > t)$  for  $t \in [0, \tau_c]$ . For  $j \geq 2$ ,  $P(\tilde{T}_{ij} > t)$  is generally not identifiable for  $t > 0$  unless the support of  $T_{i,j-1}$  is contained by the  $[0, \tau_c]$  interval. The challenge then is to find a quantity which is both interpretable and of general interest to investigators, and which is estimable. For  $j \geq 2$ , we propose estimating the conditional survival function,  $S_j(t; t_{j-1}) \equiv P(\tilde{T}_{ij} > t | T_{i,j-1} \leq t_{j-1})$ , which is identifiable for  $t \in [0, \tau_j]$ , with  $t_{j-1} + \tau_j \leq \tau_c$ . Typically, in the interests of maximizing the use of available data, one would choose  $\{t_{j-1}, \tau_j\}$  pairs such that  $t_{j-1} + \tau_j = \tau_c$  for  $j = 1, \dots, J$ , with  $t_0 = 0$ . We propose estimating  $S_j(t; t_{j-1})$  through its corresponding conditional cumulative hazard function,  $\Lambda_j(t; t_{j-1})$ , where  $S_j(t; t_{j-1}) = e^{-\Lambda_j(t; t_{j-1})}$ . Note that the estimated distribution functions corresponding to our estimators are not subject to negative

mass, which is a problem with many survival function estimators that depend on estimation of the joint survival function.

The next task is to deal with the induced dependent censoring, which arises from the fact that  $\tilde{C}_{ij}$ , the censoring time associated with the  $j$ th gap time, is functionally related to previous event times. For example, the greater the value of  $T_{i,1}$ , the greater the probability that  $\tilde{T}_{i2}$  is censored, unless  $T_{i1}$  and  $\tilde{T}_{i2}$  are independent. Thus, among the second gap times which are observed, there will be a disproportionate number of shorter times, relative to the setting where censoring was not present, indicating that standard methods of estimating the survival function would be biased.

Since  $T_{i,1}$  is not subject to induced dependent censoring, we can estimate  $S_1(t)$  through  $\hat{S}_{1:n}(t) = e^{-\hat{\Lambda}_{1:n}(t)}$ , where  $\hat{\Lambda}_{1:n}(t)$  is the familiar Nelson–Aalen cumulative hazard estimator,

$$\hat{\Lambda}_{1:n}(t) = n^{-1} \sum_{i=1}^n \int_0^t \frac{N_{i1}(ds)}{R_1(s)}$$

where  $R_1(s) = n^{-1} \sum_{i=1}^n Y_{i1}(s)$ . It has been shown that (i)  $\hat{\Lambda}_{1:n}(t)$  converges almost surely to  $\Lambda_1(t)$  uniformly in  $t \in [0, \tau_c]$  (ii)  $n^{1/2}(\hat{\Lambda}_{1:n}(t) - \Lambda_1(t))$  converges to a zero-mean normal distribution for fixed  $t \in [0, \tau_c]$  (iii)  $\{n^{1/2}(\hat{\Lambda}_{1:n}(t) - \Lambda_1(t)); t \in [0, \tau_c]\}$  converges weakly to a zero-mean Gaussian process. Since the Nelson–Aalen estimator is inappropriate when censoring is dependent, another estimator of  $\Lambda_j(t; t_{j-1})$  is required for  $j \geq 2$ . Since there are no issues of identifiability or induced dependent censoring for  $j = 1$  (as there are no preceding failure times), we focus on the second and subsequent gap times for the remainder of this article.

If there were no censoring, such that all  $J$  event times were observed for every subject, then the cumulative hazard for  $(\tilde{T}_{ij}|T_{j-1} \leq t_{j-1})$  could be estimated by

$$\hat{\Lambda}_{j:n}^*(t) = n^{-1} \sum_{i=1}^n \int_0^t \frac{N_{ij}^*(ds; t_{j-1})}{R_j^*(s)} \tag{1}$$

where  $R_j^*(s) = n^{-1} \sum_{i=1}^n I(\tilde{T}_{ij} \geq s, T_{i,j-1} \leq t_{j-1})$  and  $N_{ij}^*(s; t_{j-1}) = I(\tilde{T}_{ij} \leq s, T_{i,j-1} \leq t_{j-1})$ . Set  $G(t) = P(C_i \geq t)$  and note that:

$$E[G(s + T_{i,j-1})^{-I(j \geq 2)} N_{ij}(ds; t_{j-1}) | T_{i,j-1} \leq t_{j-1}] = E[N_{ij}^*(ds; t_{j-1}) | T_{i,j-1} \leq t_{j-1}]$$

$$E[G(s + T_{i,j-1})^{-I(j \geq 2)} Y_{ij}(s; t_{j-1}) | T_{i,j-1} \leq t_{j-1}] = E[I(\tilde{T}_{ij} \geq s) | T_{i,j-1} \leq t_{j-1}]$$

where  $N_{ij}(s; t_{j-1}) = I(\tilde{X}_{ij} \leq s, \Delta_{ij} = 1, T_{i,j-1} \leq t_{j-1})$  and  $Y_{ij}(s; t_{j-1}) = I(\tilde{X}_{ij} \geq s, T_{i,j-1} \leq t_{j-1})$ . Both of the immediately preceding equalities can be shown using conditional expectations. By replacing potentially unobservable random variables in (1) with consistent estimators of quantities with the same conditional expectation, we arrive at the proposed estimators:

$$\hat{\Lambda}_{j:n}^G(t; t_{j-1}) = n^{-1} \sum_{i=1}^n \int_0^t \frac{\hat{W}_{ij}(s)}{R_j^G(s)} N_{ij}(ds; t_{j-1}) \tag{2}$$

where  $W_{ij}(s) = Y_{ij}(s; t_{j-1})G(s + T_{i,j-1})^{-I(j \geq 2)}$ ,  $\hat{W}_{ij}(s) = Y_{ij}(s; t_{j-1})\hat{G}_n(s + T_{i,j-1})^{-I(j \geq 2)}$ ,  $\hat{G}_n(t)$  is the Kaplan–Meier [14] estimator of  $G(t)$  based on  $\{X_{i,j}, 1 - \Delta_{i,j}\}_{i=1}^n$ , and  $R_j^G(s) =$

$n^{-1} \sum_{i=1}^n \hat{W}_{ij}(s)$ . The conditional survival function can then be estimated by  $\hat{S}_{j:n}^G(t; t_{j-1}) = e^{-\hat{\Lambda}_{j:n}^G(t; t_{j-1})}$  for  $t \in [0, \tau_j]$ . The conditional survival function estimator is bounded by 0 and 1 and is monotone in  $t$ , for  $t \in [0, \tau_j]$ . Note that  $\hat{S}_{j:n}^G(t; t_{j-1})$  need not be monotone in  $t_1$ . For example, if  $S_2(t; t_1)$  is estimated for multiple values of  $t_1$ , the ordering of the conditional survival functions at equal values of  $t$ , based on different values of  $t_1$ , will depend on the nature of the correlation between  $T_{i,1}$  and  $T_{i,2}$ .

We now list the essential asymptotic results pertaining to the proposed methods, with proofs outlined in the Appendix. We assume that the observable data,  $\{\tilde{X}_{ij}, \Delta_{ij}\}_{j=1}^J$  for  $i = 1, \dots, n$ , arise from an independently and identically distributed sample. We define  $M_{ij}(t; t_{j-1}) = N_{ij}(t; t_{j-1}) - \int_0^t Y_{ij}(s; t_{j-1}) \Lambda_j(ds; t_{j-1})$ , and the following quantities pertaining to the censoring distribution:  $N_i^C(t) = I(X_{i,J} \leq t, \Delta_{iJ} = 0)$ ,  $Y_i(s) = I(X_{i,J} \geq s)$ ,  $M_i^C(t) = N_i^C(t) - \int_0^t Y_i(s) \Lambda^C(ds)$ , and  $R_C(s) = n^{-1} \sum_{i=1}^n Y_i(s)$ . First, we consider the limiting properties of  $\hat{\Lambda}_{j:n}^G(t; t_{j-1})$ .

It can be shown using results from empirical processes that  $\hat{\Lambda}_{j:n}^G(t; t_{j-1})$  converges almost surely to  $\Lambda_j(t; t_{j-1})$  uniformly in  $t \in [0, \tau_j]$ , and  $n^{1/2} \{ \hat{\Lambda}_{j:n}^G(t; t_{j-1}) - \Lambda_j(t; t_{j-1}) \}$  converges weakly to a zero-mean Gaussian process with a covariance function which can be consistently estimated by:  $\hat{\sigma}_{j:n}(s, t; t_{j-1}) = n^{-1} \sum_{i=1}^n \hat{\xi}_{ij}(s) \hat{\xi}_{ij}(t)$ , where

$$\hat{\xi}_{ij}(t) = \int_0^t \frac{\hat{W}_{ij}(s)}{R_j^G(s)} \hat{M}_{ij}(ds; t_{j-1}) + I(j \geq 2) \int_0^\tau \frac{\hat{q}_{j:n}(s, t)}{R_C(s)} \hat{M}_i^C(ds) \tag{3}$$

with

$$\hat{q}_{j:n}(s, t) = n^{-1} \sum_{\ell=1}^n \int_0^t \frac{I(s \leq r + T_{\ell j-1})}{R_j^G(r)} \hat{W}_{\ell j}(r) \hat{M}_{\ell j}(dr; t_{j-1}) \tag{4}$$

$\hat{M}_{i,j}(t; t_{j-1}) = N_{ij}(t; t_{j-1}) - \int_0^t Y_{ij}(s; t_{j-1}) \hat{\Lambda}_{j:n}^G(ds; t_{j-1})$ ,  $\hat{M}_i^C(t) = N_i^C(t) - \int_0^t Y_i(s) \hat{\Lambda}_n^C(ds)$  and  $\hat{\Lambda}_n^C(t)$  is the Nelson–Aalen estimator of  $\Lambda^C(t)$ . The essence of the proof of the normality result is expressing  $n^{1/2} \{ \hat{\Lambda}_{j:n}^G(t; t_{j-1}) - \Lambda_j(t; t_{j-1}) \}$  asymptotically as a normalized sum of independent and identically distributed random variates; an outline is provided in the Appendix.

The normality result can be used to construct confidence intervals and regions for sets of points. However, it is often of interest to study the behaviour of  $\tilde{T}_{ij}$  across all  $t \in [0, \tau_j]$  simultaneously, and it can be shown that a  $100(1 - \alpha)\%$  confidence band for  $S_j(t; t_{j-1})$  is given by

$$\exp \left\{ - \exp \left\{ \pm n^{-1/2} \frac{\hat{\sigma}_{j:n}(t, t; t_{j-1})^{1/2}}{\hat{\Lambda}_{j:n}^G(t; t_{j-1})} \kappa_{j,\alpha}^* \right\} \right\}$$

where  $\kappa_{j,\alpha}^*$  is the quantity satisfying:

$$P \left( \sup_{t \in [0, \tau_j]} |\hat{\kappa}_{j:n}^*(t)| > \kappa_{j,\alpha}^* \right) = \alpha$$

with  $\hat{\kappa}_{j:n}^*(t) = n^{-1/2} \hat{\sigma}_{j:n}(t, t; t_{j-1})^{-1/2} \sum_{i=1}^n Z_i \hat{\xi}_{ij}(t)$ , and  $\{Z_i\}_{i=1}^n$  are independent standard normal variates such that  $Z_i \perp \hat{\xi}_{ij}(t)$ . To operationalize this result, the observed data are held fixed, a large number of  $\{Z_i\}_{i=1}^n$  are generated, and the appropriate quantile is estimated empirically. Justification for this procedure rests on the concept of conditional weak convergence [15], as outlined in the Appendix.

### 3. SIMULATION STUDY

Simulation studies were conducted to assess the finite sample properties of the proposed estimators. First, for each subject,  $Q_i$  was generated from the Positive Stable distribution [16] with Laplace transform  $E[e^{-t^{Q_i}}] = \exp\{-t^\theta\}$ , using methods described by Chambers *et al.* [17]. Next,  $J = 2$  gap times were generated through the conditional distribution  $P(\tilde{T}_{ij} > t | Q_i) = \exp\{-Q_i \lambda_j t\}$ , which implies the following survival functions:

$$P(T_{i,1} > t) = \exp\{-(\lambda_1 t)^\theta\}$$

$$P(\tilde{T}_{i2} > t | T_{i1} \leq t_1) = \frac{\exp\{-(\lambda_2 t)^\theta\} - \exp\{-(\lambda_2 t + \lambda_1 t_1)^\theta\}}{1 - \exp\{-(\lambda_1 t_1)^\theta\}}$$

We chose  $\theta = 1, 0.75, 0.5$ , corresponding to Kendall's Tau ( $T_{i,1}, \tilde{T}_{i2}$ ) rank correlation,  $1 - \theta$ , of 0, 0.25 and 0.5, respectively. The hazard function for the first failure time was fixed at  $\lambda_1(t) = 0.5$ , while that for  $\tilde{T}_{i2}$  was set to  $\lambda_2(t) = 0.25, 0.5, 0.75$ . Censoring times followed the Uniform (0,10) distribution, mimicking a study of duration 10 with recruitment distributed randomly across the observation period. The proportion censored ranged from 20 to 40 per cent. We chose  $t_1 = 4$  (hence,  $\tau_2 = 6$ ) and  $t_1 = 6$  ( $\tau_2 = 4$ ). Simulated samples were of size  $n = 100$  and 200. One thousand simulations were performed for each data configuration.

In Table I, the bias (denoted by  $b(\cdot)$ ) of the proposed conditional survival function estimator is examined for  $t_1 = 4$ . In all cases examined,  $\hat{S}_{2:n}^G(t; t_1)$  is approximately unbiased. Similar results are found in Table II, where  $t_1 = 6$ . The finite-sample accuracy of the asymptotic distributional approximation for the proposed estimators is evaluated in Table III ( $t_1 = 4$ ) and Table IV ( $t_1 = 6$ ) using the same set of data configurations as Tables I–II. Generally, the average estimated standard error, denoted by  $ASE_{2:n}^G$ , very closely approximated the empirical standard deviation,  $ESD_{2:n}^G$ , with empirical coverage probabilities for the survival function close to the nominal level of 0.95. Note that confidence intervals in Tables III and IV were based on the assumption that  $\log(\hat{\Lambda}_{2:n}^G(t; t_1))$  was normally distributed. Hence, point-wise 95 per cent confidence intervals for  $\log(\Lambda_2(t; t_1))$  were computed through the Delta-method [18] as

$$\log(\hat{\Lambda}_{2:n}^G(t; t_1)) \pm 1.96 \times n^{-1/2} \frac{\hat{\sigma}_{2:n}(t, t; t_1)^{1/2}}{\hat{\Lambda}_{2:n}^G(t; t_1)}$$

This transformation resulted in a noteworthy improvement in empirical coverage probability, relative to the interval resulting from the assumption that  $n^{1/2}\{\hat{\Lambda}_{2:n}^G(t; t_1) - \Lambda_2(t; t_1)\}$  itself was normally distributed (data not shown).

Table I. Simulation Results: Bias, denoted by  $b(\cdot)$ , of proposed survival function estimator for  $\tilde{T}_{12}$ . Gap times were generated such that  $P(\tilde{T}_{ij} > t | Q_i) = \exp\{-Q_i \lambda_j t\}$ , with  $Q_i$  generated from Positive Stable distribution with Laplace transform  $E[e^{-tQ_i}] = \exp\{-t^\theta\}$ . One thousand replicates were simulated per data configuration.

$\lambda_2(t)$	$\theta$	$t$	$S_2(t; t_1)$	$t_1 = 4, n = 100$		$t_1 = 4, n = 200$	
				$\hat{S}_{2:n}^G(t; t_1)$	$b(\hat{S}_{2:n}^G(t; t_1))$	$\hat{S}_{2:n}^G(t; t_1)$	$b(\hat{S}_{2:n}^G(t; t_1))$
0.25	1.00	1.0	0.779	0.778	0.000	0.778	-0.001
		2.0	0.607	0.605	-0.002	0.604	-0.002
		3.0	0.472	0.471	-0.001	0.471	-0.002
	0.75	1.0	0.667	0.666	-0.001	0.666	-0.001
		2.0	0.510	0.510	0.000	0.508	-0.002
		3.0	0.404	0.403	-0.001	0.403	0.000
	0.50	1.0	0.507	0.505	-0.001	0.507	0.000
		2.0	0.380	0.378	-0.002	0.381	0.001
		3.0	0.304	0.307	0.003	0.306	0.002
0.50	1.00	1.0	0.607	0.607	0.000	0.606	0.000
		2.0	0.368	0.372	0.004	0.367	-0.001
		3.0	0.223	0.225	0.002	0.223	0.000
	0.75	1.0	0.510	0.509	-0.001	0.510	0.000
		2.0	0.326	0.324	-0.002	0.327	0.001
		3.0	0.222	0.222	0.000	0.221	-0.001
	0.50	1.0	0.380	0.380	0.000	0.380	0.001
		2.0	0.252	0.253	0.001	0.255	0.003
		3.0	0.185	0.197	0.012	0.187	0.002
0.75	1.00	1.0	0.472	0.474	0.001	0.474	0.001
		2.0	0.223	0.224	0.001	0.224	0.001
		3.0	0.105	0.115	0.010	0.106	0.001
	0.75	1.0	0.404	0.406	0.002	0.403	-0.001
		2.0	0.222	0.222	0.000	0.220	-0.001
		3.0	0.132	0.140	0.008	0.131	-0.001
	0.50	1.0	0.304	0.303	-0.001	0.304	0.000
		2.0	0.185	0.187	0.002	0.184	0.000
		3.0	0.127	0.142	0.015	0.128	0.001

#### 4. APPLICATION

We applied the proposed methods to a study of renal replacement therapy (RRT) patients using data obtained from the multiple organ retrieval and exchange (MORE) programme of Ontario. Patients ( $n = 5356$ ) included those who underwent RRT initiation (RRTI) in Ontario between 1 January 1990 and 31 December 1994. Each patient's follow-up began at RRTI, and two gap times were of interest: time from RRTI to transplant wait-listing (WL), time from WL to kidney transplantation (KT). Patients were followed until the earliest of the date of transplant, death, loss to follow-up or the conclusion of the observation period (31 December 1994). In total, 1095 patients were wait-listed for transplantation, 517 of whom received a

Table II. Simulation Results: Bias, denoted by  $b(\cdot)$ , of proposed survival function estimator for  $\tilde{T}_{12}$ . Gap times were generated such that  $P(\tilde{T}_{i,j} > t | Q_i) = \exp\{-Q_i \lambda_j t\}$ , with  $Q_i$  generated from Positive Stable distribution with Laplace transform  $E[e^{-t^{Q_i}}] = \exp\{-t^\theta\}$ . One thousand replicates were simulated per data configuration.

$\lambda_2(t)$	$\theta$	$t$	$S_2(t; t_1)$	$t_1 = 6, n = 100$		$t_1 = 6, n = 200$	
				$\hat{S}_{2:n}^G(t; t_1)$	$b(\hat{S}_{2:n}^G(t; t_1))$	$\hat{S}_{2:n}^G(t; t_1)$	$b(\hat{S}_{2:n}^G(t; t_1))$
0.25	1.00	1.0	0.779	0.778	-0.001	0.778	-0.001
		2.0	0.607	0.607	0.000	0.605	-0.001
		3.0	0.472	0.475	0.003	0.473	0.001
	0.75	1.0	0.683	0.686	0.003	0.682	-0.001
		2.0	0.528	0.531	0.003	0.528	-0.001
		3.0	0.422	0.431	0.009	0.422	0.000
	0.50	1.0	0.537	0.535	-0.002	0.535	-0.001
		2.0	0.412	0.411	-0.001	0.410	-0.002
		3.0	0.336	0.347	0.011	0.335	-0.001
0.50	1.00	1.0	0.607	0.608	0.001	0.605	-0.001
		2.0	0.368	0.366	-0.002	0.367	-0.001
		3.0	0.223	0.225	0.002	0.220	-0.003
	0.75	1.0	0.528	0.526	-0.002	0.528	-0.001
		2.0	0.344	0.342	-0.002	0.343	-0.001
		3.0	0.237	0.245	0.009	0.238	0.002
	0.50	1.0	0.412	0.415	0.003	0.411	-0.001
		2.0	0.283	0.286	0.003	0.280	-0.002
		3.0	0.211	0.228	0.017	0.214	0.002
0.75	1.00	1.0	0.472	0.472	0.000	0.473	0.001
		2.0	0.223	0.224	0.001	0.223	0.000
		3.0	0.105	0.117	0.012	0.107	0.002
	0.75	1.0	0.422	0.424	0.002	0.422	0.000
		2.0	0.237	0.238	0.001	0.234	-0.003
		3.0	0.143	0.157	0.014	0.143	0.000
	0.50	1.0	0.336	0.333	-0.002	0.336	0.000
		2.0	0.211	0.213	0.001	0.211	0.000
		3.0	0.148	0.169	0.021	0.153	0.005

renal transplant; 1547 patients died, and 26 were lost to follow-up. Our objective was to provide an overall estimate of the probability of (i) being wait-listed and (ii) receiving a kidney transplant among patients placed on the wait-list within a certain number of years of RRTI.

It is unrealistic to assume that a patient's RRTI  $\rightarrow$  WL and WL  $\rightarrow$  KT times are independent. Longer time between RRTI and wait-listing is positively associated with longer time between wait-listing and transplantation, as certain common factors influence both gap times. Due to the lack of statistical independence between the RRTI  $\rightarrow$  WL and WL  $\rightarrow$  KT gap times, the WL  $\rightarrow$  KT gap time is subject to induced dependent censoring; i.e. greater time until WL implies greater probability that the WT  $\rightarrow$  KT time is censored. Moreover, the marginal probability that a wait-listed patient is transplanted is not identifiable.



Table III. Simulation Results: Accuracy of asymptotic distributional approximations for proposed survival function estimator for  $\hat{T}_{i2}$ . Gap times were generated such that  $P(\hat{T}_{ij} > t | Q_i) = \exp\{-Q_i \lambda_j t\}$ , with  $Q_i$  generated from Positive Stable distribution with Laplace transform  $E[e^{-tQ_i}] = \exp\{-t^\theta\}$ . ASE denotes average approximated standard error, ESD denotes empirical standard deviation, CP denotes empirical 95 per cent coverage probability, based on 1000 simulations.

$\lambda_2(t)$	$\theta$	$t$	$t_1 = 4, n = 100$			$t_1 = 4, n = 200$		
			$ASE_{2:n}^G$	$ESD_{2:n}^G$	$CP_{2:n}^G$	$ASE_{2:n}^G$	$ESD_{2:n}^G$	$CP_{2:n}^G$
0.25	1.00	1.0	0.063	0.065	0.952	0.045	0.047	0.938
		2.0	0.100	0.104	0.943	0.071	0.073	0.944
		3.0	0.137	0.146	0.945	0.098	0.098	0.957
	0.75	1.0	0.084	0.087	0.945	0.060	0.059	0.959
		2.0	0.121	0.125	0.946	0.087	0.088	0.951
		3.0	0.158	0.165	0.935	0.112	0.114	0.949
	0.50	1.0	0.119	0.123	0.942	0.084	0.084	0.953
		2.0	0.160	0.169	0.945	0.113	0.117	0.946
		3.0	0.195	0.201	0.950	0.140	0.146	0.943
0.50	1.00	1.0	0.096	0.101	0.937	0.068	0.070	0.947
		2.0	0.161	0.166	0.945	0.116	0.118	0.952
		3.0	0.241	0.240	0.965	0.174	0.181	0.945
	0.75	1.0	0.118	0.119	0.955	0.084	0.083	0.957
		2.0	0.181	0.188	0.942	0.128	0.129	0.952
		3.0	0.245	0.253	0.955	0.177	0.181	0.944
	0.50	1.0	0.155	0.165	0.928	0.110	0.115	0.941
		2.0	0.217	0.232	0.943	0.155	0.156	0.947
		3.0	0.263	0.262	0.945	0.199	0.203	0.950
0.75	1.00	1.0	0.126	0.124	0.963	0.090	0.088	0.958
		2.0	0.231	0.250	0.942	0.166	0.175	0.933
		3.0	0.352	0.330	0.949	0.271	0.274	0.946
	0.75	1.0	0.146	0.150	0.952	0.104	0.108	0.942
		2.0	0.235	0.246	0.949	0.170	0.173	0.957
		3.0	0.322	0.316	0.943	0.245	0.256	0.942
	0.50	1.0	0.185	0.189	0.942	0.132	0.139	0.942
		2.0	0.265	0.274	0.945	0.193	0.203	0.940
		3.0	0.320	0.305	0.936	0.251	0.259	0.942

In this analysis, we seek to estimate  $S_1(t)$ , the marginal survival function for wait-listing, and  $S_2(t; t_1)$ , the conditional survival function for transplantation given wait-listing within  $t_1$  years. Regarding  $S_2(t; t_1)$ , we must choose  $t_1$ , respecting that  $\tau_2 = \tau - t_1$ , where  $\tau = 5$  years. We can select several  $t_1$ 's, such that an array of different conditional survival functions are estimated. Or, a particular choice of  $t_1$  may be suggested by the data at hand. For example, if it is found that the  $\hat{S}_1(t)$  function flattens out at  $t^*$ , then estimation of  $S_2(t; t^*)$  may be appealing. A balance will need to be struck with respect to the competing objectives of having  $\tau_2$  such that the  $\hat{S}_2(t; t_1)$  curve extends across a sufficiently long interval, and choosing  $t_1$  large enough that a sufficient number of patients are included in the estimation of  $\hat{S}_2(t; t_1)$ . Alternatively, an option is to merely set  $t_2 = \tau/2$ .

Table IV. Simulation Results: Accuracy of asymptotic distributional approximations for proposed survival function estimator for  $\tilde{T}_{i,2}$ . Gap times were generated such that  $P(\tilde{T}_{i,j} > t | Q_i) = \exp\{-Q_i \lambda_j t\}$ , with  $Q_i$  generated from Positive Stable distribution with Laplace transform  $E[e^{-tQ_i}] = \exp\{-t^\theta\}$ . ASE denotes average approximated standard error, ESD denotes empirical standard deviation, CP denotes empirical 95 per cent coverage probability, based on 1000 simulations.

$\lambda_2(t)$	$\theta$	$t$	$t_1 = 6, n = 100$			$t_1 = 6, n = 200$		
			$ASE_{2:n}^G$	$ESD_{2:n}^G$	$CP_{2:n}^G$	$ASE_{2:n}^G$	$ESD_{2:n}^G$	$CP_{2:n}^G$
0.25	1.00	1.0	0.062	0.065	0.949	0.044	0.045	0.942
		2.0	0.098	0.101	0.949	0.070	0.073	0.942
		3.0	0.135	0.137	0.942	0.097	0.100	0.945
	0.75	1.0	0.077	0.080	0.951	0.056	0.058	0.939
		2.0	0.113	0.119	0.952	0.082	0.083	0.955
		3.0	0.146	0.150	0.953	0.107	0.110	0.948
	0.50	1.0	0.109	0.113	0.938	0.077	0.080	0.939
		2.0	0.146	0.155	0.936	0.105	0.111	0.945
		3.0	0.175	0.184	0.932	0.129	0.138	0.942
0.50	1.00	1.0	0.094	0.096	0.942	0.068	0.067	0.952
		2.0	0.161	0.168	0.950	0.115	0.118	0.950
		3.0	0.236	0.251	0.938	0.174	0.181	0.950
	0.75	1.0	0.111	0.114	0.945	0.079	0.081	0.942
		2.0	0.171	0.180	0.944	0.122	0.130	0.936
		3.0	0.226	0.225	0.955	0.168	0.176	0.939
	0.50	1.0	0.141	0.146	0.936	0.101	0.111	0.926
		2.0	0.198	0.215	0.932	0.144	0.159	0.918
		3.0	0.239	0.252	0.925	0.182	0.194	0.928
0.75	1.00	1.0	0.124	0.123	0.959	0.088	0.087	0.948
		2.0	0.227	0.238	0.949	0.164	0.162	0.954
		3.0	0.341	0.323	0.944	0.265	0.270	0.946
	0.75	1.0	0.137	0.143	0.947	0.098	0.101	0.943
		2.0	0.223	0.245	0.928	0.162	0.169	0.946
		3.0	0.299	0.303	0.938	0.233	0.249	0.932
	0.50	1.0	0.170	0.178	0.938	0.120	0.124	0.944
		2.0	0.243	0.263	0.938	0.176	0.181	0.943
		3.0	0.286	0.278	0.940	0.226	0.231	0.947

Table V lists estimates of  $F_1(t) \equiv 1 - S_1(t)$ , the probability of being wait-listed within  $t$  years. Wait-list probability appears to plateau 2–3 years after RRTI. Also listed in Table V are estimates of  $F_2(t; t_1) \equiv 1 - S_2(t; t_1)$ , the probability of receiving a transplant, within  $t$  years, given wait-listing occurred within  $t_1$  years, for various values of  $t_1$ . The dependence between  $\tilde{T}_{i,1}$  and  $\tilde{T}_{i,2}$  is demonstrated by the discrepancies in the  $\hat{F}_{2:n}^G(t; t_1)$  by  $t_1$ . It appears that greater RRTI  $\rightarrow$  WL times are associated with greater WL  $\rightarrow$  KT times.

In Figure 1, we display point estimates and 95% confidence bands for  $F_1(t)$  (left panel) and  $F_2(t; t_1 = 2)$  (right panel). Given that wait-listing was within 2 years of RRTI, the estimated median time until transplant is 3 years after being wait-listed. From the corresponding plot, it is apparent that transplant probability continues to increase steadily along  $t \in [0, 3]$ . This

Table V. Estimated probability of being wait-listed within  $t$  years,  $\hat{F}_{1:n}(t)$ ; and probability of receiving a kidney transplant within  $t$  years of being wait-listed, given that the patient was wait-listed within  $t_1$  years,  $\hat{F}_{2:n}^G(t; t_1)$ . Based on  $n=5356$  patients initiating renal replacement therapy during 1990–1994 in Ontario, Canada.

$t$	$\hat{F}_{1:n}(t)$	$\hat{F}_{2:n}^G(t; t_1)$		
		$t_1 = 1$	$t_1 = 2$	$t_1 = 3$
0.5	0.07	0.05	0.04	0.03
1.0	0.16	0.17	0.12	0.11
1.5	0.22	0.30	0.23	0.21
2.0	0.25	0.38	0.33	0.29
2.5	0.26	0.47	0.40	—
3.0	0.27	0.54	0.50	—
3.5	0.29	0.63	—	—
4.0	0.30	0.70	—	—
4.5	0.30	—	—	—
5.0	0.30	—	—	—

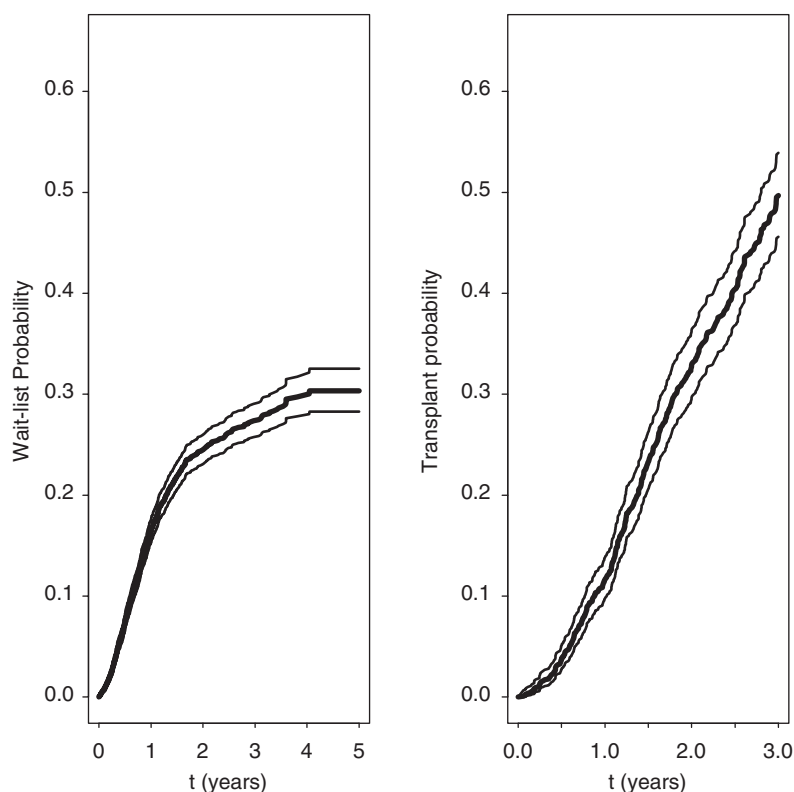


Figure 1. Left panel: Probability of being wait-listed within  $t$  years, and 95 per cent confidence bands. Right panel: Probability of receiving a transplant within  $t$  years (and 95 per cent confidence bands), for  $t \in [0, 3]$ , given that the patient was wait-listed within 2 years. Based on  $n = 5356$  patients initiating renal replacement therapy during 1990–1994 in Ontario, Canada.

indicates that, although most patients who are wait-listed do so in the first 2–3 years post-therapy initiation, time until transplant may be considerably longer than 3 years.

## 5. DISCUSSION

We propose a method for estimating conditional survival functions through their corresponding cumulative hazard functions for the times between consecutive events in sequentially ordered failure time data. The asymptotic distribution of the estimators is derived. Through simulation, the asymptotic approximations appear to be appropriate for finite samples. A method for computing simultaneous confidence bands is provided. We applied the proposed methods to a renal failure data set to estimate the probability of being wait-listed and the probability of being transplanted among patients wait-listed within various time intervals.

In the absence of within-subject gap time independence, the two main challenges facing estimation of gap time distribution functions are induced dependent censoring and identifiability for the second and subsequent gap times. Our method, like those of Wang and Wells [1] and Lin *et al.* [3], essentially adjusts for induced dependent censoring by weighting risk set contributions by the inverse of the probability of remaining uncensored, similar in spirit to Robins and Rotnitzky [12]. The estimator of Wang and Wells takes a product-integral form, and has a very complicated covariance structure; the authors recommend using the bootstrap method to obtain standard errors. In addition, Wang and Wells condition on  $[T_{i1} > t_1]$ , while our proposed method conditions on  $[T_{i1} \leq t_1]$ . Lin *et al.* estimate the joint distribution function using an empirical mean-type estimator which has a closed-form covariance structure. Van der Laan *et al.* [7] recently proposed locally efficient estimators which can accommodate censoring which is dependent on the total times and covariates. Our estimators differ from those of Lin *et al.* [3] and van der Laan *et al.* [7] in that the conditional survival function is estimated directly through a cumulative hazard estimator, rather than through a joint distribution function estimator. An advantage of the proposed techniques, compared to existing methods, is the ease of computing standard errors, which may be important to many practitioners. The formulae are simpler to compute than those of previous methods which derive conditional survival function estimators from the estimated joint and marginal distributions. In addition, none of the previous methods provide techniques for obtaining confidence bands.

With respect to identifiability, for the  $j$ th gap time ( $j \geq 2$ ), our method conditions on the  $(j-1)$ th event time having occurred in the  $[0, t_{j-1}]$  interval. This could be viewed as a limitation, in the sense that the marginal distribution may be of greater interest. However, in terms of interpretation, such conditioning may often make the survival estimators more meaningful to investigators, particularly when the conditioning serves to focus consideration on the subjects of primary interest. For instance, in the renal failure example, it could be argued that subjects not wait-listed in the first 2 years are extremely poor candidates for transplantation, and therefore should not be aggregated with patients with shorter wait-list times in an analysis of WL  $\rightarrow$  KT times. Depending on the length of follow-up and distribution of the total times, it may be possible to choose the  $t_{j-1}$ 's at points after which the marginal distribution functions of the  $T_{i,j-1}$ 's level off. Relating back to the renal failure data, the probability of wait-listing had tapered-off by the 2 year mark, indicating that little was sacrificed by not considering the small fraction of patients wait-listed after the 2 year point.

It is often of interest to compare gap time distributions among groups of subjects. Lin and Ying [19] have proposed a family of non-parametric two-sample tests for differences in gap time distributions. Huang [4] proposed a two-sample inference procedure for the gap time setting and recently developed corresponding regression methods [20], both based on the accelerated failure time model. Extension of the methods in this article to the regression setting will be communicated in a separate report.

APPENDIX

To prove asymptotic normality of the proposed estimator, we begin by defining the functional:

$$\|H_{j:n}(t)\| = \sup_{t \in [0, \tau_j]} |H_{j:n}(t)|$$

for some function  $H$ . We set  $\phi_{j:n}(t) = \{\hat{\Lambda}_{j:n}^G(t; t_{j-1}) - \Lambda_j(t; t_{j-1})\} = \phi_{j1:n}(t) + \phi_{j2:n}(t)$ , where:

$$\phi_{j1:n}(t) = n^{-1} \sum_{i=1}^n \int_0^t \frac{W_{ij}(s)}{R_j^G(s)} M_{ij}(ds; t_{j-1}) \tag{5}$$

$$\phi_{j2:n}(t) = n^{-1} \sum_{i=1}^n \int_0^t \frac{1}{R_j^G(s)} \{\hat{W}_{i,j}(s) - W_{ij}(s)\} M_{ij}(ds; t_{j-1}) \tag{6}$$

By the Triangle Inequality,  $\|\phi_{j:n}(t)\| \leq \|\phi_{j1:n}(t)\| + \|\phi_{j2:n}(t)\|$ . Combining the Uniform Strong Law of Large Numbers [21] and Lemma 1 of Lin *et al.* [22],  $\|n^{-1} \sum_{i=1}^n \int_0^t W_{ij}(s) \{R_j^G(s)^{-1} - r_j(s)^{-1}\} M_{ij}(ds; t_{j-1})\| \xrightarrow{\text{a.s.}} 0$ , where  $r_j(s) = \lim_{n \rightarrow \infty} R_j^G(s)$ . Therefore,  $\phi_{j1:n}(t) = n^{-1} \sum_{i=1}^n \int_0^t W_{ij}(s) r_j(s)^{-1} M_{ij}(ds; t_{j-1}) + o(1)$ , which converges almost surely to 0 uniformly in  $t \in [0, \tau_j]$  by the USLLN. Using results in Fleming and Harrington [23], followed by the Functional Delta Method [24], it can be shown that:

$$n^{1/2} \{\hat{W}_{i,j}(s) - W_{ij}(s)\} = I(j \geq 2) W_{ij}(s) n^{-1/2} \sum_{\ell=1}^n \int_0^{s+T_{\ell,j-1}} \frac{M_{\ell}^C(dr)}{R_C(r)} + o_p(1) \tag{7}$$

Combining (6) and (7) yields:

$$\phi_{j2:n}(t) = I(j \geq 2) n^{-1} \sum_{i=1}^n \int_0^t \frac{q_{j:n}(s, t)}{R_C(s)} M_i^C(ds) + o_p(n^{-1/2})$$

where:

$$q_{j:n}(s, t) = n^{-1} \sum_{\ell=1}^n \int_0^t \frac{I(s \leq r + T_{\ell,j-1})}{R_j^G(r)} W_{\ell j}(r) M_{\ell j}(dr; t_{j-1})$$

Again combining the USLLN with Lemma 1 from Lin *et al.* [22],  $\|n^{-1} \sum_{i=1}^n \int_0^t \{q_{j:n}(s, t) R_C(s)^{-1} - q_j(s, t) r_C(s)^{-1}\} M_i^C(ds)\| \xrightarrow{\text{a.s.}} 0$ , where  $r_C(s) = E[Y_i(s)]$  and  $q_j(s, t) = \lim_{n \rightarrow \infty} q_{j:n}(s, t)$ . Since  $M_i^C(t)$  is a martingale with respect to the filtration  $\mathcal{F}_i^C(t) = \sigma\{Y_i(s), N_i^C(s-) : s \in [0, t]\}$ ,

$\|\phi_{j2:n}(t)\| \xrightarrow{\text{a.s.}} 0$  by standard Martingale theory [25]. Hence  $\|\phi_{j:n}(t)\|$  converges almost surely to 0 as  $n \rightarrow \infty$ . With respect to the finite-dimensional distributions, it can be shown using the theory of empirical processes [22, 26] that  $\|n^{-1/2} \sum_{i=1}^n \int_0^t W_{ij}(s) \{R_j^G(s)^{-1} - r_j(s)^{-1}\} M_{ij}(ds; t_{j-1})\| = o_p(1)$ .

Therefore,

$$n^{1/2} \phi_{j1:n}(t) = n^{-1/2} \sum_{i=1}^n \int_0^t W_{ij}(s) r_j(s)^{-1} M_{ij}(ds; t_{j-1}) + o_p(1)$$

Furthermore, we can show using empirical process results that

$$n^{-1/2} \sum_{i=1}^n \int_0^t \{q_{j:n}(s, t) R_C(s)^{-1} - q_j(s, t) r_C(s)^{-1}\} M_i^C(ds) = o_p(1),$$

to obtain:

$$n^{1/2} \phi_{j2:n}(t) = I(j \geq 2) n^{-1/2} \sum_{i=1}^n \int_0^\tau \frac{q_j(s, t)}{r_C(s)} M_i^C(ds) + o_p(1)$$

since  $q_{j:n}(s, t) \xrightarrow{\text{a.s.}} q_j(s, t)$  for  $s \in [0, \tau_j]$  and  $R_C(s) \xrightarrow{\text{a.s.}} r_C(s)$  for  $s \in [0, \tau_c]$  by the USLLN. Therefore,  $n^{1/2} \phi_{j:n}(t) = n^{-1/2} \sum_{i=1}^n \xi_{ij}(t) + o_p(1)$ , with:

$$\xi_{ij}(t) = \int_0^t \frac{W_{ij}(s)}{r_j(s)} M_{ij}(ds; t_{j-1}) + I(j \geq 2) \int_0^\tau \frac{q_j(s, t)}{r_C(s)} M_i^C(ds)$$

where  $\xi_{1j}(t), \dots, \xi_{nj}(t)$  are asymptotically independent and identically distributed mean-zero random variates. The finite-dimensional distributions of  $n^{1/2} \phi_{j:n}(t)$  converge to that of a zero-mean normal variate by the multivariate central limit theorem. To complete the weak convergence proof, we demonstrate tightness of  $n^{1/2} \phi_{j:n}(t)$ , which follows from tightness of  $n^{1/2} \phi_{j1:n}(t)$  and  $n^{1/2} \phi_{j2:n}(t)$  separately. Tightness of  $n^{1/2} \phi_{j1:n}(t)$  follows from that of each summand in (5), and  $n^{1/2} \int_0^t W_{ij}(s) r_j(s)^{-1} M_{ij}(ds; t_{j-1})$  is tight since it is manageable [21]; manageability follows from the fact that each term is expressible as a product of monotone functions [22, 26]. Since  $n^{1/2} \phi_{j2:n}(t)$  is an  $\mathcal{F}^C(t)$ -martingale asymptotically, it is asymptotically tight [25]. Hence,  $n^{1/2} \phi_{j:n}(t)$  is asymptotically tight, and converges weakly to a zero-mean Gaussian process with covariance function  $E[\xi_{1j}(s) \xi_{1j}(t)]$ , by the Functional Central Limit Theorem [21].

To justify the proposed confidence bands, we begin with the result that  $n^{1/2} \phi_{j:n}(t) = n^{-1/2} \sum_{i=1}^n \xi_{ij}(t) + o_p(1)$ . Applying the Functional Delta Method,

$$n^{1/2} \{\log \hat{\Lambda}_{j:n}^G(t; t_{j-1}) - \log \Lambda_j(t; t_{j-1})\} = n^{-1/2} \Lambda_j(t; t_{j-1})^{-1} \sum_{i=1}^n \xi_{ij}(t) + o_p(1)$$

After standardizing and applying a variance-stabilizing transform,

$$\begin{aligned} \kappa_{j:n}(t) &\equiv n^{1/2} \frac{\Lambda_j(t; t_{j-1})}{\sigma_j(t; t_{j-1})^{1/2}} \{\log \hat{\Lambda}_{j:n}^G(t; t_{j-1}) - \log \Lambda_j(t; t_{j-1})\} \\ &= n^{-1/2} \sigma_j(t; t_{j-1})^{-1/2} \sum_{i=1}^n \xi_{ij}(t) + o_p(1) \end{aligned}$$

Set  $\hat{\kappa}_{j:n}(t) = n^{-1/2} \hat{\sigma}_{j:n}(t, t; t_{j-1})^{-1/2} \sum_{i=1}^n \hat{\xi}_{ij}(t)$ , where  $\hat{\xi}_{ij}(t)$  is given in (3), and note that  $\|\hat{\kappa}_{j:n}(t) - \kappa_{j:n}(t)\| \xrightarrow{a.s.} 0$  using the USLLN along with continuity arguments. Now, set:

$$\hat{\kappa}_{j:n}^*(t) = n^{-1/2} \hat{\sigma}_{j:n}(t, t; t_{j-1})^{-1/2} \sum_{i=1}^n Z_i \hat{\xi}_{ij}(t)$$

where  $Z_i \sim N(0, 1)$  and the  $\{Z_i\}_{i=1}^n$  are distributed independently of the observed data. By the Multiplier Central Limit Theorem [15], the limiting distribution of  $\|\hat{\kappa}_{j:n}^*(t)\|$  is the same as the unconditional limiting distribution of  $\|\hat{\kappa}_{j:n}(t)\|$ . Therefore, computing the empirical quantile  $\kappa_{j:n}^*$  such that  $P(\|\hat{\kappa}_{j:n}^*(t)\| > \kappa_{j:n}^*) = \alpha$  and inverting the resulting confidence band for  $n^{1/2} \Lambda_j(t; t_{j-1}) \log \Lambda_j(t; t_{j-1}) \sigma_j(t, t; t_{j-1})^{-1/2}$  yields a  $100(1 - \alpha)$  per cent confidence band for  $\Lambda_j(t; t_{j-1})$ , which can then be transformed to produce a confidence band for  $S_j(t; t_{j-1})$ .

#### ACKNOWLEDGEMENTS

The authors thank the Multiple Organ Retrieval and Exchange (MORE) Program and David Churchill and Christian Rabbat of McMaster University for providing access to the wait-list and transplant data. We also thank Danyu Lin for his insightful comments on this work, and the reviewers for their many constructive comments on the original version of this manuscript. This research was supported by National Institutes of Health grant R01 HL-57444.

#### REFERENCES

1. Wang W, Wells MT. Nonparametric estimation of successive duration times under dependent censoring. *Biometrika* 1999; **85**:561–572.
2. Wang MC. Gap time bias in incident and prevalent cohorts. *Statistica Sinica* 1999; **9**:999–1010.
3. Lin DY, Sun W, Ying Z. Nonparametric estimation of the gap time distributions for serial events with censored data. *Biometrika* 1999; **86**:59–70.
4. Huang Y. Two-sample multistate accelerated sojourn times model. *Journal of the American Statistical Association* 2000; **95**:619–627.
5. Visser M. Nonparametric estimation of the bivariate survival function with an application to vertically transmitted AIDS. *Biometrika* 1996; **83**:507–518.
6. Huang Y, Louis TA. Nonparametric estimation of the joint distribution of survival time and mark variables. *Biometrika* 1998; **85**:785–798.
7. 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.
8. Wang MC, Chang S-H. Nonparametric estimation of a recurrent survival function. *Journal of the American Statistical Association* 1999; **94**:146–153.
9. Peña EA, Strawderman RL, Hollander M. Nonparametric estimation with recurrent event data. *Journal of the American Statistical Association* 2001; **96**:1299–1315.
10. Nelson W. Theory and applications of hazard plotting for censored failure data. *Technometrics* 1972; **14**: 945–965.
11. Aalen O. Nonparametric inference for a family of counting processes. *Annals of Statistics* 1978; **6**:701–726.
12. Robins JM, Rotnitzky A. Recovery of information and adjustment for dependent censoring using surrogate markers. In *AIDS Epidemiology-Methodological Issues, Proceedings of the Biopharmaceutical Section*, Jewell N, Dietz K, Farewell V (eds). American Statistical Association: Alexandria, VA, 1992; 24–33.
13. Lin DY, Fleming TR, Wei LJ. Confidence bands for survival curves under the proportional hazards model. *Biometrika* 1994; **81**:73–81.
14. Kaplan EL, Meier P. Nonparametric estimation from incomplete samples. *Journal of the American Statistical Association* 1958; **53**:457–481.
15. van der Vaart AW, Wellner JA. *Weak Convergence and Empirical Processes*. Springer: New York, 1996.
16. Kanter M. Stable densities under change of scale and total variation inequalities. *The Annals of Probability* 1975; **3**:697–707.
17. Chambers JM, Mallows CL, Stuck BW. A method for simulating stable random variables. *Journal of the American Statistical Association* 1976; **71**:340–344.

18. Sen PK, Singer JM. *Large Sample Methods in Statistics*. Chapman & Hall: New York, 1993.
19. Lin DY, Ying Z. Nonparametric tests for the gap time distributions of serial events based on censored data. *Biometrics* 2001; **57**:369–375.
20. Huang Y. Censored regression with the multistate accelerated sojourn times model. *Journal of the Royal Statistical Society, Series B* 2002; **64**:17–29.
21. Pollard D. *Empirical Processes: Theory and Applications*. Institute of Mathematical Statistics: Hayward, 1990.
22. Lin DY, Wei LJ, Yang I, Ying Z. Semiparametric regression for the mean and rate functions of recurrent events. *Journal of the Royal Statistical Society, Series B* 2000; **62**:711–730.
23. Fleming TR, Harrington, DP. *Counting Processes and Survival Analysis*. Wiley: New York, 1991.
24. Andersen PK, Borgan O, Gill RD, Keiding N. *Statistical Models Based on Counting Processes*. Springer: New York, 1993.
25. Pollard D. *Convergence of Stochastic Processes*. Springer: New York, 1984.
26. Biliyas Y, Gu M, Ying Z. Towards a general asymptotic theory for the Cox model with staggered entry. *Annals of Statistics* 1997; **25**:662–682.