Biometrics 65, 753–761
September 2009
DOI: 10.1111/j.1541-0420.2008.01157.x

Flexible Estimation of Differences in Treatment-Specific Recurrent Event Means in the Presence of a Terminating Event

Qing Pan[1,*] and Douglas E. Schaubel[2,]**

[1]Department of Statistics, George Washington University, Washington, DC 20052, U.S.A.
[2]Department of Biostatistics, University of Michigan, Ann Arbor, Michigan 48109, U.S.A.

[*]*email:* qpan@gwu.edu

[**]*email:* deschau@umich.edu

Summary. In this article, we consider the setting where the event of interest can occur repeatedly for the same subject (i.e., a recurrent event; e.g., hospitalization) and may be stopped permanently by a terminating event (e.g., death). Among the different ways to model recurrent/terminal event data, the marginal mean (i.e., averaging over the survival distribution) is of primary interest from a public health or health economics perspective. Often, the difference between treatment-specific recurrent event means will not be constant over time, particularly when treatment-specific differences in survival exist. In such cases, it makes more sense to quantify treatment effect based on the cumulative difference in the recurrent event means, as opposed to the instantaneous difference in the rates. We propose a method that compares treatments by separately estimating the survival probabilities and recurrent event rates given survival, then integrating to get the mean number of events. The proposed method combines an additive model for the conditional recurrent event rate and a proportional hazards model for the terminating event hazard. The treatment effects on survival and on recurrent event rate among survivors are estimated in constructing our measure and explain the mechanism generating the difference under study. The example that motivates this research is the repeated occurrence of hospitalization among kidney transplant recipients, where the effect of expanded criteria donor (ECD) compared to non-ECD kidney transplantation on the mean number of hospitalizations is of interest.

Key words: Additive rates model; Competing risks; Marginal mean; Proportional hazards model; Rate regression; Semiparametric model.

1. Introduction

Recurrent events are frequently of interest in clinical and epidemiologic studies. Examples include repeated infections among HIV patients and multiple hospitalizations in a health services utilization study. A large variety of semiparametric recurrent event models exist in the literature. These methods can generally be classified as intensity models or mean/rate models. Intensity models consider the instantaneous probability of event occurrence conditional on the event history (e.g., Prentice, Williams, and Peterson, 1981; Andersen and Gill, 1982), while mean/rate (the rate being the derivative of mean) models consider the marginal mean number of events (e.g., Pepe and Cai, 1993; Lawless and Nadeau, 1995; Lin et al., 2000). Depending on the assumed form of the covariate effects, recurrent event models can be classified as multiplicative or additive. Proportional means/rates models assume covariate effects on a multiplicative scale, while additive models assume them on an additive scale.

Often in biomedical applications, the recurrent event sequence can be stopped permanently by a terminating event (e.g., death). Various approaches have been proposed for modeling recurrent events in the presence of a terminating event (e.g., Cook and Lawless, 1997; Li and Lagakos, 1997; Ghosh and Lin, 2000, 2002), and this area has attracted much attention recently (e.g., Huang and Wang, 2004; Liu, Wolfe, and Huang, 2004; Ye, Kalbfleisch, and Schaubel, 2007). The existing approaches generally fall into one of three categories. First, there are methods for modeling the marginal mean number of events (e.g., Ghosh and Lin, 2000, 2002). In this case, the mean averages over surviving and deceased subjects. In the second category of methods, one models the conditional recurrent event rate given survival. This is among the approaches suggested by Cook and Lawless (1997), and was mentioned by Lin et al. (2000). This method has been applied quite frequently (e.g., Schaubel and Cai, 2005). A variation of this approach employs a latent (frailty) variable (e.g., Liu et al., 2004; Ye et al., 2007), conditional on which the recurrent and terminal events are assumed independent. The marginal and conditional methods explicitly acknowledge the fact that subjects no longer experience recurrent events after death and that time to death may differ among the groups being compared. The third approach considers death to be a censoring event, with the recurrent events essentially being a

latent process unobservable after death (e.g., Ghosh and Lin, 2003).

The example motivating our method is to compare the mean number of hospitalizations for transplant recipients with an expanded criteria donor (ECD) kidney and patients transplanted with a non-ECD kidney. An ECD is defined (Port et al., 2002) as a deceased donor either (i) age $\geqslant$60, or (ii) age 50–59 and with at least two of the following three characteristics: hypertensive, serum creatinine concentration >1.5mg/dl, or death due to stroke. The ECD (0, 1) classification is a well-accepted quality index for donated kidneys in the nephrology community. In fact, wait-listed patients willing to accept an ECD kidney are essentially listed separately and generally have reduced waiting time until transplant. The relative frequency of ECD transplantation has increased over time, and several authors have examined the impact of an ECD on the posttransplant death hazard. Our analysis targets the impact of ECD transplantation on the mean number of posttransplant hospitalizations. Hospitalization frequency is a composite index of a patient's health status and resource utilization, and therefore serves as a concrete and objective measure of posttransplant performance. Among transplant recipients, hospitalizations can occur repeatedly for the same patient and are often terminated by death.

Ghosh and Lin (2000) proposed nonparametric tests for treatment-specific differences with respect to the marginal mean number of recurrent events. The Kaplan–Meier estimator for survival probabilities and a Nelson–Aalen estimator of the conditional recurrent event rate function given survival are integrated to calculate the recurrent event means. Ghosh and Lin (2002) later proposed a semiparametric proportional means model. The recurrent event rate after death was still taken to be zero, and the marginal rates in the outcome are implicitly averaged over the recurrent event rates of living and deceased subjects.

Often investigators are more interested in the absolute difference between recurrent event means (as opposed to their ratio), which suggests using an additive (as opposed to multiplicative) model. Schaubel, Zeng, and Cai (2006) extended the method of Lin and Ying (1994) to develop an additive rates model, but it was not designed to handle terminating events. One could develop an extension of Schaubel et al. (2006) to accommodate terminating events. However, treatment-specific differences in mean number of events are often not constant over time, particularly when the treatment-specific survival functions differ. That is, as follow-up time, t, increases, due to subjects dying, the composition of the study population will shift and the pattern of the shift will be treatment-specific if treatment affects survival. Therefore, treatment effects on the recurrent event mean would not be expected to be constant over time in the presence of treatment effects on survival. In cases where the treatment effect depends on follow-up time, the cumulative effect is of much greater interest for patient decision making (at time $t = 0$) than the instantaneous effect.

In this article, we propose a novel semiparametric method for comparing treatment-specific recurrent event mean functions in the presence of a terminating event. The proposed method estimates the treatment effect as a process over time, without forcing the effect to be a constant difference or ratio. The method combines a proportional hazards model for the terminating event and an additive model for the conditional recurrent event rate given survival, with the treatment effects measured by differences in treatment-specific recurrent event means.

In comparing treatments, investigators are often interested in differences between treatment-specific means, which motivates our measure. However, under this method, an increase in the event mean could be due to patients surviving longer, or could be due to patients experiencing events at an increased rate while they survive, and the proposed estimator alone cannot distinguish between these two phenomena. Researchers wanting to further examine the nature of the treatment effect on the event mean would examine separately the treatment effects on the terminating event hazard and the conditional event rates given survival. This additional step is always necessary even when the proposed measure turns out to be nonsignificant because a lower survival probability can serve to offset higher event rate among survivors, and hence cancel out the treatment effect on the mean.

The remainder of this article is organized as follows. Section 2 describes the proposed treatment effect measure and methods of estimation. Asymptotic properties of the estimator are listed in Section 3, with proofs provided in Web Appendices A and B. Section 4 evaluates the performance of the asymptotic results in moderate size samples. We compare mean numbers of hospitalizations among kidney transplant recipients using the proposed methods in Section 5. A discussion and some concluding remarks are contained in Section 6.

2. Proposed Methods

In this section, we describe the proposed methods, after establishing notation and explaining the data structure of interest.

2.1 *Notation and Setup*

We begin by setting up the necessary notation. We use D_i for the time of the terminating event (e.g., death), and C_i for the censoring time. The observed counting process for the terminating event is represented by $N_i^D(t) = I(D_i \leqslant t,\ D_i < C_i)$. We let $Y_i(t) = I(C_i \wedge D_i \geqslant t)$ be the at risk indicator for subject i at time t, and set $\widehat{\pi}(t) = n^{-1}\sum_{i=1}^{n} Y_i(t)$. The number of recurrent events as of time t is represented by $N_i^{R*}(t) = \int_0^t dN_i^{R*}(s)$, where $dN_i^{R*}(s) = N_i^{R*}(s) - N_i^{R*}(s^-)$. In the data structure under consideration, $N_i^{R*}(t) = N_i^{R*}(t \wedge D_i)$; that is, recurrent events cannot occur after death. What we observe is a quantity subject to right censoring, $N_i^R(t) = N_i^{R*}(t \wedge D_i \wedge C_i)$. For the ith subject ($i = 1, \ldots, n$), we let Z_i and X_i denote the vector of covariates for the terminating event model and recurrent event model, respectively. We set $Z_i = (Z_{i1},\ Z_{i2}')'$, $X_i = (X_{i1},\ X_{i2}')'$, where Z_{i1}, X_{i1} are (1/0) indicators for the treatment/placebo and Z_{i2} and X_{i2} represent adjustment covariates. Correspondingly, we let $\beta_0 = (\beta_{01},\ \beta_{02}')'$, and $\theta_0 = (\theta_{01},\ \theta_{02}')'$ represent the regression coefficients for Z_i and X_i. In addition, for convenience we denote the covariate vectors for a treated subject by $Z_i^1 = (1,\ Z_{i2}')'$ and $X_i^1 = (1,\ X_{i2}')'$. Similarly, for a placebo subject, $Z_i^0 = (0,\ Z_{i2}')'$ and $X_i^0 = (0,\ X_{i2}')'$. The objective is to compare the treatment and placebo with respect to the marginal mean number of recurrent events. That is, the outcome being modeled is the mean number of recurrent events $E\{N_i^{R*}(t) \mid Z_i\}$, which is the integral of recurrent event

rate function over time $E\{N_i^{R*}(t) \mid Z_i\} = \int_0^t E\{dN_i^{R*}(s) \mid Z_i\}$. It should be mentioned that we are not modeling the recurrent event intensity function, $E\{dN_i^{R*}(t) \mid Z_i, \mathcal{N}_i^{R*}(t^-)\}$, where $\mathcal{N}_i^{R*}(t) = \{N_i^{R*}(s); s \in (0,t]\}$, representing the event history up to time t, an approach considered for example by Andersen and Gill (1982).

As described in Section 1, in the presence of a terminating event, the recurrent event mean is a function of survival probability and the conditional recurrent event rate given survival. In estimating the mean function, it is natural to model these two entities separately. We denote the conditional recurrent event rate given survival by

$$dR_i(t) = E\big\{dN_i^{R*}(t) \,\big|\, D_i > t\big\},$$

with the recurrent event mean function then given by

$$\mu_i(t) = \int_0^t S_i(u) dR_i(u),$$

where $S_i(t) = Pr(D_i > t)$. We assume that the terminating event hazard function follows a proportional hazards model,

$$d\Lambda_i(t) = d\Lambda(t \mid Z_i) = d\Lambda_0(t) e^{\beta_0' Z_i}, \tag{1}$$

where $d\Lambda_0(t)$ is the unspecified baseline hazard function and the vector β_0 represents the true value of the regression coefficients. The parameter β_0 is estimated by $\widehat{\beta}$, the solution to the partial likelihood score function, $U^D(\beta) = 0$, where

$$\begin{aligned} U^D(\beta) &= \sum_{i=1}^n \int_0^\tau \{Z_i - \overline{Z}(t;\beta)\} dN_i^D(t) \\ \overline{Z}(\beta;t) &= \frac{S^{(1)}(t;\beta)}{S^{(0)}(t;\beta)} \\ S^{(k)}(t;\beta) &= n^{-1} \sum_{i=1}^n Y_i(t) e^{\beta' Z_i} Z_i^{\otimes k}, \end{aligned}$$

with $Z_i^{\otimes 0} = 1$, $Z_i^{\otimes 1} = Z_i$, and $Z_i^{\otimes 2} = Z_i Z_i'$. The Breslow–Aalen baseline hazard estimator, $\widehat{\Lambda}_0(t;\widehat{\beta})$, is employed, where

$$\widehat{\Lambda}_0(t;\beta) = n^{-1} \sum_{i=1}^n \int_0^t S^{(0)}(s;\beta)^{-1} dN_i^D(s).$$

The subject-specific survival function is then estimated by $\widehat{S}_i(t) = \exp\{-\widehat{\Lambda}_0(t) e^{\widehat{\beta}' Z_i}\}$.

Both multiplicative and additive models have been used in modeling the recurrent event mean. For certain outcomes such as costs or number of hospitalizations, investigators are typically more interested in the absolute (as opposed to relative) dollar amount or hospital admission numbers. As such, we use an additive model to estimate the recurrent event rates among surviving subjects,

$$dR_i(t) = dR(t \mid X_i) = dR_0(t) + \theta_0' X_i dt, \tag{2}$$

where $dR_0(t)$ is the baseline recurrent event rate function and the vector θ_0 represents the true additive effects of the corresponding covariate vector X_i. Usually, there are common covariates affecting both the terminating event hazard and recurrent event rates. As such, the covariate vectors Z_i and X_i usually overlap and will often be identical. Adapting the model of Schaubel et al. (2006) to the recurrent/terminal event setting, we estimate θ_0 as follows:

$$\begin{aligned} \widehat{\theta} &= \widehat{B}^{-1} U^R \\ \widehat{B} &= \sum_{i=1}^n \int_0^\tau Y_i(s) \{X_i - \overline{X}(s)\}^{\otimes 2} ds \\ \overline{X}(s) &= \widehat{\pi}(s)^{-1} n^{-1} \sum_{i=1}^n Y_i(s) X_i \\ U^R &= \sum_{i=1}^n \int_0^\tau \{X_i - \overline{X}(t)\} dN_i^R(t). \end{aligned}$$

Models for the conditional event rate, $dR_i(t)$, and the terminating event hazard, $d\Lambda_i(t)$, assume covariate effects which are constant over time, t. Two points are important in this regard. First, unlike $\mu_i(t)$, covariate effects on both $dR_i(t)$ and $d\Lambda_i(t)$ will often be constant over time in practice. In a sense, the recurrent event rate conditional on survival, $dR_i(t)$, is an easier measure to model directly, compared to the marginal recurrent event mean, because the survival population is more homogeneous than the combined population of both living and deceased subjects. The model $d\Lambda_i(t)$ is a standard Cox model frequently employed in biomedical studies, while the quantity $dR_i(t)$ is analogous to the cause-specific hazard in competing risk studies. Second, models (1) and (2) could always be extended to allow for time-varying effects, and the procedures proposed in this article would still be applicable.

2.2 *Proposed Measure and its Estimator*

We propose a treatment effect measure, which is the difference in treatment-specific marginal recurrent event means. The means are given by

$$\begin{aligned} \mu_1(t) &= E\left[N_i^R(t) \,\middle|\, Z_{i1} = 1\right] \\ \mu_0(t) &= E\left[N_i^R(t) \,\middle|\, Z_{i1} = 0\right]. \end{aligned}$$

We model $dR_i(t)$ and $d\Lambda_i(t)$ separately, compute fitted means, then average over the observed adjustment covariates. From this perspective, it is useful to write the treatment-specific marginal means as

$$\begin{aligned} \mu_j(t) &= E\Big[E\Big[N_i^R(t) \,\Big|\, Z_i^j, X_i^j\Big]\Big] \\ &= E\left[\int_0^t S\big(r \,\big|\, Z_i^j\big) dR\big(r \,\big|\, X_i^j\big)\right], \end{aligned}$$

for $j = 0, 1$. For the iterated expectations above, the inner one is the expected value of $N_i^R(t)$ conditional on a specific set of Z_i^j, X_i^j values and the outer marginal expectation is taken with respect to the marginal distribution of adjustment covariates $(Z_{i2}', X_{i2}')'$. By substituting in survival and conditional rate function estimators, we obtain the proposed treatment-specific mean estimators:

$$\begin{aligned} \widehat{\mu}_1(t) &= n^{-1} \sum_{i=1}^n \int_0^t \widehat{S}\big(r \,\big|\, Z_i^1\big) \big\{ d\widehat{R}_0(r) + \big(\widehat{\theta}_1 + X_{i2}' \widehat{\theta}_2\big) dr \big\} \\ \widehat{\mu}_0(t) &= n^{-1} \sum_{i=1}^n \int_0^t \widehat{S}\big(r \,\big|\, Z_i^0\big) \big\{ d\widehat{R}_0(r) + X_{i2}' \widehat{\theta}_2 \, dr \big\}, \end{aligned}$$

where the baseline rate estimator is given by $\widehat{R}_0(t) = n^{-1}\sum_{i=1}^{n}\int_0^t \widehat{\pi}(s)^{-1}Y_i(s)\{dN_i^R(s) - \widehat{\theta}'X_i ds\}$. The treatment effect on the recurrent event mean can be measured by

$$\psi(t) = \mu_1(t) - \mu_0(t),$$

which can be estimated as

$$\widehat{\psi}(t) = \widehat{\mu}_1(t) - \widehat{\mu}_0(t)$$

for $t \in (0,\, \tau]$ using $\widehat{S}(t \,|\, Z_i)$ from the proportional hazards model and $d\widehat{R}_0(t), \widehat{\theta}_1$ and $\widehat{\theta}_2$ from the additive conditional rates model.

2.3 *Analyzing Treatment Effects*

Based on the proposed measure alone, it is not clear if an estimated treatment effect is the result of treatment-specific differences in survival and/or from treatment-specific differences in the conditional event rate given survival. For example, the marginal recurrent event mean could be greater for treated patients than placebo patients because either (i) conditional on survival the event rates are equal, but treated subjects live longer; or (ii) treated and placebo subjects have approximately equal survival, but $dR(t \,|\, X_i^1) > dR(t \,|\, X_i^0)$. Therefore, the hazard ratio for the terminating event (e^{β_1}) and the rate difference for recurrent events among survivors (θ_1) provide important supplemental information. Collectively, the three estimators $\widehat{\psi}(t), e^{\widehat{\beta}_1}$ and $\widehat{\theta}_1$ describe treatment effects from each of these angles.

3. Asymptotic Properties

In this section, we describe the essential asymptotic properties of the proposed estimators. We begin by listing the assumed regularity conditions, for $i = 1, \ldots, n$.

(a) $\{N_i^{R*}(\cdot),\, D_i,\, C_i,\, Z_i,\, X_i\}$ are independent and identically distributed.
(b) $E[dN_i^{R*}(t) \,|\, D_i > t,\, C_i > t,\, X_i] = E[dN_i^{R*}(t) \,|\, D_i > t,\, X_i]$.
(c) $\lim_{\delta\to 0}\frac{1}{\delta}Pr\{t \leqslant D_i < t+\delta \,|\, D_i > t, C_i > t, Z_i\} = \lim_{\delta\to 0}\frac{1}{\delta}Pr\{t \leqslant D_i < t+\delta \,|\, D_i > t, Z_i\}$.
(d) $Pr(Y_i(\tau) = 1) > 0$.
(e) $\int_0^\tau d\Lambda_0(t) < \infty, \int_0^\tau dR_0(t) < \infty$ and $N_i^R(\tau) < \infty$.
(f) Elements of Z_{i2} and X_{i2} are bounded almost surely.
(g) Positive-definiteness of the matrices, $A(\beta)$ and B, where

$$A(\beta) = E\left[\int_0^\tau \{Z_i - \overline{z}(t;\beta)\}^{\otimes 2}Y_i(t)e^{\beta' Z_i}\, d\Lambda_0(t)\right]$$
$$\overline{Z}(t;\beta) \xrightarrow{a.s.} \overline{z}(t;\beta)$$
$$B = E\left[\int_0^\tau Y_i(s)\{X_i - \overline{x}(s)\}^{\otimes 2}ds\right]$$
$$\overline{X}(t) \xrightarrow{a.s.} \overline{x}(t).$$

Condition (a) is the basis for the central limit theorem and is usually satisfied; an exception would be clustered data. Conditions (b) and (c) correspond to independent censoring assumptions for the recurrent and terminating event processes, respectively. Condition (d) is a standard identifiability condition.

It should be noted that our methods also assume that there are no unmeasured factors which predict both $dR_i(t)$ and $d\Lambda_i(t)$. The potential impact resulting from violations of the no-unmeasured-predictors assumption is discussed and evaluated empirically in Section 4.

We now describe the essential asymptotic results for the proposed estimators, with proofs outlined in Web Appendices A and B.

THEOREM 1. *Under conditions (a) to (g), $\widehat{\psi}$ is a uniformly consistent estimator of ψ; that is, $\widehat{\psi}(t) \xrightarrow{a.s.} \psi(t)$ for $t \in (0,\, \tau]$.*

The proof of Theorem 1 proceeds through a Taylor series expansion, followed by repeated application of the uniform strong law of large numbers.

THEOREM 2. *Under conditions (a) to (g), $n^{\frac{1}{2}}\{\widehat{\psi}(t) - \psi(t)\}$ converges weakly to a zero-mean Gaussian process with covariance function $E[\{\xi_{i1\cdot}(s;\beta_0,\, \theta_0) - \xi_{i0\cdot}(s;\beta_0,\, \theta_0)\}\,\{\xi_{i1\cdot}(t;\beta_0,\, \theta_0) - \xi_{i0\cdot}(t;\beta_0,\, \theta_0)\}]$, where*

$$\begin{aligned}
\xi_{ij\cdot}(t;\beta,\theta) &= \xi_{ij1}(t;\beta,\theta) \\
&\quad + \xi_{ij2}(t;\beta,\theta) + \xi_{ij3}(t;\beta,\theta) + \xi_{ij4}(t;\beta,\theta) \\
\xi_{ij1}(t;\beta,\theta) &= -E\left[e^{\beta' Z_k^0}\int_0^t S\big(u^- \,\big|\, Z_k^0\big)\right. \\
&\quad \left.\times \int_0^u \big\{Z_k^0 - \overline{z}(r;\beta)\big\}' d\Lambda_0(r) dR\big(u \,\big|\, X_k^j\big)\right] \\
&\quad \times A(\beta)^{-1}U_i^D(\beta) \\
\xi_{ij2}(t;\beta,\theta) &= E\left[\int_0^t S\big(u^- \,\big|\, Z_k^j\big)\big\{X_k^j - \overline{x}(u)\big\}' du\right] B^{-1}U_i^R(\theta) \\
\xi_{ij3}(t;\beta,\theta) &= \int_0^t E\big\{S\big(u^- \,\big|\, Z_k^j\big)\big\}\pi(u)^{-1} dM_i^R(u;\theta) \\
\xi_{ij4}(t;\beta,\theta) &= -\int_0^t E\big[e^{\beta' Z_k^j}\big\{\mu\big(t \,\big|\, Z_k^j, X_k^j\big) - \mu\big(r \,\big|\, Z_k^j, X_k^j\big)\big\}\big] \\
&\quad \times \frac{dM_i^D(r;\beta)}{s^{(0)}(r;\beta)},
\end{aligned}$$

where k indexes a subject, with subject-specific asymptotic score contributions given by

$$\begin{aligned}
U_i^D(\beta) &= \int_0^\tau \{Z_i - \overline{z}(t;\beta)\} dM_i^D(t;\beta) \\
dM_i^D(t;\beta) &= dN_i^D(t) - Y_i(t)e^{\beta' Z_i}\, d\Lambda_0(t) \\
U_i^R(\theta) &= \int_0^t \{X_i - \overline{x}(u)\} dM_i^R(u;\theta) \\
M_i^R(t;\theta) &= N_i^R(t) - \int_0^t Y_i(u)\{dR_0(u) + \theta' X_i du\} \\
s^{(d)}(t;\beta) &= \lim_{n\to\infty} S^{(d)}(t;\beta),
\end{aligned} \tag{3}$$

for $d = 0,\, 1,\, 2$. Because $n^{\frac{1}{2}}(\widehat{\psi} - \psi) = n^{\frac{1}{2}}(\widehat{\mu}_1 - \mu_1) - n^{\frac{1}{2}}(\widehat{\mu}_0 - \mu_0)$, we work on $n^{\frac{1}{2}}(\widehat{\mu}_1 - \mu_1)$ and $n^{\frac{1}{2}}(\widehat{\mu}_0 - \mu_0)$ separately as follows:

$$
\begin{aligned}
&n^{\frac{1}{2}}\{\widehat{\mu}_j(t;\widehat{\beta},\widehat{\theta})-\mu_j(t)\}\\
&\quad= n^{\frac{1}{2}}\{\widehat{\mu}_j(t;\widehat{\beta},\widehat{\theta})-\widehat{\mu}_j(t;\beta_0,\widehat{\theta})\}\\
&\qquad+ n^{\frac{1}{2}}\{\widehat{\mu}_j(t;\beta_0,\widehat{\theta})-\widehat{\mu}_j(t;\beta_0,\theta_0)\}\\
&\qquad+ n^{-\frac{1}{2}}\sum_{i=1}^{n}\int_0^t \widehat{S}\big(r^-\,\big|\,Z_i^j\big)\{d\widehat{R}_0(r;\theta_0)-dR_0(r)\}\\
&\qquad+ n^{-\frac{1}{2}}\sum_{i=1}^{n}\int_0^t \big\{\widehat{S}\big(r^-;\beta_0\,\big|\,Z_i^j\big)-S\big(r^-\,|\,Z_i^j\big)\big\}\\
&\qquad\times\big\{dR_0(r)+\theta_0' X_i^j dr\big\}
\end{aligned}
$$

for $j = 0, 1$. Taylor series expansions are applied, along with several applications of the strong law of large numbers (Sen and Singer, 1993). Results from each of the four parts are combined to show that $n^{\frac{1}{2}}\{\widehat{\mu}_j(t)-\mu_j(t)\}$ is asymptotically equivalent to $n^{-\frac{1}{2}}\sum_{i=1}^{n}\xi_{ij\cdot}(t;\beta_0,\theta_0)$. After taking the difference of $n^{\frac{1}{2}}(\widehat{\mu}_1-\mu_1)$ and $n^{\frac{1}{2}}(\widehat{\mu}_0-\mu_0)$, $n^{-\frac{1}{2}}\sum_{i=1}^{n}\{\xi_{i1\cdot}(t;\beta_0,\theta_0)-\xi_{i0\cdot}(t;\beta_0,\theta_0)\}$ can be shown to converge to a zero-mean normal distribution for fixed t by the central limit theorem. A demonstration of tightness completes the proof of weak convergence using results from empirical process theory (Pollard, 1990; van der Vaart and Wellner, 1996). The $\xi_{ij\cdot}(s;\beta_0,\,\theta_0)$ quantities can be consistently estimated by replacing all limiting values by their empirical counterparts, then averaging across $i = 1, \ldots, n$.

4. Simulation Study

For each data configuration, we generated $n = 200$ independent and identically distributed subjects with both terminating and recurrent events. The terminating event hazard follows the following proportional hazards model,

$$d\Lambda(t\,|\,Z_i) = d\Lambda_0(t)e^{\beta_1 Z_{i1}+\beta_2 Z_{i2}},$$

where Z_{i1} (treatment) is distributed as Bernoulli (0.5), the adjustment covariate Z_{i2} follows a uniform (0, 10) distribution and $d\Lambda_0(t) = 0.04$. We set the coefficient β_1 at 0.5 and 1, to examine scenarios of low or high treatment effect on survival. Censoring is uniformly distributed on (0, 20), which leads to an average censoring percentage of approximately 42%. The recurrent events follow a Poisson process, with rate function

$$dR(t\,|\,X_i,Q_i) = dR_0(t)+Q_i dt+\theta' X_i dt, \tag{4}$$

where Q_i follows a gamma distribution with mean 0.25 and variance $V(Q_i)$, and $V(Q_i) = 0.5$ or 1. The Q_i variate represents a frailty term shared by all the recurrent event times for the same subject and may be thought of as an unmeasured predictor. Note that the frailty term Q_i does not affect $d\Lambda(t\,|\,Z_i)$ and the assumptions of our proposed methods hold under this setup. The above model was simulated by generating gap times between successive events as:

$$T_{i,j+1} = T_{i,j}-\log(U_{ij})\{dR_0(t)+Q_i+\theta' X_i\}^{-1},$$

for $j = 0, \ldots, 50$, where $U_{ij}\sim$ Unif (0, 1), $X_i = Z_{i1}$ and $T_{i,0}\equiv 0$. We varied the baseline recurrent event rate from 0.125 to 0.25. The covariate X_i is the same as Z_{i1} in the proportional hazards model for the terminating event, both representing the treatment or exposure of interest. The regression coefficient for X_i is set at $\theta = 0.5$.

Table 1
Simulation results: performance of proposed estimator

β_1	$V(Q_i)$	$dR_0(t)$	t	$\psi(t)$	Bias	ESD	ASE	CP
1	0.5	0.25	5	1.16	0.02	0.34	0.34	0.95
			10	1.01	0.05	0.57	0.56	0.95
			15	0.53	0.06	0.74	0.74	0.95
		0.125	5	1.28	0.01	0.34	0.33	0.95
			10	1.34	0.02	0.54	0.54	0.95
			15	1.05	0.04	0.70	0.71	0.95
	1	0.25	5	1.16	−0.01	0.45	0.43	0.95
			10	1.01	0.00	0.74	0.71	0.94
			15	0.53	0.03	0.96	0.91	0.94
		0.125	5	1.28	0.00	0.44	0.43	0.96
			10	1.34	0.02	0.71	0.69	0.95
			15	1.05	0.05	0.89	0.89	0.95
0.5	0.5	0.25	5	1.71	−0.02	0.36	0.35	0.96
			10	2.42	−0.03	0.62	0.61	0.96
			15	2.65	−0.03	0.83	0.84	0.97
		0.125	5	1.76	−0.02	0.33	0.34	0.96
			10	2.58	−0.04	0.58	0.59	0.96
			15	2.91	−0.03	0.77	0.80	0.97
	1	0.25	5	1.71	0.02	0.47	0.45	0.95
			10	2.42	0.03	0.80	0.77	0.95
			15	2.65	0.05	1.04	1.02	0.95
		0.125	5	1.76	−0.03	0.47	0.44	0.95
			10	2.58	−0.05	0.80	0.75	0.95
			15	2.91	−0.04	1.05	0.99	0.95

ESD = empirical standard deviation.
ASE = average asymptotic standard error.
CP = coverage probability.

Table 1 lists the performance of our proposed estimator $\widehat{\psi}(t)$ in eight scenarios of different β_1, $V(Q_i)$, and $dR_0(t)$ combinations. The average observed number of recurrent events per subject ranges from 3.0 to 4.4. Three evenly spaced time points 5, 10, and 15 are picked to examine the performance of the estimators at early, middle, and late follow-up times. For all configurations examined, $\widehat{\psi}(t)$ is very close to their true values (obtained by numerical integration), and average asymptotic standard errors (ASEs) agree well with empirical standard deviations (ESDs). Correspondingly, empirical coverage probabilities (CPs) are close to the nominal value of 0.95. Overall, the asymptotic properties appear to be applicable to moderate size samples based on the simulation results.

Next, we evaluated the sensitivity of our proposed methods to the no-unmeasured-predictors assumption. Specifically, we set up a model with the gamma frailty, Q_i, affecting both $d\Lambda_i(t)$ and $dR_i(t)$, in violation of our underlying assumptions. The simulated proportional hazards model (1) changes to

$$d\Lambda(t\,|\,Z_i) = Q_i e^{\beta' Z_i}\,d\Lambda_0(t), \tag{5}$$

such that Q_i represents an unmeasured covariate affecting, now, both the death hazard and conditional event rate given survival. In this setup, parameters in the conditional recurrent event rate model are no longer estimated consistently under the proposed approach. The estimators $\widehat{\theta}$ and $\widehat{R}_0(t)$ are roots of the estimating functions $\sum_{i=1}^{n}\int_0^{\tau} X_i dM_i^R(t)$ and $\sum_{i=1}^{n}\int_0^t dM_i^R(s)$, respectively, where $M_i^R(t)$ is as defined in

Table 2
Simulation results: robustness of proposed estimator under a misspecified model

$V(Q_i)$	β_1	$dR_0(t)$	$\psi(t)$	Bias	ESD	ASE	CP
2	1	0.25	4.18	0.01	0.39	0.38	0.95
		0.125	4.22	0.01	0.33	0.33	0.94
	0.5	0.25	4.30	0.00	0.37	0.38	0.94
		0.125	4.31	0.01	0.33	0.32	0.94
1	1	0.25	3.65	0.04	0.43	0.41	0.94
		0.125	3.71	0.03	0.35	0.34	0.95
	0.5	0.25	3.86	0.03	0.41	0.40	0.95
		0.125	3.88	0.05	0.34	0.34	0.96
0.5	1	0.25	2.91	0.01	0.44	0.43	0.94
		0.125	3.01	0.04	0.37	0.36	0.94
	0.5	0.25	3.25	0.01	0.43	0.42	0.94
		0.125	3.30	−0.01	0.35	0.35	0.96

Estimators are evaluated at $t = 10$, the mean of the censoring distribution.

equation (3). These estimating functions are unbiased and lead to consistent estimation when $E[dM_i^R(t) \mid X_i, D_i > t] = 0$. However, under models (4) and (5), $E[dM_i^R(t) \mid X_i, D_i > t] = P(C_i > t \mid X_i)E[Q_i \mid X_i, D_i > t]dt$, which need not equal 0, hence leading to inconsistent estimation.

The impact of such violations is evaluated through simulation. In addition to the scenarios examined in the simulation study of the correct model, an extreme case where $V(Q_i) = 2$, which is eight times $E(Q_i)$, was also examined to assess the impact of highly variable frailty terms on the performance of $\widehat{\psi}(t)$. Table 2 demonstrates the robustness of $\widehat{\psi}(t)$ under this misspecified model. The estimator is biased, although the bias is not large, particularly in relative terms. The proposed standard error estimators remain close to the ESDs and hence CPs are still close to 0.95. Generally, the proposed methods appear fairly robust to the no-unmeasured-predictors assumption.

5. Application

We applied our proposed method to the study of renal transplant patients. Patients began follow-up ($t = 0$) at the time of kidney transplantation. Some patients received an ECD transplant, while others got a non-ECD kidney. Patients who receive a kidney transplant are suspected of being subject to an increased mortality hazard and hospitalization rate in the weeks immediately following the transplant, due to the risk of surgical complications, which are suspected to be more serious in ECD recipients. From a public health perspective, it is of interest to compare ECD and non-ECD transplanted patients with respect to mean number of hospitalizations. Due to the strong time dependence in the effect of ECD kidney transplantation, the instantaneous effect on the hospitalization rate is generally of less interest than the cumulative effect. Moreover, survival probabilities are known to be reduced for ECD relative to non-ECD recipients (Port et al., 2002). Given these facts, the ECD effect on the mean number of hospitalizations would not be expected to be constant over time.

We combine demographic, clinical, and mortality data from the Scientific Registry of Transplant Recipients (SRTR) and hospitalization history information from Centers for Medicare and Medicaid Services. The study population is restricted to patients whose primary payer was Medicare. To increase homogeneity, we also exclude from our target population repeat and multiorgan transplant recipients. All Medicare patients aged ⩾18 who received a kidney transplant from a deceased donor in year 2000 were included in our study sample. In total, 3816 recipients with follow-up and complete covariate information were included and tracked from the time of transplant until death, loss to follow-up, or at the end of the observation period (December 31, 2005). Among the 3816 patients, 970 (25.42%) were observed to die during the 6-year follow-up period, with the remaining 2846 (74.58%) recipients censored, either due to loss to follow-up or administratively at the date December 31, 2005. Among the 3816 patients in our analysis, 3213 have no hospitalization in the follow-up period. Among the 603 hospitalized patients, 255 had one to three hospitalizations, and 279 had four to ten hospitalizations. There were three patients with more than 30 hospitalizations, with the maximum number of hospitalizations being 34. On average, each recipient experienced 0.85 hospitalizations during the follow-up period.

As stated, the treatment of interest is deceased donor source: ECD ($Z_{i1} = 1$) or non-ECD ($Z_{i1} = 0$). We adjusted for the same set of covariates in the proportional hazards and additive rates models: candidate demographics (gender,

Table 3
Analysis of kidney transplant data: estimated regression parameters from proportional hazards and additive rates models

Covariate, $Z_{ik} = X_{ik}$	$e^{\widehat{\beta}_k}$	p	$\widehat{\theta}_k$	$\widehat{SE}(\widehat{\theta}_k)$	p
ECD	1.28	0.002	84.4	29.4	0.004
Female	1.00	0.951	−6.6	19.2	0.732
Age 18–24	0.34	0.003	−80.7	41.1	0.045
Age 25–34	0.55	0.0001	−87.6	29.0	0.003
Age 35–44	0.83	0.092	−56.2	26.4	0.033
Age 45–54	1	–	0	–	–
Age 55–64	1.62	<0.0001	50.3	28.9	0.082
Age 65–70	2.12	<0.0001	109.5	40.9	0.007
Age ⩾ 70	2.88	<0.0001	146.5	47.8	0.002
African American	0.99	0.908	5.7	24.3	0.813
COPD	1.25	0.457	7.1	136.5	0.960
Angina	1.40	0.0002	71.1	39.5	0.072
Pretransplant malignancy	2.67	0.092	750.7	517.9	0.147
Cerebral vascular disease	1.19	0.323	95.4	71.5	0.182
Peripheral vascular disease	1.03	0.848	31.7	56.1	0.571
Polycystic kidneys	0.56	0.0006	−138.7	22.7	<0.001
Diabetes	1.59	<0.0001	123.3	32.2	<0.001
Hypertension	1.19	0.067	16.5	24.3	0.500
Years on dialysis	1.04	0.002	−0.3	3.0	0.932
Functional status: minor disability	0.31	<0.0001	−114.2	106.3	0.282
Intensive care unit	1.24	0.631	−220.1	136.2	0.106

θ: additional hospitalizations per 1000 per year survived.

Table 4
Analysis of kidney transplant data (ECD versus non-ECD): difference in mean numbers of hospitalizations

t	$\widehat{\psi}(t)$	$\widehat{SE}\{\widehat{\psi}(t)\}$	p
1 year	79	28	0.005
2 year	153	54	0.005
3 year	217	79	0.006
4 year	270	104	0.010
5 year	318	128	0.014

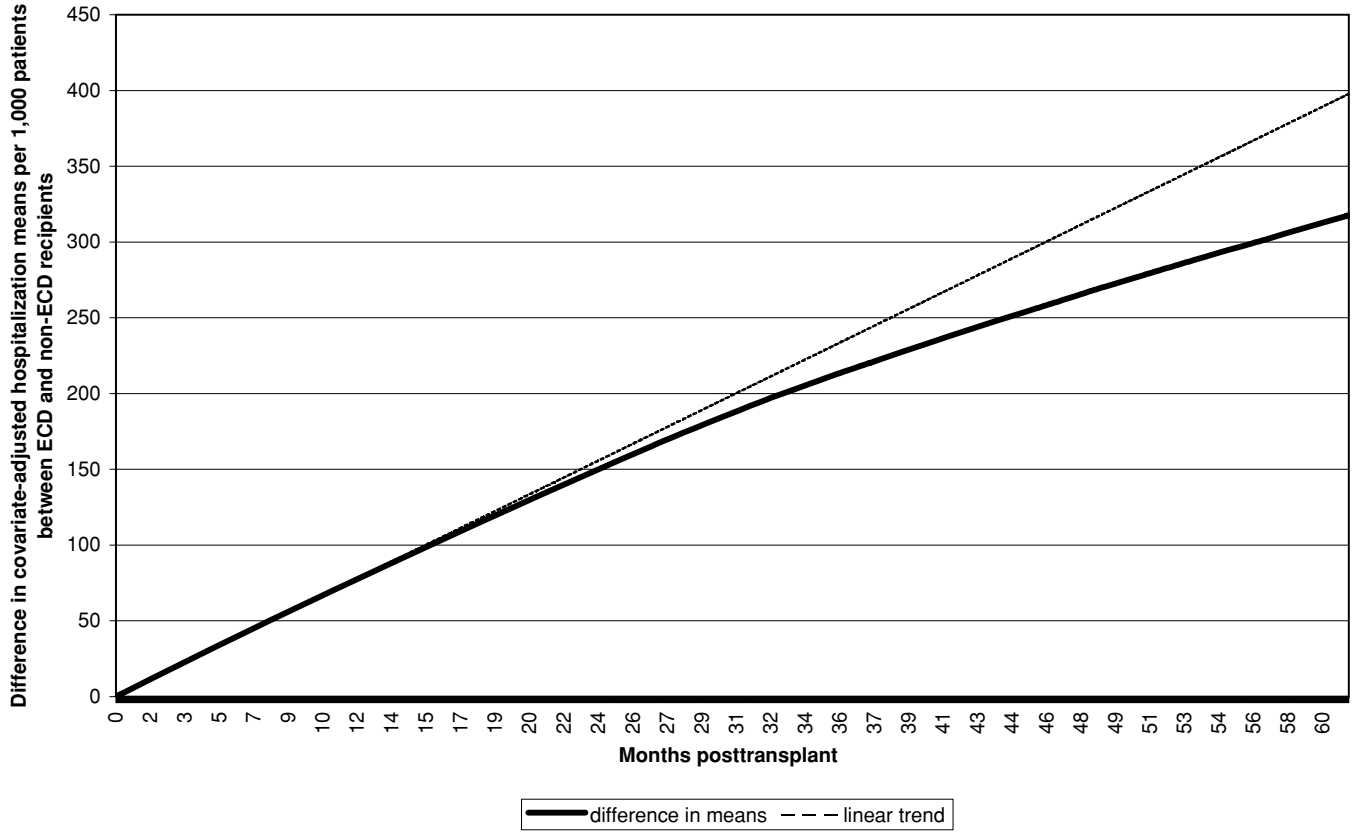
ψ: differences in hospitalization means per 1000 patients between ECD and non-ECD recipients at the end of 1 to 5 years.

age, race), pretransplant years on dialysis, various comorbid conditions (drug-treated chronic obstructive pulmonary disease [COPD], angina, symptomatic cerebral vascular disease, symptomatic peripheral vascular disease, pretransplant malignancy), primary diagnosis (polycystic kidney disease, diabetes, hypertension), functional status (fully active to severely disabled), and stay in the intensive care unit.

Hazard ratios and conditional recurrent event rate differences for each covariate are listed in Table 3, as well as their corresponding p-values. Recipients of an ECD kidney have a 1.3 times higher hazard of death compared to non-ECD recipients. At the same time, ECD kidney recipients experienced 84.4 more hospital admissions per 1000 patients per year survived. Lower survival probabilities for ECD patients could lead to a reduced mean number of hospitalizations, because hospitalizations are terminated by death. However, ECD patients also experience a significantly elevated conditional hospitalization rate given survival. The combination of these two effects results in the marginal effect of ECD on mean hospitalizations.

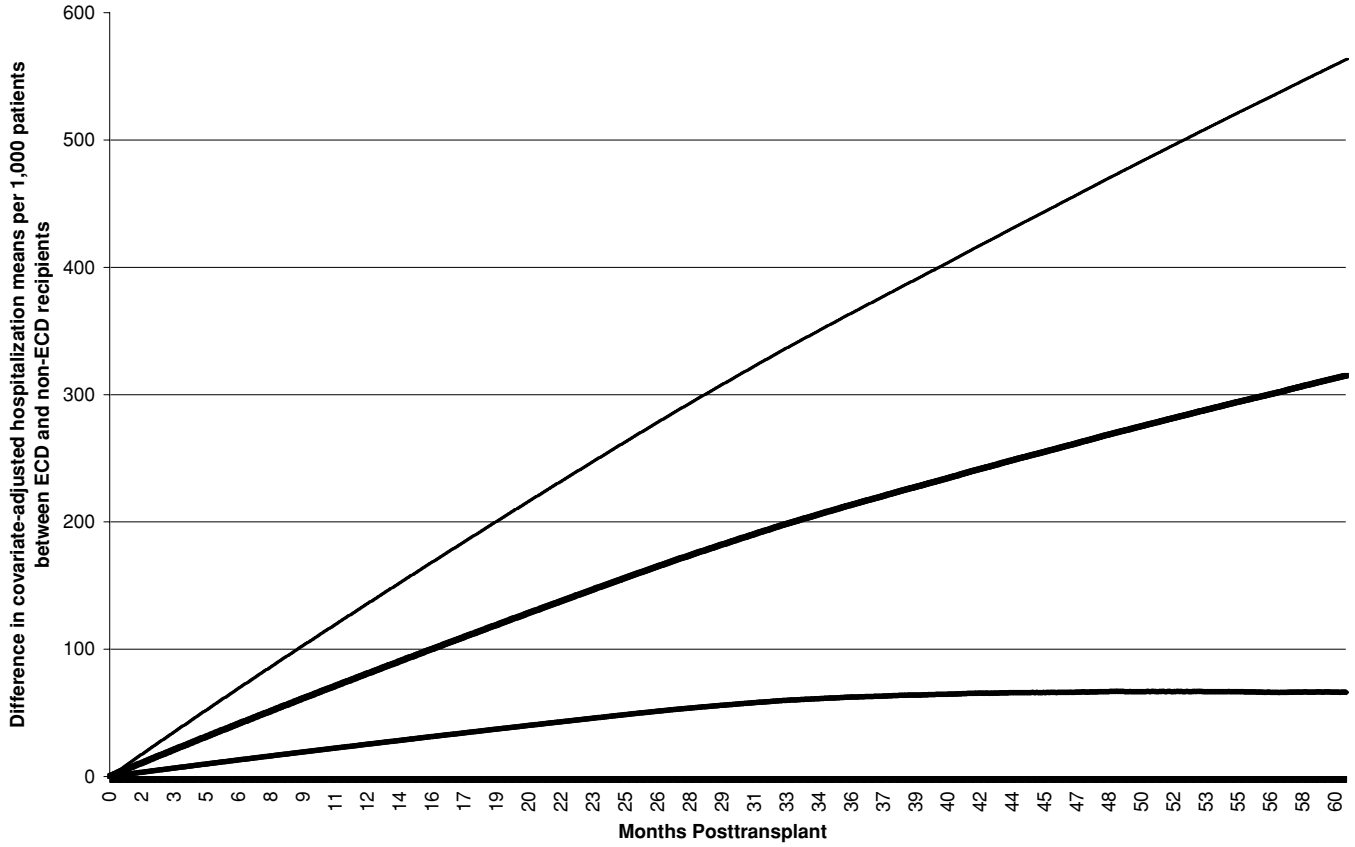
Based the proposed estimator $\widehat{\psi}(t)$, ECD recipients on average experience 79 more hospital admissions per 1000 persons compared to non-ECD recipients at the end of one year. This difference increases with time posttransplant and reaches 318 per 1000 patients at $t = 5$ years, as shown in Table 4. At each time point, the difference between ECD transplant patients and non-ECD transplant patients is highly significant. In summary, receiving an ECD kidney leads to a significantly greater number of hospitalizations posttransplant.

The difference in mean number of hospitalizations (averaging over the observed adjustment covariate pattern) for ECD and non-ECD transplant recipients is plotted in Figure 1. We can see that ECD recipients have more hospitalizations immediately after transplant. The difference keeps increasing with time and reaches 318 hospital admissions per 1000 patients at the end of the 5-year follow-up period. The increase in



Linear trend: The dashed line represents the difference in means that would result if the difference across the (0, 400] day interval persisted across the (400, 1825] day interval.

Figure 1. ECD effect on mean number of hospitalizations.



The two outside lines represent 95% pointwise confidence intervals for the year 0–5 interval.

Figure 2. 95% pointwise confidence intervals for the difference between ECD and non-ECD mean number of hospitalizations.

hospitalization rate due to ECD dominates the decrease in survival probability for ECD recipients and leads to positive ECD effects on the marginal event mean.

We present $\widehat{\psi}(t)$ and pointwise 95% confidence intervals in Figure 2. It is clear that the increase in mean number of hospitalizations associated with ECD transplantation is highly significant.

6. Discussion

We propose semiparametric methods to compare marginal treatment-specific mean numbers of recurrent events in the presence of a terminating event. The proposed methods involve modeling the terminating event hazard and the conditional recurrent event rate given survival separately, then integrating to estimate treatment-specific marginal recurrent event means. A measure of the combined effects is proposed, with its asymptotic properties derived and evaluated in moderate size samples. We demonstrate that the estimator works reasonably well under misspecified models; that is, models failing to incorporate unmeasured predictors of both the terminating event hazard function and conditional recurrent event rate. The proposed estimator is applied to national kidney transplant data to study the effect of ECD transplantation on posttransplant hospitalization admission numbers. ECD recipients are found to have significantly more hospitalizations during the whole follow-up period, with the difference between ECD and non-ECD recipients increasing with time.

The proposed method estimates the difference in treatment-specific marginal means, while incorporating treatment-specific differences in survival. Thus, $\psi(t)$ combines the actual survival probabilities and recurrent event rates for the treatment and placebo, respectively, and reflects the difference in marginal recurrent event numbers which would be observed between treatment and placebo patients with the same adjustment covariates. At the same time, researchers may also be interested in contrasts between treatment-specific survival and/or recurrent event rates among survivors. For a complete interpretation of their findings, researchers should carefully consider all three estimators, $\widehat{\psi}(t), \widehat{\beta}_1$, and $\widehat{\theta}_1$.

Our strongest assumption is that there are no unmeasured predictors of both the terminating event hazard and conditional event rate and this assumption will fail frequently, particularly in observational studies. Several issues are important in this regard. First, the unmeasured predictor must be a risk factor for both the terminating and recurrent event conditional on all measured covariates. If the most important predictors are included in the adjustment covariate vector, bias may be minimized and the estimated treatment effect may be a fairly accurate approximation to the reality. Second, the most popular alternative to a no-unmeasured-predictor assumption is to incorporate

a frailty. However, most frailty approaches for recurrent event data assume that the events follow a Poisson process, which is restrictive in its own right. To the best of our knowledge, there is only one method (Ye et al., 2007) that employs a frailty in the absence of the Poisson-process assumption, and this method does not propose cumulative effect measures. Third, our sensitivity analysis reveals that if an unmeasured frailty affects both the terminating event hazard and conditional rate, bias is relatively small.

7. Supplementary Materials

Web Appendices A and B, referenced in Section 3, are available under the Paper Information link at the *Biometrics* website `http://www.biometrics.tibs.org`.

Acknowledgements

This research was supported in part by National Institutes of Health grant R01 DK-70869 (DES). We thank the editor, the associate editor, and the referees for helpful comments, which led to substantial improvement of the manuscript. We also thank Jack Kalbfleisch for his thoughtful suggestions. We thank the SRTR for access to the renal transplant data, as well as the programmers and analysts at the Arbor Research Collaborative for Health and the University of Michigan Kidney Epidemiology and Cost Center for assistance with preparing the analysis files. The SRTR is funded by contract number 231-00-0116 from the Health Resources and Services Administration, U.S. Department of Health and Human Services.

References

Andersen, P. K. and Gill, R. D. (1982). Cox's regression model for counting processes: A large sample study. *Annals of Statistics* **10,** 1110–1120.

Cook, R. J. and Lawless, J. F. (1997). Marginal analysis of recurrent events and a terminating event. *Statistics in Medicine* **16,** 911–924.

Ghosh, D. and Lin, D. Y. (2000). Nonparametric analysis of recurrent events and death. *Biometrics* **56,** 554–562.

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

Ghosh, D. and Lin, D. Y. (2003). Semiparametric analysis of recurrent events data in the presence of dependent censoring. *Biometrics* **59,** 877–885.

Huang, C. Y. and Wang, M. C. (2004). Joint modeling and estimation for recurrent event processes and failure time data. *Journal of the American Statistical Association* **99,** 1153–1165.

Lawless, J. F. and Nadeau, C. (1995). Some simple robust methods for the analysis of recurrent event. *Technometrics* **37,** 158–168.

Li, Q. and Lagakos, S. (1997). Use of the Wei-Lin-Weissfeld method for the analysis of a recurring and a terminating event. *Statistics in Medicine* **16,** 925–940.

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

Lin, D. Y., Wei, L. J., Yang, I., and Ying, Z. (2000). Semiparametric regression for the mean and rate functions of recurrent events. *Journal of the Royal Statistical Society, Series B* **62,** 711–730.

Liu, L., Wolfe, R. A., and Huang, X. (2004). Shared frailty models for recurrent events and a terminal event. *Biometrics* **60,** 747–756.

Pepe, M. S. and Cai, J. (1993). Some graphical displays and marginal regression analyses for recurrent failure times and time-dependent covariates. *Journal of the American Statistical Association* **88,** 811–820.

Pollard, D. (1990). *Empirical Processes: Theory and Applications.* Hayward, California: Institute of Mathematical Statistics.

Port, F. K., Bragg-Gresham, J. L., Metzger, R. A., Dykstra, D. M., Gillespie, B. W., Young, E. W., Delmonico, F. L., Wynn, J. J., Merion, R. M.,Wolfe, R. A., and Held, P. J. (2002). Donor characteristics associated with reduced graft survival: An approach to expanding the pool of donor kidneys. *Transplantation* **74,** 1281–1286.

Prentice, R. L., Williams, B., and Peterson, A. V. (1981). On the regression analysis of multivariate failure time data. *Biometrika* **68,** 373–379.

Schaubel, D. E. and Cai, J. (2005). Analysis of clustered recurrent event data with application to hospitalization rates among renal failure patients. *Biostatistics* **6,** 404–419.

Schaubel, D. E., Zeng, D., and Cai, J. (2006). A semiparametric additive rates model for recurrent event data. *Lifetime Data Analysis* **12,** 389–406.

Sen, P. K. and Singer, J. M. (1993). *Large Sample Methods in Statistics.* New York: Chapman and Hall.

van der Vaart, A. W. and Wellner, J. A. (1996). *Weak Convergence and Empirical Processes.* New York: Springer.

Ye, Y., Kalbfleisch, J. D., and Schaubel, D. E. (2007). Semiparametric analysis of correlated recurrent and terminal events. *Biometrics* **63,** 78–87.

Received August 2007. *Revised June* 2008.
Accepted June 2008.