The Canadian Journal of Statistics

Vol. xx, No. yy, 20??, Pages 1–25

La revue canadienne de statistique

Semiparametric regression methods for temporal processes subject to multiple sources of censoring

Tianyu ZHAN[1]* and Douglas E. SCHAUBEL[2]*

[1]*Department of Biostatistics, University of Michigan, 1415 Washington Hts., Ann Arbor, MI, U.S.A., 48109*

[2]*Department of Biostatistics, Epidemiology, and Informatics, University of Pennsylvania, 423 Guardian Dr., Philadelphia, PA, U.S.A., 19104*

Key words and phrases: Additive hazards model; dependent censoring; inverse weighting; prevalence; semiparametric model; survival analysis; temporal process regression.

MSC 2010: Primary 62N01; secondary 62D05

Abstract: Process regression methodology is underdeveloped relative to the frequency with which pertinent data arise. In this article, the response is a binary indicator process representing the joint event of being alive and remaining in a specific state. The process is indexed by time (e.g., time since diagnosis) and observed continuously. Data of this sort occur frequently in the study of chronic disease. A general area of application involves a recurrent event with non-negligible duration (e.g., hospitalization and associated length of hospital stay) and subject to a terminating event (e.g., death). We propose a semiparametric multiplicative model for the process version of the probability of being alive and in the (transient) state of interest. Under the

proposed methods, the regression parameter is estimated through a procedure that does not require estimating the baseline probability. Unlike the majority of process regression methods, the proposed methods accommodate multiple sources of censoring. In particular, we derive a computationally convenient variant of Inverse Probability of Censoring Weighting based on the additive hazards model. We show that the regression parameter estimator is asymptotically normal, and that the baseline probability function estimator converges to a Gaussian process. Simulations demonstrate that our estimators have good finite sample performance. We apply our method to national end-stage liver disease (ESLD) data. *The Canadian Journal of Statistics* xx: 1–25; 20??

Résumé: Insérer votre résumé ici. We will supply a French abstract for those authors who can't prepare it themselves. *La revue canadienne de statistique* xx: 1–25; 20??

1. INTRODUCTION

In biomedical applications, the response of interest can often be cast as a binary indicator process indexed by time. We consider the setting wherein the indicator at time t takes the value 1 (denoting 'success' in some form) when the patient is alive and in a particular state, and 0 otherwise. Examples include the following: (i) In a study of leukaemia patients, the response could be coded as 1 if the patient is alive and in remission t days following diagnosis, and 0 otherwise. (ii) In a study of morbidity among end-stage renal disease patients, the response at time t equals 1 if the patient is alive-and-not-hospitalized at time t, and 0 other-

* *Author to whom correspondence may be addressed.*

E-mail: douglas.schaubel@pennmedicine.upenn.edu

wise. (iii) An end-stage liver disease (ESLD) patient could be coded as having response 1 if active on the liver transplant waiting list (at time t days after initial waiting list registration), and 0 otherwise. When covariate effects are of chief interest, temporal process regression is a natural way to cast the afore-described data structure. Although several methods amenable to this data structure have been developed in the last 10-15 years (beginning with Fine, Yan, & Kosorok, 2004), few modeling choices are available relative to the frequency with which this data structure arises in practice. In this article, we develop semiparametric process regression methods that can be used to model settings such as (i), (ii), and (iii) above, in a flexible manner and making fewer assumptions regarding the censoring process.

Formalizing the above-described data structure, suppose that (for a hypothetical subject) D represents time of death, and $\mathcal{E}(t)$ is an indicator taking the value 1 if the subject is in the state of interest at follow-up time t (and 0 if not). We can define $A(t) = \mathcal{E}(t)I(D > t)$, such that the survival time in the state of interest, D^A, can be written $D^A = \int_{[0,\infty)} A(t)dt = \int_{[0,D]} \mathcal{E}(t)dt$. This can be considered a special case of quality adjusted survival time (Gelber, Gelman, & Goldhirsch, 1989; Glasziou, Simes, & Gelber, 1990; Zhao & Wang, 2008), where 'quality' is defined as yes versus no (1 versus 0). In this article, we develop methods with the goal of directly analyzing the process, $E[A(t)]$.

The methods we propose are motivated by the end-stage liver disease (ESLD) setting. The preferred therapy for ESLD is deceased-donor liver transplantation. However, due to a shortage of donor livers, medically suitable patients are placed on a waiting list. A wait-listed patient is eligible to receive a transplant only when 'active'; patients may be deactivated for several reasons, most of which are related to a decline in health status which renders the patient at least temporarily unsuitable for transplantation. Hence, keeping the patient active on the waiting list represents a successful outcome, in the sense that the patient not only continues to survive but also remains eligible for the preferred treatment. To date, there has been little study of the probability of remaining active on the waiting list.

The response we consider could be framed as a temporal process, $A(t)$, where t is continuous. In contrast to a counting process, $A(t)$ need not be a non-decreasing function. In the context of our afore-described motivating example, we let $A(t)$ be the indicator of being both alive and active on the transplant waiting list at time t. In several existing temporal process regression methods, the expectation of $A(t)$ is linked to linear components through a continuous link function; for example, Fine, Yan,& Kosorok, 2004), Yan & Fine (2005), and Yan & Huang (2009). This could be viewed as a generalized linear model indexed by time. The regression coefficients $\boldsymbol{\beta}(t)$ could be solved at observed jump points. In our work, we consider a semiparametric model for $E[A(t)]$, where covariates have multiplicative effects on a completely unspecified probability function

indexed by time.

In this manuscript, we develop semiparametric regression methods for a temporal process subject to dependent censoring. Two types of censoring are considered. Specifically, we let C_1 denote censoring that is independent conditional on external covariates. Dependent censoring, denoted by C_2, is correlated with the process of interest even given covariates introduced in the process regression model. To avoid bias due to dependent censoring, we derive a variant of Inverse Probability of Censoring Weighting (IPCW) (Robins & Rotnitzky, 1992) based on a semiparmametric additive hazard model (Lin & Ying, 1994). We also derive a stabilized version of the proposed inverse weights (Hernán, Brumback, & Robins, 2000; Robins & Finkelstein, 2000; Zhang & Schaubel, 2011) to simplify calculations and, hence, considerably reduce computing time in large data sets. Analogous to a weighted partial likelihood score equation (Cox, 1972; Sasieni, 1993), the regression estimator could be estimated by the solution to an estimating equation free of the baseline probability.

Existing methods pertinent to the data structure of our interest include Scheike & Zhang (2007), which involved directly modeling a state occupation probability in a multi-state model. In addition, pseudo-observation approaches were developed by Andersen, Klein, & Rosthøj et al. (2003) and Grand & Putter (2016). These methods are fully parametric and, as such, require full specification of the baseline state occupation probability over time (unlike the proposed methods).

Pseudo-value approaches also tend to be computationally cumbersome in large data sets, and they tend to have rather restrictive assumptions on the censoring mechanism.

Our methods have several novel features. First, the baseline probability function is represented in the model nonparametrically. This is a potentially big advantage, since covariate effects typically take centre stage in process regression (and other regression settings), with little interest in modeling the baseline probability. We essentially profile out the baseline probability function, which results in major computation reduction relative to a fully parametric probability model. Second, the response indicator we consider is the joint event of survival and state occupation. The limited number of process regression methods that considered a terminating event typically modeled the state indicator conditional on survival. Notwithstanding the utility of such approaches, it is useful to develop methods for the joint outcome of survival and state occupation (a response for which few methods have been developed). Third, existing process regression methods typically assume independent censoring, while the proposed methods allow for both independent and dependent censoring. Fourth, in contrast with the vast majority of methods that accommodate dependent censoring, we construct the inverse weight under an additive hazards model.

The remainder of the article is organized as follows. We set up notation and describe our proposed methods in the next section. In Section 3, we derive the

asymptotic properties of the regression parameter estimator and baseline probability function estimator, with proofs provided in the Supplemental Materials document. Simulation studies are performed to evaluate finite-sample properties in Section 4. In Section 5, we apply the proposed methods to national ESLD data. Finally, concluding remarks are given in Section 6. Note that R code for carrying out the proposed methods is provided in the Supplementary Materials.

2. MODEL AND METHODS

We begin by formalizing the data structure described in Section 1. We then describe the proposed inference methods.

2.1. Notation and Assumed Models

Suppose that there are n independent subjects ($i = 1, 2, \ldots, n$). Let D_i be the death (terminal event) time of subject i, and let $\mathcal{E}_i(t)$ be an indicator function taking value 1 when subject i is occupying the state of interest at follow-up time t. The outcome of interest is the joint event, being alive and occupying the state of interest, which we denote by $A_i(t) = \mathcal{E}_i(t)I(D_i > t)$. In the end-stage liver disease example, D_i represents death (in the absence of liver transplantation), while $\mathcal{E}_i(t) = 1$ if subject i is active on the liver transplant waiting list as of t days following initial waiting list registration, and 0 otherwise. We let $\boldsymbol{Z}_i(t)$ be a covariate vector, with any time-dependent elements being external (Kalbfleisch & Prentice, 2002). The probability of interest is the probability that a subject i is

alive and occupying the state of interest at time t,

$$\pi_i(t) = P[A_i(t) = 1|\boldsymbol{Z}_i(t)].$$

We assume that the covariate $\boldsymbol{Z}_i(t)$ has a multiplicative effect on a completely unspecified baseline probability function, $\pi_0(t)$, such that

$$\pi_i(t) = \pi_0(t)\text{exp}[\boldsymbol{\beta}_0^T \boldsymbol{Z}_i(t)], \tag{1}$$

where $\boldsymbol{\beta}_0$ is the p-dimensional parameter vector of chief interest. Model (1) is reminiscent of the Cox proportional hazards model. However, there are some important differences, including the fact that $\pi_i(t)$ is interpreted as a marginal probability, rather than a conditional probability rate, and that $\pi_i(t)$ need not be monotone.

Two types of censoring are considered. Let C_{1i} be the administrative censoring, which is assumed to be conditionally independent of $A_i(t)$ given $\boldsymbol{Z}_i(t)$; i.e.,

$$E[A_i(t)|\boldsymbol{Z}_i(t), C_{1i}, C_{1i} \geq t] = E[A_i(t)|\boldsymbol{Z}_i(t)]. \tag{2}$$

This is also known as covariate-dependent censoring, in the sense that C_{1i} is allowed to depend on the covariate employed in the model of interest. We let C_{2i} represent dependent censoring time; that is, C_{2i} is not assumed to be conditionally independent $A_i(t)$ given $\boldsymbol{Z}_i(t)$. For example, in the context of our motivating example, a patient's pre-transplant $A_i(t)$ process is censored if and when the

patient receives a liver transplant; i.e., the liver transplant hazard and mortality hazard may be correlated, even conditional on $\boldsymbol{Z}_i(t)$. We let $C_i = C_{1i} \wedge C_{2i}$ represent the censoring time, where $a \wedge b = \min(a, b)$. Here we consider follow-up time $t \in [0, \tau]$, where τ is a pre-specified constant satisfying $Pr(C_i \geq \tau) > 0$ for $i = 1, 2, ..., n$. In practice, τ could be chosen as the maximum observed censoring time. To further characterize C_{2i}, we let $\boldsymbol{X}_i^{\dagger}(t)$ represent the time-dependent covariate at time t. Note that $\boldsymbol{X}_i^{\dagger}(t)$ would typically contain the elements of $\boldsymbol{Z}_i(t)$, as well as additional factors (the most important being internal time-varying covariates assumed to predict both D_i and C_{2i}). We denote the covariate history as of time t by $\widetilde{\boldsymbol{X}}_i(t) = \{\boldsymbol{X}_i^{\dagger}(s), s \in [0, t)\}$. Finally, we let $\lambda_i^{C_2}(t)$ be the cause specific hazard function of C_{2i} which is defined as

$$\lambda_i^{C_2}(t) = \lim_{\delta \to 0} \frac{1}{\delta} Pr[t \leq C_{2i} < t + \delta | C_{2i} \geq t, D_i \geq t, \widetilde{\boldsymbol{X}}_i(t)].$$

We assume that, conditional on $\widetilde{\boldsymbol{X}}_i(t)$, the cause-specific hazard of C_{2i} at time t does not further depend on the possibly unobserved, D_i or $\mathcal{E}_i(u), u \in (t, \tau]$, i.e.,

$$\lambda_i^{C}\{t|\widetilde{\boldsymbol{X}}_i(t)\} = \lambda_i^{C}\{t|\widetilde{\boldsymbol{X}}_i(t), C_{1i}, C_{1i} \geq t, D_i, D_i \geq t, \mathcal{E}_i(u), u \in (t, \tau]\}. \quad (3)$$

This represents the critical 'no unmeasured confounders' for censoring assumption (Robins, 1993; Robins & Finkelstein, 2000). The following semiparametric additive hazards model (Lin & Ying, 1994) is assumed for C_{2i}:

$$\lambda_i^{C_2}(t; \boldsymbol{\theta}_0) = \lambda_0^{C_2}(t) + \boldsymbol{\theta}_0^T \boldsymbol{X}_i(t), \quad (4)$$

where $\lambda_0^{C_2}(t)$ is the baseline hazard function for C_{2i} and the covariate $\boldsymbol{X}_i(t)$ is chosen (e.g., through model selection techniques) to satisfy $\lambda_i^{C_2}[t|\boldsymbol{X}_i(t)] = \lambda_i^{C_2}[t|\widetilde{\boldsymbol{X}}_i(t)]$. Note that $\boldsymbol{X}_i(t)$ need not be based on the covariate status at time t and could, in fact, contain elements representing the covariate history. Finally, we define $\Lambda_i^{C_2}(t) = \int_0^t \lambda_i^{C_2}(s)ds$ as the cumulative hazard function corresponding to C_{2i}, and $\Lambda_0^{C_2}(t) = \int_0^t \lambda_0^{C_2}(s)ds$ as the cumulative baseline hazard function.

Model (4) facilitates the calculation of the weight function, since the baseline cumulative hazard cancels out after a particular stabilizing factor is introduced; we provide specifics later. Note that model checking could proceed using techniques proposed in Yin (2007).

As described, there are multiple sources of censoring. C_1 is intended to represent independent censoring; e.g., administrative censoring, or random loss to follow-up. C_1 is assumed to be conditionally independent of the process $A(t)$ given the covariate $\boldsymbol{Z}(t)$. Note that, any time-dependent elements of $\boldsymbol{Z}(t)$ must be external (Kalbfleisch and Prentice, 2002); e.g., air temperature, precipitation status, etc. In contrast, C_2 is intended to represent dependent censoring; e.g., a binary non-reversible treatment which is applied after follow-up begins. C_2 is assumed to be independent of $A(t)$ given $\tilde{\boldsymbol{X}}(t)$, which will contain internal time-dependent covariates (e.g., blood pressure, serum creatinine). Covariate selection for the C_1 and C_2 models should proceed separately.

In the next subsection, we describe the proposed methods for the scenario where C_1 is known. We then subsequently describe the proposed techniques for the more frequently occurring set-up when C_1 is not known.

2.2. Case 1: C_1 Known

We first consider the case where the independent censoring time, C_{1i}, is known for all subjects. In such cases, C_{1i} is known even if D_i occurs first. This set-up would apply, for example, in a clinical trial with staggered entry but no drop-out or random loss to follow-up. This does not match most observational data, but it is a useful starting point in terms of outlining the proposed estimation techniques. Consider the following two estimating functions,

$$\sum_{i=1}^{n} \int_0^{\tau} \boldsymbol{Z}_i(t)[A_i(t) - \pi_i(t)]I(C_i \geq t)dt,$$

$$\sum_{i=1}^{n} [A_i(t) - \pi_i(t)]I(C_i \geq t).$$

These two estimating functions do not have expectation zero under assumptions (1) and (2), since dependent censoring C_{2i} is potentially correlated with $A_i(t)$, even conditional on $\boldsymbol{Z}_i(t)$. To handle this issue, we utilize Inverse Probability of Censoring Weighting (IPCW) (Robins & Rotnitzky, 1992) to accommodate dependent censoring. Define

$$W_i^A(t; \boldsymbol{\theta}_0) = I(C_{2i} \geq t)\text{exp}\{\Lambda_i^{C_2}[t \wedge D_i; \boldsymbol{\theta}_0]\} \quad (5)$$

as, heuristically, the inverse probability of being uncensored by C_{2i} as of time t. Note that dependent censoring cannot occur after death; this is an assumed property of the data structure we consider. This makes sense intuitively, from the perspective that C_{2i} is driven by internal factors (including a subject's survival), which, by definition, shut down at death. For instance, in our motivating example, a patient cannot receive a liver transplant after dying.

Now, consider the revised estimating functions

$$\sum_{i=1}^{n} \int_0^{\tau} \boldsymbol{Z}_i(t)[A_i(t) - \pi_i(t)]I(C_{1i} \geq t)W_i^A(t;\boldsymbol{\theta}_0)dt, \tag{6}$$

$$\sum_{i=1}^{n} [A_i(t) - \pi_i(t)]I(C_{1i} \geq t)W_i^A(t;\boldsymbol{\theta}_0). \tag{7}$$

Under the assumption given in (3), these two weighted estimating functions have expectation zero. Basically, the proof follows from the fact that $W_i^A(t;\boldsymbol{\theta}_0)$ can be written as one minus a Martingale component of C_{2i}, which is independent of $[A_i(t) - \pi_i(t)]I(C_{1i} \geq t)$ conditional on $\boldsymbol{X}_i(t)$ (Robins & Finkelstein, 2000). A proof is provided in Section 4.6 of the Supplemental Materials.

In contrast to the majority of the existing literature, we derive a stabilizer that is merely a function of t, which is valid, since $E\{g(t)[A_i(t) - \pi_i(t)]I(C_{1i} \geq t)W_i^A(t;\boldsymbol{\theta}_0)|\boldsymbol{Z}_i(t)\} = 0$ will hold. We denote $g(t) = \exp[-\Lambda_0^{C_2}(t \wedge D_i)]$ for this purpose, such that

$$W_i^B(t;\boldsymbol{\theta}_0) = I(C_{2i} \geq t)\exp\left\{\Lambda_i^{C_2}(t \wedge D_i;\boldsymbol{\theta}_0) - \Lambda_0^{C_2}(t \wedge D_i)\right\}. \tag{8}$$

Under the assumed additive hazard model from (4), $W_i^B(t;\boldsymbol{\theta}_0) = I(C_{2i} \geq t)\text{exp}[\int_0^{t\wedge D_i} \boldsymbol{\theta}_0^T \boldsymbol{X}_i(s)ds]$, since $\Lambda_0^{C_2}(t)$ cancels out. This nice property enables us to get unbiased estimating equations without estimating the baseline cumulative hazard $\Lambda_0^{C_2}(t)$ and, hence, should increase computational efficiency.

Solving (7) for $\pi_0(t)$ by treating $\boldsymbol{\beta}$ as known, then substituting the estimated $\pi_0(t)$ into (6) gives us the following estimating equation,

$$\sum_{i=1}^{n} \int_0^{\tau} \{\boldsymbol{Z}_i(t) - \bar{\boldsymbol{Z}}[t;\boldsymbol{\beta}, W(\boldsymbol{\theta}_0)]\} I(C_{1i} \geq t) W_i(t;\boldsymbol{\theta}_0) A_i(t) dt = \mathbf{0}, \quad (9)$$

where $\bar{\boldsymbol{Z}}[t;\boldsymbol{\beta}, W(\boldsymbol{\theta})] = \boldsymbol{Z}^{(1)}[t;\boldsymbol{\beta}, W(\boldsymbol{\theta})]/Z^{(0)}[t;\boldsymbol{\beta}, W(\boldsymbol{\theta})]$, $\boldsymbol{Z}^{(k)}[t;\boldsymbol{\beta}, W(\boldsymbol{\theta})] = n^{-1}\sum_{i=1}^{n} \boldsymbol{Z}_i(t)^{\otimes k} I(C_{1i} \geq t) W_i(t;\boldsymbol{\theta}) \text{exp}\{\boldsymbol{\beta}^T \boldsymbol{Z}_i(t)\}$, for $k = 0, 1, 2$, where $\boldsymbol{a}^{\otimes 0} = 1$, $\boldsymbol{a}^{\otimes 1} = \boldsymbol{a}$, and $\boldsymbol{a}^{\otimes 2} = \boldsymbol{a}\boldsymbol{a}^T$, and with $W_i(t;\boldsymbol{\theta})$ set to either $W_i^A(t;\boldsymbol{\theta})$ or $W_i^B(t;\boldsymbol{\theta})$. Having estimated $\boldsymbol{\beta}_0$ through the root of (9), denoted by $\widehat{\boldsymbol{\beta}}$, we could estimate $\pi_0(t)$ by solving (7),

$$\widehat{\pi}_0(t) = \frac{\sum_{i=1}^{n} A_i(t) I(C_{1i} \geq t) W_i(t;\boldsymbol{\theta}_0)}{\sum_{i=1}^{n} I(C_{1i} \geq t) W_i(t;\boldsymbol{\theta}_0) \text{exp}[\widehat{\boldsymbol{\beta}}^T \boldsymbol{Z}_i(t)]}. \quad (10)$$

Equation (10) does not constrain $\widehat{\pi}_0(t)$ to be ≤ 1. The frequency with which out-of-range $\widehat{\pi}_0(t)$ and $\widehat{\pi}_i(t)$ occur will be depend on the data set at hand. This issue does not interfere with $\widehat{\boldsymbol{\beta}}$ being a consistent estimator, which makes sense from the perspective that, out-of-range or not, $\widehat{\pi}_0(t)$ was profiled out. Note that out-of-range estimators have a long history of being tolerated in other areas of statistics. Further discussion on this issue is provided in Section 6.

We define the integral of baseline probability function up to time L,

$$P_0(L) = \int_0^L \pi_0(t)dt,$$

which could be interpreted as the restricted mean time survived in the state of interest for a subject with covariate equal to the reference level for all elements. The quantity $P_0(L)$ is estimated by $\widehat{P}_0(L) = \int_0^L \widehat{\pi}_0(t)dt$.

Based on model (4), one could estimate $\boldsymbol{\theta}_0$ and $d\Lambda_0^{C_2}(t)$ by $\widehat{\boldsymbol{\theta}}$ and $d\widehat{\Lambda}_0^{C_2}(t;\widehat{\boldsymbol{\theta}})$. From the works of Lin & Ying (1994), $\widehat{\boldsymbol{\theta}}$ and $d\widehat{\Lambda}_0^{C_2}(t;\widehat{\boldsymbol{\theta}})$ are given by

$$\widehat{\boldsymbol{\theta}} = \left[\sum_{i=1}^n \int_0^\infty Y_i(t)\{\boldsymbol{X}_i(t) - \bar{\boldsymbol{X}}(t)\}^{\otimes 2}dt\right]^{-1}\left[\sum_{i=1}^n \int_0^\infty \{\boldsymbol{X}_i(t) - \bar{\boldsymbol{X}}(t)\}dN_i^{C_2}(t)\right]$$

$$d\widehat{\Lambda}_0^{C_2}(t;\widehat{\boldsymbol{\theta}}) = \frac{\sum_{i=1}^n\{dN_i^{C_2}(t) - Y_i(t)\widehat{\boldsymbol{\theta}}^T\boldsymbol{X}_i(t)dt\}}{\sum_{i=1}^n Y_i(t)},$$

where $N_i^{C_2}(t) = I(C_{2i} \le t \wedge X_i)$, $dN_i^{C_2}(t) = N_i^{C_2}(t^- + dt) - N_i^{C_2}(t^-)$, $Y_i(t) = I(X_i \ge t)$ and $X_i = D_i \wedge C_i$. Let $\boldsymbol{X}^{(k)}(t) = n^{-1}\sum_{i=1}^n \boldsymbol{X}_i(t)^{\otimes k}Y_i(t)$, $\bar{\boldsymbol{X}}(t) = \boldsymbol{X}^{(1)}(t)/X^{(0)}(t)$. One could further estimate $d\Lambda_i^{C_2}(t;\boldsymbol{\theta}_0)$ by $d\widehat{\Lambda}_i^{C_2}(t;\widehat{\boldsymbol{\theta}}) = d\widehat{\Lambda}_0^{C_2}(t;\widehat{\boldsymbol{\theta}}) + \widehat{\boldsymbol{\theta}}^T\boldsymbol{X}_i(t)dt$. After that, the estimated weights $\widehat{W}_i^A(t;\widehat{\boldsymbol{\theta}})$ or $\widehat{W}_i^B(t;\widehat{\boldsymbol{\theta}})$ can be calculated.

In general we would recommend $W_i^B(t;\theta)$ over $W_i^A(t;\theta)$, since the former is more convenient computationally and generally more precise, with both characteristics owing to the fact that $\Lambda_0^{C_2}(t)$ is bypassed.

2.3. Case 2: C_1 Not Known

Next we consider a more realistic scenario where independent censoring time is random, with the randomness implying that C_{1i} is unknown when D_i occurs first. Setting the missing censoring time to either D_i or the maximum follow-up time, τ, would introduce bias, since the indicator $I(C_{1i} > D_i)$ is correlated with the target process. As for C_{2i}, if D_i happens first, then C_{2i} could be treated as infinity or considered to be subject to a dependent censoring hazard of 0 for $t > D_i$. The reason is that C_{2i} relies on time varying covariate vector $\boldsymbol{X}_i$ containing internal covariates, which would shut down if death occurs. In this case, the hazard for C_{2i} is zero after D_i. Moreover, the inverse weighting function $W_i(t;\boldsymbol{\theta}_0)$ remains constant after D_i if it is observed earlier than C_i.

Our solution is to impute missing C_{1i} from its assumed model (Rubin, 2004). Specifically, we assume that the hazard function for C_1 follows the proportional hazards model,

$$\lambda_i^{C_1}(t;\boldsymbol{\gamma}_0) = \lambda_0^{C_1}(t)\text{exp}[\boldsymbol{\gamma}_0^T \boldsymbol{Z}_i(t)], \tag{11}$$

where $\lambda_0^{C_1}(t)$ is the baseline hazard and $\boldsymbol{\gamma}_0$ is an unknown regression parameter.

For subjects with $C_i \leq D_i$, we set the imputed censoring time as the known censoring time. In the mth imputed dataset, for subjects with $C_i > D_i$, we impute $\widehat{C}_{1i}^{\langle m \rangle}$ from the estimated conditional survival function,

$$\widehat{G}(t;\widehat{\boldsymbol{\gamma}}) = I(t \geq D_i)\text{exp}[-\widehat{\Lambda}_i^{C_1}(t;\widehat{\boldsymbol{\gamma}}) + \widehat{\Lambda}_i^{C_1}(D_i;\widehat{\boldsymbol{\gamma}})].$$

Standard partial likelihood (Cox, 1975) techniques can be fitted to the observed censoring time data $\{X_i, I(C_{1i} \leq D_i \wedge C_{2i}), \boldsymbol{Z}_i(t);\ \ t \in [0,\tau]\}_{i=1}^n$ to compute $\widehat{\boldsymbol{\gamma}}$. The baseline cumulative hazard function for $\Lambda_0^{C_1}(t)$ is estimated through the method of Breslow (1972). Then, we set $C_{1i}^{\langle m\rangle} = I(C_i \leq D_i)C_i + I(C_i > D_i)\widehat{C}_{1i}^{\langle m\rangle}$. Note that the $C_{1i}^{\langle m\rangle}$ are bounded above at τ, with τ defined in Section 2.1. In total, we will create M imputation datasets. Within each imputed dataset m, we substitute $C_{1i}^{\langle m\rangle}$ for C_{1i} and set C_{2i} as τ if $D_i < C_{1i}$. Estimators arising from the mth imputed data set are denoted by $\widehat{\boldsymbol{\beta}}^{\langle m\rangle}$ and $\widehat{\pi}_0^{\langle m\rangle}(t)$. We then estimate $\boldsymbol{\beta}_0$ and $\pi_0(t)$ by averaging the M imputation-specific estimators,

$$\widehat{\boldsymbol{\beta}}^M = M^{-1}\sum_{m=1}^{M}\widehat{\boldsymbol{\beta}}^{\langle m\rangle}; \tag{12}$$

$$\widehat{\pi}_0^M(t) = \frac{\sum_{m=1}^{M}\sum_{i=1}^{n} A_i(t)I(C_{1i}^{\langle m\rangle} \geq t)W_i(t;\boldsymbol{\theta}_0)}{\sum_{m=1}^{M}\sum_{i=1}^{n} I(C_{1i}^{\langle m\rangle} \geq t)W_i(t;\boldsymbol{\theta}_0)\exp[\boldsymbol{Z}_i(t)^T\widehat{\boldsymbol{\beta}}^M]}. \tag{13}$$

Note that the multiple imputation method we employ does not sample the parameters assumed to underly the C_1 distribution but, instead, imputes $C_{1i}^{\langle m\rangle}$ values from the same estimated survival curve. This procedure has been referred to as Improper Imputation (Wang & Robins, 1998). As a consequence of this choice, the well-established variance formula for multiple imputation (Rubin, 2004) does not apply, necessitating an explicit derivation of variances estimators corresponding to (12) and (13) through the combination of several asymptotic expansions. These issues are dealt with in the next section, along with our treatment of the large-sample properties of the proposed estimators.

Some commentary regarding our combination of inverse weighting and imputation is useful at this juncture. The sources of censoring, C_1 and C_2, have very different implications in terms of their impact on parameter estimation. Due to the marginal nature of $\pi_i(t)$, subjects contribute relevant follow-up until time C_{1i}, which may occur after D_i. A similar issue arises in Gray (1988) and Fine & Gray (1999) in the context of inference targeting the subdistribution hazard function in the competing risks setting. Both Gray (1988) and Fine & Gray (1999) used a weight function, rather than imputation. It is important to note that the weight was not an inverse weight; if anything, it could be described as an 'inverse-inverse' weight, since it corresponds to a conditional survival probability (as opposed to the inverse thereof). Essentially, the risk sets contributions are weighted with respect the conditional probability of remaining uncensored at time t (consider $t > D_i$), given that $C_{i1} > D_i$. From this perspective, imputing C_{1i} does in fact appear consistent in spirit with the weights used in subdistribution modeling, which can be cast heuristically as mean imputation at the risk set level. In fact, Ruan & Gray (2008) later proposed subdistribution hazard methods that involved imputing censoring times. An analogous imputation scheme was later employed by Schaubel & Zhang (2010).

Note that C_{1i} marks the end of relevant follow-up and, hence, is a variate we want to observe. One could use either weighting (Ghosh and Lin, 2002) or imputation techniques to recover the missingness of C_{1i}, but we choose the latter

due to computational ease. Provided we either observe C_{1i} or can validly impute its value, the $[0, C_{1i}]$ experience could be analyzed without inverse weighting if not for dependent censoring, C_{2i}. In line with the setting where IPCW is typically employed, we inverse weight the uncensored experience to reflect data that would have been observed if C_{1i} were the only source of censoring. The events D_i and C_{2i} serve as competing risks in the sense of Prentice et al. (1978). From this angle, dependent censoring does not occur after D_i; hence, the weight function not incrementing for $t > D_i$, per (5).

3. ASYMPTOTIC PROPERTIES

We provide asymptotic results and proofs for the known C_1 case in the Supplemental Materials. Here, we focus on the random C_1 setting under the weight $W_i^B(t; \boldsymbol{\theta}_0)$.

Theorem 1. *Under assumptions (1), (2), (3), and (11),* $\widehat{\boldsymbol{\beta}}^M$ *is a consistent estimator of* $\boldsymbol{\beta}_0$*, and as* $n \to \infty$*,* $n^{1/2}(\widehat{\boldsymbol{\beta}}^M - \boldsymbol{\beta}_0)$ *converges in distribution to a mean-zero normal random variable with a variance-covariance matrix* $\boldsymbol{\Sigma}(\boldsymbol{\theta}_0, \boldsymbol{\beta}_0, \boldsymbol{\gamma}_0, M) = E[\boldsymbol{f}_1^{\boldsymbol{\beta}_B}(\boldsymbol{\theta}_0, \boldsymbol{\beta}_0, \boldsymbol{\gamma}_0, M)^{\otimes 2}]$*, where*

$$\boldsymbol{f}_i^{\boldsymbol{\beta}_B}(\boldsymbol{\theta}, \boldsymbol{\beta}, \boldsymbol{\gamma}, M) = \boldsymbol{\Omega}[\boldsymbol{\beta}, W^B(\boldsymbol{\theta})]^{-1}[\boldsymbol{\Phi}_{1i}^{\boldsymbol{\beta}_B}(\boldsymbol{\theta}, \boldsymbol{\beta}, \boldsymbol{\gamma}, M) + \boldsymbol{\Phi}_{2i}^{\boldsymbol{\beta}_B}(\boldsymbol{\theta}, \boldsymbol{\beta})];$$

$$\boldsymbol{\Phi}_{1i}^{\boldsymbol{\beta}_B}(\boldsymbol{\theta}, \boldsymbol{\beta}, \boldsymbol{\gamma}, M) = \int_0^\tau \{\boldsymbol{Z}_i(t) - \bar{\boldsymbol{z}}[t; \boldsymbol{\beta}, W^B(\boldsymbol{\theta})]\} W_i^B(t; \boldsymbol{\theta}) \frac{1}{M} \sum_{m=1}^{M} dM_i^{\langle m \rangle}(t; \boldsymbol{\beta}, \boldsymbol{\gamma});$$

$$dM_i^{\langle m \rangle}(t; \boldsymbol{\beta}, \boldsymbol{\gamma}) = \left[A_i(t) - \pi_0(t) exp\{\boldsymbol{\beta}^T \boldsymbol{Z}_i(t)\}\right] I(C_{1i}^{\langle m \rangle} \geq t; \boldsymbol{\gamma}) I(C_{2i} \geq t) dt.$$

Note that we define:

$$\boldsymbol{\Phi}_{2i}^{\boldsymbol{\beta}_B}(\boldsymbol{\theta},\boldsymbol{\beta}) = \boldsymbol{H}^B[\boldsymbol{\beta},\boldsymbol{\theta},W^B(\boldsymbol{\theta})][\boldsymbol{\Omega}^{C_2}]^{-1}\boldsymbol{u}_i^{C_2}(\boldsymbol{\theta}).$$

For $\pi_0^M(t)$, we have the following result.

Theorem 2. *Under assumptions (1), (2), (3), and (11), as* $n \to \infty$, $n^{1/2}(\widehat{\pi}_0^M - \pi_0)$ *converges weakly to a mean-zero Gaussian process with a variance and covariance matrix between* $n^{1/2}[\widehat{\pi}_0^M(s) - \pi_0(s)]$ *and* $n^{1/2}[\widehat{\pi}_0^M(t) - \pi_0(t)]$*, given by* $\sigma(s,t,\boldsymbol{\theta}_0,\boldsymbol{\beta}_0,\boldsymbol{\gamma}_0,M) = E[\xi_1(s,\boldsymbol{\theta}_0,\boldsymbol{\beta}_0,\boldsymbol{\gamma}_0,M)\xi_1(t,\boldsymbol{\theta}_0,\boldsymbol{\beta}_0,\boldsymbol{\gamma}_0,M)]$*, where*

$$\xi_i(t,\boldsymbol{\theta},\boldsymbol{\beta},\boldsymbol{\gamma},M) = \frac{\widetilde{f}_i^{\pi_1,B}(t,\boldsymbol{\theta}) - [E\widetilde{f}_1^{\pi_1,B}(t,\boldsymbol{\theta})]\widetilde{f}_i^{\pi_2,B}(t,\boldsymbol{\theta},\boldsymbol{\beta},\boldsymbol{\gamma},M)}{E[\widetilde{f}_1^{\pi_2,B}(t,\boldsymbol{\theta},\boldsymbol{\beta},\boldsymbol{\gamma},M)]};$$

$$\widetilde{f}_i^{\pi_2,B}(t,\boldsymbol{\theta},\boldsymbol{\beta},\boldsymbol{\gamma},M) = \widetilde{f}_i^{\pi_{21},B}(t,\boldsymbol{\theta},\boldsymbol{\beta},\boldsymbol{\gamma},M) + \widetilde{f}_i^{\pi_{22},B}(t,\boldsymbol{\theta},\boldsymbol{\beta},\boldsymbol{\gamma},M) + \widetilde{f}_i^{\pi_{23},B}(t,\boldsymbol{\theta},\boldsymbol{\beta});$$

$$\widetilde{f}_i^{\pi_{21},B}(t,\boldsymbol{\theta},\boldsymbol{\beta},\boldsymbol{\gamma},M) = \frac{1}{M}\sum_{m=1}^{M} I(C_{1i}^{\langle m\rangle} \geq t;\boldsymbol{\gamma})e^{\boldsymbol{\beta}^T\boldsymbol{Z}_i(t)}W_i^B(t;\boldsymbol{\theta});$$

$$\widetilde{f}_i^{\pi_{22},B}(t,\boldsymbol{\theta},\boldsymbol{\beta},\boldsymbol{\gamma},M) = \boldsymbol{z}^{(1)}[t;\boldsymbol{\beta},W^B(\boldsymbol{\theta})]^T\boldsymbol{f}_i^{\boldsymbol{\beta}_B}(\boldsymbol{\theta},\boldsymbol{\beta},\boldsymbol{\gamma},M).$$

For computational convenience, we suggest estimating the variances given above with the weight function treated as known. In this case, the proposed variance estimator will be tend to be conservative (Hernán, Brumback, & Robins, 2000; Pan & Schaubel, 2008; Zhang & Schaubel, 2011). We evaluate this approximation in the next section.

4. SIMULATION STUDIES

We report on simulations to evaluate performance of our methods. The results of three weights are evaluated: $W_i(t) = 1$, which does not correctly accommodate the censoring mechanisms and is included for comparison purposes only; $W_i^A(t;\boldsymbol{\theta}_0)$ defined in (5); and the stabilized weights, $W_i^B(t;\boldsymbol{\theta}_0)$, defined in (8).

For each simulation setting, two scenarios ($n = 500$, $n = 1,000$) are generated. The covariate $\boldsymbol{Z} = (Z_1, Z_2)'$ has elements that are Bernoulli(0.5). The terminal event, D, is generated by the hazard function $\lambda_0^D \exp\{\boldsymbol{\alpha}_0'\boldsymbol{Z}\}$, where $\lambda_0^D = 0.015$ or 0.02, and $\boldsymbol{\alpha}_0 = (-0.609, 0.609)'$. The target model represents the probability of being alive and active, $\pi(t) = \pi_0(t)\exp\{\boldsymbol{\beta}_0'\boldsymbol{Z}\}$, where $\pi_0(t) = 0.3 - 0.0025t$, for $t = 1, 2, ..., 100$. The event of being active on waiting list given the subject is alive is sampled from the conditional probability, $P\{\mathcal{E}(t) = 1|D > t, \boldsymbol{Z}\} = \pi(t)\exp[\lambda^D \times t]$.

We generated two censoring times, C_1 and C_2, for each individual. The independent censoring time, C_1, is generated from the hazard function $\lambda_0^{C_1}\exp\{\boldsymbol{\gamma}_0'\boldsymbol{Z}\}$, where $\lambda_0^{C_1} = 0.015$ and $\boldsymbol{\gamma}_0 = (0.609, -0.609)'$. For the dependent censoring time, C_2, we first generate X_t, where $X_t = \min\{D, -40 \times \log[Z_1\epsilon_1 + (1 - Z_2)(1 - \epsilon_1)] + 5\epsilon_2\}$, $\epsilon_1 = \int_1^D [1 - A(t)]\, dt/100$, $\epsilon_2 \sim$ Uniform(0,1). Let $X(t) = I(X_t \geq t)$, with $X(t)$ being dependent on $A(t)$ even conditional on $\boldsymbol{Z}$ due to its mutual association with ϵ_1. Next, we generate time-dependent censoring time C_2 from the hazard function $\lambda_0^{C_2} + \phi_1 Z_1 + \phi_2 Z_2 +$

$\phi_3 X(t)$. Note that ϵ_1 and ϵ_2 are mutually independent.

We set $\lambda_0^D = 0.015$, $\lambda_0^{C_2} = 0.005$, $\phi_1 = \phi_2 = -0.002$, and $\phi_3 = 0.025$ for heavy censoring C_2, which results in about 36% of subjects being censored by C_2 and 36% of subjects censored by C_1. Moreover, two magnitudes of $\boldsymbol{\beta}_0$ are considered: 0.916 and 0.405. Next we consider a light censoring case for C_2, where $\lambda_0^D = 0.02$, $\lambda_0^{C_2} = 0.003$, $\phi_1 = \phi_2 = -0.001$, and $\phi_3 = 0.015$. This set-up results in 23% of subjects are censoring by C_2 and 37% of subjects are censoring by C_1. In Tables 1 and 2 we treat censoring time as random and obtain imputed estimators based on $M = 5$.

In each setting, the biases of $\widehat{\boldsymbol{\beta}}$ and $\widehat{P}_0(50)$ are very small for both $\widehat{W}_i^A(t; \widehat{\boldsymbol{\theta}})$ and $\widehat{W}_i^B(t; \widehat{\boldsymbol{\theta}})$, indicating that our estimators are consistent. Moreover, the average asymptotic standard errors (ASEs) are generally close to, but slightly larger than the empirical standard deviations (ESDs). This results from our treating the estimated weight function as fixed. The empirical coverage probabilities (ECPs) are also around 0.95, implying the accuracy of large-sample confidence intervals. Due to the substantial biases of not adjusting for time-dependent confounders, the unweighted method exhibits bias and has inaccurate estimated variance and poor coverage probabilities.

Some further notes are in order. First, $M = 1$ is valid, unlike when the variance formula in Rubin (2004) is employed. Increasing M will generally increase precision and improve empirical coverage probability, albeit with diminishing

returns; this was demonstrated empirically in Zhan (2017). Second, simulation results with $n = 100$ were not very good (Zhan, 2017). We do not recommend the methods for sample sizes less than $n = 500$.

5. ANALYSIS OF END-STAGE LIVER DISEASE DATA

We applied the proposed methods to model the waiting list active/inactive process using 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 U.S., as submitted by the members of the Organ Procurement and Transplantation Network (OPTN), and has been described elsewhere. The Health Resources and Services Administration (HRSA), U.S. Department of Health and Human Services provides oversight to the activities of the OPTN and SRTR contractors.

In the end-stage liver disease (ESLD) setting, the number of available deceased-donor livers is always less than the number of patients in need of liver transplantation. Once an ESLD patient is wait-listed, the patient's status can oscillate between active and inactive based on their medical condition. A wait-listed patient can receive deceased-donor organ offers only when active. In this application, follow-up time t represents time since registration on the waiting list. The process $\mathcal{E}(t) = 1$ when the patient is active and 0 when inactive. We model $E[A(t)|\boldsymbol{Z}]$, where $A(t) = 1$ when the patient is alive and active on the waiting list.

TABLE 1: Simulations results for $\widehat{\boldsymbol{\beta}}^{5}$ with random censoring time based on $M = 5$, and $5,000$ replicates.

Censoring	n	$\boldsymbol{\beta}_0$	Weights	Bias	ASE	ESD	CP
Heavy	500	(0.916, -0.916)	1	(0.160, -0.051)	(0.081, 0.080)	(0.077, 0.075)	(0.486, 0.924)
			W^A	(0.000, 0.006)	(0.098, 0.097)	(0.092, 0.094)	(0.958, 0.952)
			W^B	(0.017, -0.013)	(0.090, 0.089)	(0.085, 0.086)	(0.956, 0.950)
		(0.405, -0.405)	1	(0.166, -0.053)	(0.082, 0.080)	(0.079, 0.077)	(0.475, 0.910)
			W^A	(0.003, 0.002)	(0.106, 0.106)	(0.104, 0.107)	(0.949, 0.949)
			W^B	(0.020, -0.013)	(0.096, 0.097)	(0.092, 0.095)	(0.946, 0.947)
	1,000	(0.916, -0.916)	1	(0.162, -0.051)	(0.057, 0.057)	(0.056, 0.053)	(0.179, 0.874)
			W^A	(0.001, 0.004)	(0.070, 0.069)	(0.064, 0.067)	(0.964, 0.958)
			W^B	(0.016, -0.010)	(0.064, 0.064)	(0.059, 0.061)	(0.958, 0.955)
		(0.405, -0.405)	1	(0.166, -0.055)	(0.059, 0.057)	(0.057, 0.055)	(0.183, 0.857)
			W^A	(0.001, 0.006)	(0.076, 0.077)	(0.073, 0.075)	(0.955, 0.953)
			W^B	(0.017, -0.012)	(0.069, 0.070)	(0.064, 0.067)	(0.952, 0.952)
Light	500	(0.916, -0.916)	1	(0.097, -0.062)	(0.090, 0.089)	(0.085, 0.086)	(0.824, 0.902)
			W^A	(-0.006, 0.007)	(0.103, 0.103)	(0.098, 0.097)	(0.958, 0.962)
			W^B	(0.010, -0.009)	(0.098, 0.098)	(0.093, 0.094)	(0.958, 0.954)
		(0.405, -0.405)	1	(0.107, -0.071)	(0.091, 0.091)	(0.088, 0.088)	(0.793, 0.885)
			W^A	(0.013, -0.012)	(0.107, 0.107)	(0.103, 0.106)	(0.950, 0.953)
			W^B	(0.025, -0.025)	(0.101, 0.101)	(0.096, 0.096)	(0.950, 0.949)
	1,000	(0.916, -0.916)	1	(0.099, -0.059)	(0.064, 0.063)	(0.061, 0.060)	(0.665, 0.865)
			W^A	(-0.006, 0.008)	(0.073, 0.073)	(0.067, 0.069)	(0.964, 0.953)
			W^B	(0.005, -0.004)	(0.070, 0.070)	(0.064, 0.065)	(0.968, 0.963)
		(0.405, -0.405)	1	(0.108, -0.071)	(0.065, 0.064)	(0.062, 0.062)	(0.614, 0.818)
			W^A	(0.013, -0.013)	(0.076, 0.076)	(0.071, 0.073)	(0.957, 0.954)
			W^B	(0.020, -0.023)	(0.072, 0.072)	(0.067, 0.068)	(0.958, 0.950)

A patient's active process may be censored by liver transplantation, with such censoring representing dependent censoring due to mutual correlation between $A(t)$ and liver transplantation. Note that, due to the nature of the liver allocation system in the U.S., a patient's rank on the waiting list is determined by their Model for End-stage Liver Disease (MELD) score. In particular, the waiting list is sequenced in decreasing order of (current) MELD score. Since higher MELD scores correspond to higher pre-transplant mortality, a model of pre-transplant outcomes based on baseline (time zero) patient characteristics will generally be subject to dependent censoring via deceased-donor liver transplantation, C_2. Patients are subject to independent right censoring due to administrative censoring and living donor transplantation, aggregated into C_1. Note that living-donor transplants are not allocated using the MELD system and, as such, are not determined by internal time-varying covariates.

The sample size of this study is $n = 53,991$. There were 13,180 subjects observed to die, 17,982 patients who were independently censored, and 22,829 subjects who were dependently censored. Baseline covariates include blood type, gender, race, BMI status, hospitalization, age, region, and values at wait-listing for MELD score, serum albumin, and serum sodium. Comorbid conditions are also included, for example hepatitis C, noncholestatic, cholestatic, acute hepatic necrosis, metastatic disease, and malignant neoplasm. The covariate information at time zero is used to characterize the process of being alive and active on the

waiting list, and to implement the imputation of C_1. As for time dependent covariates, we include more predictors to the baseline covariates set: MELD all score, albumin levels, sodium, ascites, encephalitis, and dialysis status; we exclude baseline MELD, baseline albumin, and baseline sodium. Moreover, continuous variables are centred at their mean values.

The stabilized weights, $W^B(t;\boldsymbol{\theta})$, were used to remove bias due to dependent censoring. To further mitigate the impact of outliers, the weights were capped by 150. For subjects with observed death, we imputed C_1 as described in Section 2. Due to the size of the data set, we used $M = 1$.

Covariate effects along with P-values are listed in Table 3. Each 5-year increase in age at wait-listing was associated with a significant 2% decrease in alive/active probability. Relative to the United Network for Organ Sharing (UNOS) Region with the greatest number of wait-listed patients (Region 5), Regions 1, 10, and 11 had significant reductions in alive/active probability, at 17%, 6%, and 16%, respectively. Region 7 had a 6% increase ($p = 0.01$). The probability of being alive and active on the waiting list decreased by 1% for each (integer) increase in MELD score, and increased by 9% per unit increase in serum albumin.

The estimated baseline probability function is plotted in Figure 1. We estimated the integral of baseline probability of being alive and active over $[0, 5]$ at 3.92 years; this indicates that a 'baseline' patient (i.e., a patient with all covariates

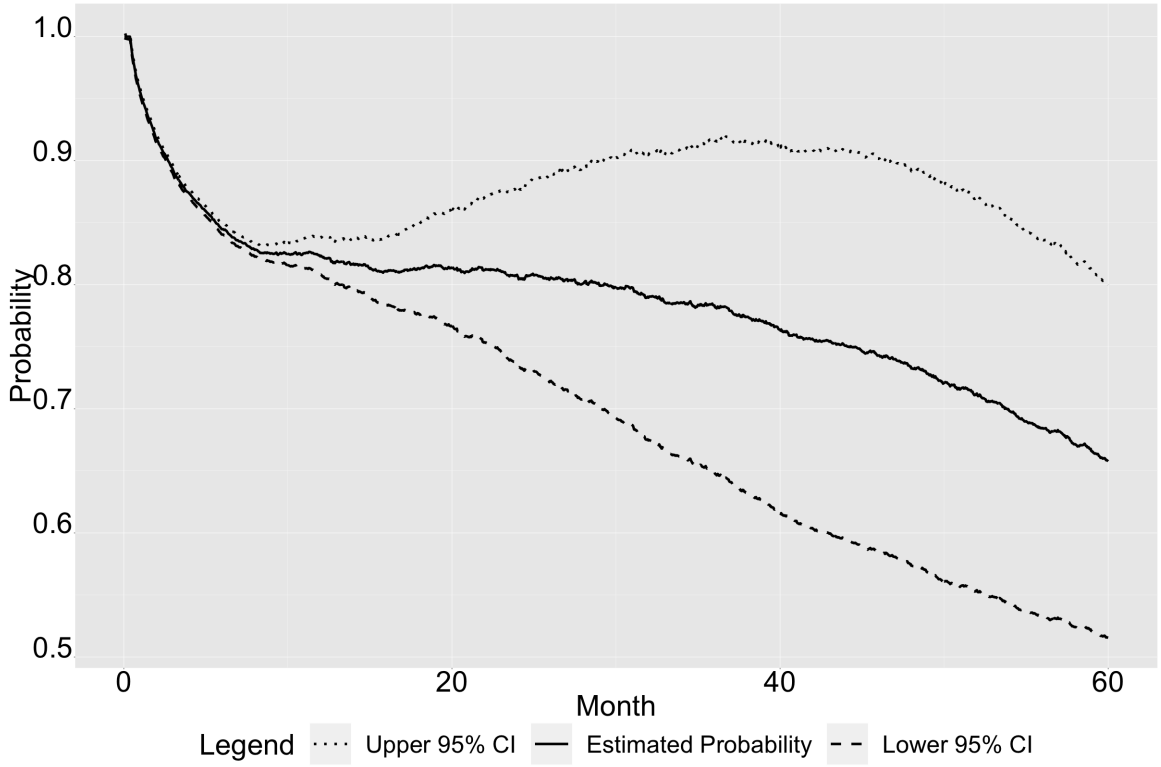


FIGURE 1: Estimated baseline probability of the probability of being and alive and active on the waiting list (solid line), along with point-wise 95% confidence intervals (dashed line).

equal to the reference) would be expected to be alive and active on the waiting list for approximately 4 of the first 5 years after wait-listing, in the absence of liver transplantation. Point-wise 95% confidence interval of $\widehat{\pi}_0(t)$ calculated at each day is also shown.

6. CONCLUDING REMARKS

In this article, we propose semiparametric temporal process regression methods. Relative to existing process regression methods, our methods are distinguished by several features. In particular, the baseline probability (as a function of time) is unspecified and is essentially profiled out in the estimation of the regression

coefficient (presumably of chief interest in most studies). The method accommodates dependent right censoring, and it does so through a computationally attractive additive hazards model. In our context, the additive hazard model facilities the calculation of weight function, since the baseline cumulative hazard function cancels out. Moreover, the proposed methods accommodate independent right censoring through imputation rather than a weight function.

The logit link function used in direct binomial regression (e.g., Schieke & Zhang, 2007) may be preferable to the log link function used in the proposed method, since the estimated baseline probability, $\widehat{\pi}_0(t)$ in (10), is not bounded by 1. Several considerations are important in this regard. First, out-of-range fitted values may or may not be a problem; this depends on the data at hand. Second, fitted values are frequently of little or no interest to clinical and epidemiologic investigators. Third, our regression parameter estimator is consistent even if fitted values are out of range, owing to the fact that $\pi_0(t)$ is profiled out of the estimating equation for $\widehat{\boldsymbol{\beta}}$; this phenomenon was demonstrated through simulation in Zhan & Schaubel (2018). Fourth, our model is semiparametric, with an unspecified baseline probability. Direct binomial regression, being fully parametric, requires correct specification of the baseline probability (often of little interest in applications). Use of the log link is central to the computational techniques that enable not specifying the baseline probability in our approach. Use of the log link in (1) enables the use of proportional hazards software (e.g., coxph in R

(Therneau T, 2015) and proc phreg in SAS (SAS Institute Inc.)) after appropriate data augmentation, since (9) is analogous to a weighted Cox score function.

Our methods make the distinction between independent censoring, C_1, and dependent censoring, C_2. This is necessary since the variates play very different roles in our framework. C_1 represent the end of a subject's potential follow-up. This is the case in several existing methods, including the popular subdistribution hazard modeling of Fine & Gray (1999). The methods of Fine & Gray (1999) do not impute unobserved C_1 but, instead, apply a weight function which represents the conditional probability of being uncensored (given that the subject was uncensored at the time of death). This weight function is essentially playing the same role as our imputation of C_1. In contrast, C_2 is a nuisance process, with its corresponding inverse weight seeking to recover the data that would be observed if the process underlying C_2 were absent.

ACKNOWLEDGEMENTS

The authors thank the Editor, Associate Editor, and two Reviewers for their thoughtful and constructive comments, which improved the article considerably.

Supplementary Material

Supplementary material available online includes the proofs of Theorems 1, 2, 3, and 4 from Section 3; R code for carrying out the proposed methods; and additional simulation results.

BIBLIOGRAPHY

Andersen, P. K., Klein, J. P., & Rosthøj, S. (2003). Generalised linear models for correlated pseudoobservations, with applications to multistate models. *Biometrika*, 90, 15–27.

Breslow, N. E. (1972). Contribution to the discussion of the paper by D.R. Cox. *Journal of the Royal Statistical Society, Series B*, 34, 216–217.

Cox, D. R. (1972). Regression models and life tables (with discussion). *Journal of the Royal Statistical Society*, 34, 187–220.

Cox, D. R. (1975). Partial likelihood. *Biometrika*, 62, 269–276.

Fine, J. P. & Gray, R. J. (1999). A proportional hazards model for the subdistribution of a competing risk. *Journal of the American Statistical Association*, 94, 496–509.

Fine, J. P., Yan, J., & Kosorok, M. R. (2004). Temporal process regression. *Biometrika*, 91, 683–703.

Gelber, R. D., Gelman, R. S., & Goldhirsch, A. (1989). A quality-of-life-oriented endpoint for comparing therapies. *Biometrics*, 45, 781–795.

Glasziou, P., Simes, R., & Gelber, R. (1990). Quality adjusted survival analysis. *Statistics in Medicine*, 9, 1259–1276.

Ghosh, D. & Lin, D. Y. (2002). Marginal regression models for recurrent and terminal events. *Statistica Sinica*, 12, 663–688.

Grand, M. K. & Putter, H. (2016). Regression models for expected length of stay. *Statistics in Medicine*, 35, 1178–1192.

Gray, R. J. (1988). A class of k-sample tests for comparing the cumulative incidence of a competing risk. *The Annals of Statistics*, 16, 1141–1154.

Hernán, M. A., Brumback, B., & Robins, J. M. (2000). Marginal structural models to estimate the causal effect of zidovudine on the survival of HIV-positive men. *Epidemiology*, 11, 561–570.

Kalbfleisch, J. D. & Prentice, R. L. (2002). *The Statistical Analysis of Failure Time Data, 2nd ed.*. Jonh Wiley & Sons, Hoboken, New Jersey.

Lin, D. & Ying, Z. (1994). Semiparametric analysis of the additive risk model. *Biometrika*, 81, 61–71.

Pan, Q. & Schaubel, D. E. (2008). Proportional hazards models based on biased samples and estimated selection probabilities. *Canadian Journal of Statistics*, 36, 111–127.

Prentice, R. L., Kalbfleisch, J. D., Peterson Jr, A. V., Flournoy, N., Farewell, V. T., & Breslow, N. E. (1978). The analysis of failure times in the presence of competing risks. *Biometrics*, 34, 541–554.

Robins, J. M. (1993). Information recovery and bias adjustment in proportional hazards regression analysis of randomized trials using surrogate markers. In American Statistical Association (Ed.), *Proceedings of the Biopharmaceutical Section*, American Statistical Association,Alexandria, Virgina, 24–33.

Robins, J. M. & Finkelstein, D. M. (2000). Correcting for noncompliance and dependent censoring in an AIDS clinical trial with inverse probability of censoring weighted (IPCW) log-rank tests. *Biometrics*, 56, 779–788.

Robins, J. M. & Rotnitzky, A. (1992). Recovery of information and adjustment for dependent censoring using surrogate markers. In Jewell, N.P., Dietz, K., & Farewell, V.T. (Eds.) *AIDS Epidemiology*, Birkhäuser, Boston, MA, 297-331.

Ruan, P. K. & Gray, R. J. (2008). Analyses of cumulative incidence functions via nonparametric multiple imputation. *Statistics in Medicine*, 27, 5709–5724.

Rubin, D. B. (2004). *Multiple Imputation for Nonresponse in Surveys*, Vol. 81 of Wiley Classics Library, John Wiley & Sons, Hoboken, New Jersey.

SAS Institute Inc., Version 9.3, Cary, NC, USA.

Sasieni, P. (1993). Maximum weighted partial likelihood estimators for the cox model. *Journal of the American Statistical Association*, 88, 144–152.

Schaubel, D. E. & Zhang, M. (2010). Estimating treatment effects on the marginal recurrent event mean in the presence of a terminating event. *Lifetime Data Analysis*, 16, 451–477.

Scheike, T. H. & Zhang, M. J. (2007). Direct modelling of regression effects for transition probabilities in multistate models. *Scandinavian Journal of Statistics*, 34, 17–32.

Therneau, T. (2015). A Package for Survival Analysis in S. version 2.38, https://CRAN.R-project.org/package=survival.

Wang, N. & Robins, J. M. (1998). Large-sample theory for parametric multiple imputation procedures. *Biometrika*, 85, 935–948.

Yan, J. & Fine, J. P. (2005). Functional association models for multivariate survival processes. *Journal of the American Statistical Association*, 100, 184–196.

Yan, J. & Huang, J. (2009). Partly functional temporal process regression with semiparametric profile estimating functions. *Biometrics*, 65, 431–440.

Yin, G. (2007). Model checking for additive hazards model with multivariate survival data. *Journal of Multivariate Analysis*, 98, 1018–1032.

Zhan, T. (2017). Ph.D Dissertation, University of Michigan, Department of Biostatistics.

Zhan, T., & Schaubel, D. E. (2018). Semiparametric temporal process regression of survival-out-of-hospital. *Lifetime Data Analysis*, 25, 1–19.

Zhang, M. & Schaubel, D. E. (2011). Estimating differences in restricted mean lifetime using observational data subject to dependent censoring. *Biometrics*, 67, 740–749.

Zhao, Y. & Wang, H. (2008). Empirical likelihood inference for the regression model of mean quality-adjusted lifetime with censored data. *Canadian Journal of Statistics*, 36, 463–478.

Received 9 July 2009

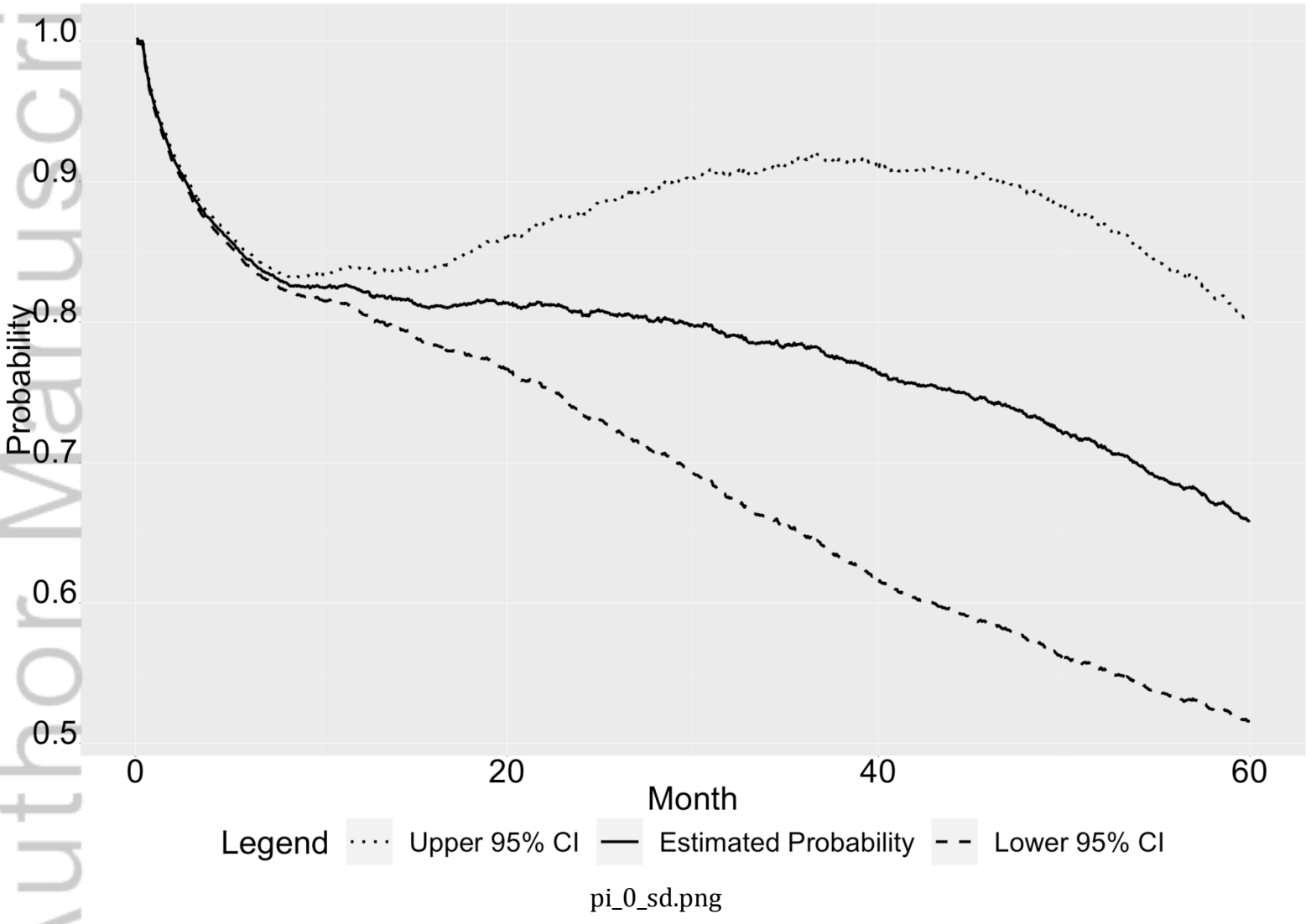
Accepted 8 July 2010

TABLE 2: Simulations results for $\widehat{P}_0^5(50)$ with random censoring time based on $M = 5$, and $5,000$ replicates.

Censoring	$\Pi_0(50)$	n	$\boldsymbol{\beta}_0$	Weights	Bias	ASE	ESD	CP
Heavy	11.812	500	(0.916, -0.916)	1	-1.871	0.748	0.715	0.296
				W^A	0.069	0.968	0.888	0.963
				W^B	-0.319	0.899	0.864	0.931
			(0.405, -0.405)	1	-1.906	0.750	0.721	0.286
				W^A	0.063	0.957	0.886	0.963
				WB	-0.336	0.896	0.840	0.932
		1,000	(0.916, -0.916)	1	-1.896	0.531	0.512	0.062
				W^A	0.015	0.683	0.627	0.962
				W^B	-0.324	0.641	0.596	0.927
			(0.405, -0.405)	1	-1.901	0.533	0.516	0.058
				W^A	0.012	0.673	0.623	0.966
				W^B	-0.311	0.638	0.586	0.932
Light	11.812	500	(0.916, -0.916)	1	-1.162	0.852	0.812	0.714
				W^A	0.138	1.009	0.931	0.961
				W^B	-0.204	0.953	0.907	0.943
			(0.405, -0.405)	1	-1.153	0.858	0.825	0.719
				W^A	0.128	1.008	0.947	0.960
				W^B	-0.201	0.955	0.904	0.949
		1,000	(0.916, -0.916)	1	-1.193	0.604	0.572	0.491
				W^A	0.089	0.714	0.657	0.965
				W^B	-0.186	0.678	0.629	0.953
			(0.405, -0.405)	1	-1.179	0.608	0.580	0.506
				W^A	0.078	0.711	0.665	0.962
				W^B	-0.175	0.678	0.625	0.954

TABLE 3: Analysis of Liver transplant data: Covariate effects on being active on the waiting list and being alive (based on $M = 1$ imputation)

Covariate	Value	$\widehat{\boldsymbol{\beta}}$	$\widehat{\mathrm{SE}}(\widehat{\boldsymbol{\beta}})$	p	$\exp(\widehat{\boldsymbol{\beta}})$
Blood type (v.s. O)	A	0.0113	0.0140	0.421	1.01
	AB	-0.0057	0.0390	0.8843	0.99
	B	0.0160	0.0202	0.4265	1.02
Gender	Female	0.0234	0.0131	0.0748	1.02
Race (v.s. White)	Black	0.0036	0.0257	0.8898	1.00
	Hispanic	0.0266	0.0182	0.1441	1.03
	Asian	0.0284	0.0291	0.3287	1.023
	Other	-0.0396	0.0698	0.5707	0.96
Diagnosis	HCV	-0.0245	0.0215	0.2544	0.98
	Noncholestatic	0.0388	0.0204	0.0568	1.04
	Cholestatic	0.0450	0.0256	0.0792	1.05
	Acute hepatic necrosis	0.1116	0.0429	0.0094*	1.12
	Metastatic disease	0.0246	0.0500	0.6229	1.03
	Malignant neoplasm	-0.1130	0.0464	0.0149*	0.89
BMI (v.s. $(20, 25)$)	$[0, 20]$	0.0178	0.0277	0.5217	1.02
	$[25, 30)$	-0.0011	0.0159	0.9436	1.000
	$[30, \infty)$	-0.0513	0.0169	0.0024*	0.95
Hospitalization status	ICU	0.1652	0.0500	0.001*	1.18
(v.s. not hospitalized)	Not ICU	-0.0279	0.0323	0.3868	0.97
Age	per 5 years	-0.0205	0.0030	$<.0001$*	0.98
Region (v.s. 5)	1	-0.1924	0.0453	$<.0001$*	0.83
	2	0.0204	0.0213	0.3394	1.02
	3	-0.0259	0.0276	0.3493	0.97
	4	-0.0423	0.0222	0.0572	0.96
	6	-0.0646	0.0479	0.1772	0.94
	7	0.0611	0.0248	0.0135*	1.06
	8	-0.0511	0.0304	0.0932	0.95
	9	-0.0130	0.0232	0.5771	0.99
	10	-0.0587	0.0285	0.0397*	0.94
	11	-0.1753	0.0328	$<.0001$*	0.84
MELD	per unit score	-0.0078	0.0016	$<.0001$*	0.99
Albumin	per mmol/L	0.0850	0.0107	$<.0001$*	1.09
Sodium	per g/dL	0.0026	0.0018	0.1481	1.00



pi_0_sd.png