

Semiparametric Analysis of Recurrent Events Data in the Presence of Dependent Censoring

Debashis Ghosh

Department of Biostatistics, University of Michigan, 1420 Washington Heights, Ann Arbor, Michigan, U.S.A.
email: ghoshd@umich.edu

and

D. Y. Lin

Department of Biostatistics, University of North Carolina, CB#7420, Chapel Hill, North Carolina, U.S.A.
email: lin@bios.unc.edu

SUMMARY. Dependent censoring occurs in longitudinal studies of recurrent events when the censoring time depends on the potentially unobserved recurrent event times. To perform regression analysis in this setting, we propose a semiparametric joint model that formulates the marginal distributions of the recurrent event process and dependent censoring time through scale-change models, while leaving the distributional form and dependence structure unspecified. We derive consistent and asymptotically normal estimators for the regression parameters. We also develop graphical and numerical methods for assessing the adequacy of the proposed model. The finite-sample behavior of the new inference procedures is evaluated through simulation studies. An application to recurrent hospitalization data taken from a study of intravenous drug users is provided.

KEY WORDS: Accelerated failure time model; Competing risks; Counting process; Empirical process; Multiple events; Survival analysis.

1. Introduction

Recurrent events data arise in many longitudinal follow-up studies. Medical examples include tumor recurrences in cancer patients (Byar, 1980), repeated opportunistic infections in HIV-infected subjects (Li and Lagakos, 1997) and recurrent seizures in epileptic patients (Albert, 1991). There exist statistical methods for analyzing the marginal distributions of times to recurrent events (e.g., Wei, Lin, and Weissfeld, 1989), the gap times between successive recurrences (e.g., Prentice, Williams, and Peterson, 1981) and the intensity/rate functions of the recurrent event process (e.g., Andersen and Gill, 1982; Pepe and Cai, 1993; Lawless, Nadeau, and Cook, 1997).

In general, the recurrent event times are subject to right censoring, in that the observation of the recurrent events is terminated at or before the end of the study. If censoring is due to study termination or random loss to follow-up, then the censoring time can be reasonably regarded as independent of the recurrent event times. Dependent or informative censoring arises if the censoring time depends on the observed or unobserved recurrent event times. This would be the case if, for example, the subjects who are at higher risks of recurrent events tend to be withdrawn from the study earlier. In a typical medical study, the recurrent event times are subject to both independent and dependent censoring, and the correlation structure between the dependent censoring time and recurrent event process is complex. It is important to

realize that the aforementioned methods for analyzing recurrent events data are not valid in the presence of dependent censoring.

It is difficult to deal with dependent censoring in the analysis of recurrent events data, because the distribution of the recurrent event process is not identifiable nonparametrically. Limited methodological advances have been made. Chang (2000) proposed a method for comparing two gap time distributions, by postulating a multivariate scale-change model for the joint distribution of the dependent censoring time and the gap times of recurrent events. Wang, Qin, and Chiang (2001) studied the proportional rate model for recurrent events by characterizing the correlation structure between the recurrent event process and dependent censoring time through a random effect.

In this article, we propose a novel semiparametric regression model for formulating the joint distribution of the recurrent event process and dependent censoring. This model is described in the next section. In Section 3, we construct appropriate estimating functions for the regression parameters and derive the asymptotic properties of the resulting estimators. In Section 4, we show how to assess the adequacy of the proposed model. In Section 5, we present the results of our simulation studies. In Section 6, we apply the proposed methods to a longitudinal study on drug abuse. We conclude with some discussion in Section 7.

2. Data and Model

Let $N^*(t)$ denote the number of recurrent events that occur over the time interval $[0, t]$ in the absence of any censoring, let C and D denote, respectively, the times to independent and dependent censoring, and let \mathbf{Z} denote a set of covariates. We assume that C is independent of $N^*(\cdot)$ and D given \mathbf{Z} . By contrast, D is allowed to depend on $N^*(\cdot)$, even conditionally on \mathbf{Z} .

Let $\{N_i^*(\cdot), D_i, C_i, \mathbf{Z}_i\} (i = 1, \dots, n)$ be n independent replicates of $\{N^*(\cdot), D, C, \text{ and } \mathbf{Z}\}$. We suppose that there exist unknown constant vectors $\boldsymbol{\eta}_0$ and $\boldsymbol{\theta}_0$ such that, for given \mathbf{Z}_i and t , the bivariate random vectors $\{D_i e^{-\boldsymbol{\eta}'_0 \mathbf{Z}_i}, N_i^*(t e^{\boldsymbol{\theta}'_0 \mathbf{Z}_i})\}' (i = 1, \dots, n)$ have a common, but completely unspecified, joint distribution. In other words,

$$\begin{bmatrix} D_i e^{-\boldsymbol{\eta}'_0 \mathbf{Z}_i} \\ N_i^*(t e^{\boldsymbol{\theta}'_0 \mathbf{Z}_i}) \end{bmatrix} \stackrel{d}{=} \begin{bmatrix} D_0 \\ N_0^*(t) \end{bmatrix}, \quad i = 1, \dots, n, \tag{1}$$

where $\{D_0, N_0^*(t)\}'$ has an arbitrary bivariate distribution, and $\stackrel{d}{=}$ means equal in distribution. The independent censoring time C is allowed to depend on covariates \mathbf{Z} in an arbitrary manner. Model (1) is semiparametric, since the joint distribution of $\{D_0, N_0^*(\cdot)\}$ is completely unspecified. Under this joint model, the marginal distribution for the dependent censoring time satisfies the familiar accelerated failure time (AFT) model for survival data (Kalbfleisch and Prentice, 1980, pp. 32–34; Cox and Oakes, 1984, pp. 64–65), while that of the recurrent event times satisfy the AFT model for counting processes (Lin, Wei, and Ying, 1998). The latter marginal model can be expressed as $E\{N_i^*(t e^{\boldsymbol{\theta}'_0 \mathbf{Z}_i})\} = \mu_0(t) (i = 1, \dots, n)$, where $\mu_0(\cdot)$ is an arbitrary baseline mean frequency function. If \mathbf{Z} consists of a single treatment indicator, then $\boldsymbol{\eta}_0 > 0$ (or $\boldsymbol{\theta}_0 > 0$) implies that the treatment stochastically lengthens the time to dependent censoring (or the times to recurrent events) by a scale of $e^{\boldsymbol{\eta}_0}$ (or $e^{\boldsymbol{\theta}_0}$).

There are two major forms of dependent censoring. The first case is when subjects are voluntarily withdrawn from the study for reasons that are related to the recurrent event process. The subjects can potentially experience further events after the censoring time, but these events are not observable to the investigators. The second scenario is when dependent censoring occurs because of death; there are no further recurrent events after death. In this case, $N^*(\cdot)$ pertains to the latent recurrent event process, which is analogous to latent failure times studied in the classical competing-risks literature (Kalbfleisch and Prentice, 1980, Section 7.2).

The data consist of $\{N_i(\cdot), X_i, \delta_i, \mathbf{Z}_i\} (i = 1, \dots, n)$, where $N_i(t) = N_i^*(t \wedge D_i \wedge C_i)$, $X_i = D_i \wedge C_i$ and $\delta_i = I(D_i \leq C_i)$. Our task is to use these data to make inference about $\boldsymbol{\beta}_0 = (\boldsymbol{\eta}_0, \boldsymbol{\theta}_0)'$.

3. Inference Procedures

3.1 Estimation of $\boldsymbol{\beta}_0$

Since D is only subject to independent censoring by C , the existing methods for the AFT model can be used to estimate

$\boldsymbol{\eta}_0$. Writing $\tilde{X}_i(\boldsymbol{\eta}) = X_i e^{-\boldsymbol{\eta}' \mathbf{Z}_i}$, we shall estimate $\boldsymbol{\eta}_0$ from the log-rank estimating function

$$\mathbf{U}_1(\boldsymbol{\eta}) = \sum_{i=1}^n \delta_i \left[\mathbf{Z}_i - \frac{\sum_{j=1}^n I\{\tilde{X}_j(\boldsymbol{\eta}) \geq \tilde{X}_i(\boldsymbol{\eta})\} \mathbf{Z}_j}{\sum_{j=1}^n I\{\tilde{X}_j(\boldsymbol{\eta}) \geq \tilde{X}_i(\boldsymbol{\eta})\}} \right]. \tag{2}$$

Let $\hat{\boldsymbol{\eta}}$ be a zero-crossing of $\mathbf{U}_1(\boldsymbol{\eta})$. Then, $\hat{\boldsymbol{\eta}}$ is consistent and asymptotically normal (Tsiatis, 1990).

To motivate our estimator for $\boldsymbol{\theta}_0$, we first examine the estimating function of Lin et al. (1998). For $i = 1, \dots, n$ and $k = 1, 2, \dots$, let T_{ik} be the time to the k th recurrent event for the i th subject. Clearly, $N_i^*(t) = \sum_{k=1}^{\infty} I(T_{ik} \leq t)$ and $N_i(t) = \sum_{k=1}^{\infty} I(T_{ik} \leq t \wedge X_i)$. Define $\tilde{T}_{ik}(\boldsymbol{\theta}) = T_{ik} e^{-\boldsymbol{\theta}' \mathbf{Z}_i}$ and $\tilde{X}_i(\boldsymbol{\theta}) = X_i e^{-\boldsymbol{\theta}' \mathbf{Z}_i} (i = 1, \dots, n; k = 1, 2, \dots)$. Then, the estimating function of Lin et al. (1998) takes the form

$$\tilde{\mathbf{U}}(\boldsymbol{\theta}) = \sum_{i=1}^n \int_0^{\infty} \left[\mathbf{Z}_i - \frac{\sum_{j=1}^n I\{\tilde{X}_j(\boldsymbol{\theta}) \geq t\} \mathbf{Z}_j}{\sum_{j=1}^n I\{\tilde{X}_j(\boldsymbol{\theta}) \geq t\}} \right] d\tilde{N}_i(t; \boldsymbol{\theta}), \tag{3}$$

where $\tilde{N}_i(t; \boldsymbol{\theta}) = \sum_{k=1}^{\infty} I\{\tilde{T}_{ik}(\boldsymbol{\theta}) \leq t \wedge \tilde{X}_i(\boldsymbol{\theta})\}$.

It is straightforward to show that

$$\tilde{\mathbf{U}}(\boldsymbol{\theta}) = \sum_{i=1}^n \int_0^{\infty} \left[\mathbf{Z}_i - \frac{\sum_{j=1}^n I\{\tilde{X}_j(\boldsymbol{\theta}) \geq t\} \mathbf{Z}_j}{\sum_{j=1}^n I\{\tilde{X}_j(\boldsymbol{\theta}) \geq t\}} \right] d\tilde{M}_i(t; \boldsymbol{\theta}),$$

where

$$\tilde{M}_i(t; \boldsymbol{\theta}) = \int_0^t I\{\tilde{X}_i(\boldsymbol{\theta}) \geq s\} d\{N_i^*(s e^{\boldsymbol{\theta}' \mathbf{Z}_i}) - \mu_0(s)\},$$

and $\mu_0(\cdot)$ is as defined in Section 2. If there were no dependent censoring such that X were independent of $N^*(\cdot)$ given \mathbf{Z} , then

$$\begin{aligned} E\{\tilde{M}_i(t; \boldsymbol{\theta}_0)\} &= \int_0^t I\{\tilde{X}_i(\boldsymbol{\theta}) \geq s\} \\ &\quad \times d[E\{N_i^*(s e^{\boldsymbol{\theta}'_0 \mathbf{Z}_i}) - \mu_0(s)\}] = 0. \end{aligned}$$

Consequently, $\tilde{\mathbf{U}}(\boldsymbol{\theta}_0)$ would be a sum of integrals with respect to zero-mean processes, so that $\tilde{\mathbf{U}}(\boldsymbol{\theta})$ would be an asymptotically unbiased estimating function. In the presence of dependent censoring,

$$\begin{aligned} E\{\tilde{M}_i(t; \boldsymbol{\theta}_0)\} &= E \int_0^t I\{\tilde{X}_i(\boldsymbol{\theta}_0) \geq s\} \\ &\quad \times [E\{dN_i^*(s e^{\boldsymbol{\theta}'_0 \mathbf{Z}_i}) \mid D_i e^{-\boldsymbol{\theta}'_0 \mathbf{Z}_i} \geq s\} - d\mu_0(s)]. \end{aligned}$$

In general, there does not exist a function $\mu_0(\cdot)$ such that $E\{dN_i^*(t e^{\boldsymbol{\theta}'_0 \mathbf{Z}_i}) \mid D_i e^{-\boldsymbol{\theta}'_0 \mathbf{Z}_i} \geq t\} = d\mu_0(t)$ for all i , the only exception being when model (1) holds with $\boldsymbol{\eta}_0 = \boldsymbol{\theta}_0$. Thus, $\tilde{\mathbf{U}}(\boldsymbol{\theta})$ is not a valid estimating function unless $\boldsymbol{\eta}_0 = \boldsymbol{\theta}_0$.

The above derivation suggests that (3) would be a valid estimating function under dependent censoring if we subjected the transformed recurrent event times $\tilde{T}_{ik}(\boldsymbol{\theta})(k=1,2,\dots)$ to right censoring by $X_i e^{-\boldsymbol{\eta}'\mathbf{Z}_i}$ instead of $X_i e^{-\boldsymbol{\theta}'\mathbf{Z}_i}$. However, a difficulty would arise if $\boldsymbol{\eta}'\mathbf{Z}_i < \boldsymbol{\theta}'\mathbf{Z}_i$, because then X_i would be multiplied by a larger factor than the T_{ik} ($k=1,2,\dots$), so that some of the censored recurrent event times might need to be uncensored. To circumvent this difficulty, we subject the $\tilde{T}_{ik}(\boldsymbol{\theta})(k=1,2,\dots)$ to right censoring by $\tilde{X}_i(\mathbf{b}) \equiv X_i e^{-\boldsymbol{\eta}'\mathbf{Z}_i - d}$, where $d = \max_{1 \leq i \leq n} (\boldsymbol{\theta} - \boldsymbol{\eta})'\mathbf{Z}_i$. If \mathbf{Z} consists of a single treatment indicator, then $d = 0$ if $\eta \geq \theta$ and $d = \theta - \eta$ otherwise. Thus, the (uncensored) recurrent event times in the treatment group may be artificially censored if $\eta > \theta$, and the (uncensored) recurrent event times in the control group may be artificially censored if $\eta < \theta$. This kind of artificial censoring has been employed by Robins and Rotnitzky (1992), Lee, Li and Elashoff (1993), Lin, Robins, and Wei (1996), and Chang (2000), among others.

For fixed $\boldsymbol{\eta}$, we propose estimating $\boldsymbol{\theta}_0$ using the estimating function

$$\mathbf{U}_2(\boldsymbol{\theta}; \boldsymbol{\eta}) = \sum_{i=1}^n \int_0^\infty \{\mathbf{Z}_i - \bar{\mathbf{Z}}^{(2)}(t; \boldsymbol{\beta})\} dN_{2i}(t; \boldsymbol{\beta}),$$

where $\bar{\mathbf{Z}}^{(2)}(t; \boldsymbol{\beta}) = \sum_{j=1}^n I\{\tilde{X}_j(\boldsymbol{\beta}) \geq t\} \mathbf{Z}_j / \sum_{j=1}^n I\{\tilde{X}_j(\boldsymbol{\beta}) \geq t\}$, and $N_{2i}(t; \boldsymbol{\beta}) = \sum_{k=1}^\infty I\{\tilde{T}_{ik}(\boldsymbol{\theta}) \leq t \wedge \tilde{X}_i(\boldsymbol{\beta})\}$. Given $\hat{\boldsymbol{\eta}}$, we define $\hat{\boldsymbol{\theta}}$ as a zero-crossing of $\mathbf{U}_2(\boldsymbol{\theta}; \hat{\boldsymbol{\eta}})$. Write $\hat{\boldsymbol{\beta}} = (\hat{\boldsymbol{\theta}}, \hat{\boldsymbol{\eta}})'$.

Let $d_0 = \max_{1 \leq i \leq n} (\boldsymbol{\theta}_0 - \boldsymbol{\eta}_0)'\mathbf{Z}_i$. Under model (1), the expectations $E\{dN_i^*(se^{\boldsymbol{\theta}'\mathbf{Z}_i}) | D_i e^{-\boldsymbol{\eta}_0'\mathbf{Z}_i - d_0} \geq t\} (i=1, \dots, n)$ have a common value, say $r_0(t) dt$, regardless of the value of \mathbf{Z}_i . Simple algebraic manipulation yields

$$\mathbf{U}_2(\boldsymbol{\theta}; \boldsymbol{\eta}) = \sum_{i=1}^n \int_0^\infty \{\mathbf{Z}_i - \bar{\mathbf{Z}}^{(2)}(t; \boldsymbol{\beta})\} dM_{2i}(t; \boldsymbol{\beta}), \quad (4)$$

where

$$M_{2i}(t; \boldsymbol{\beta}) = \int_0^t I\{\tilde{X}_i(\boldsymbol{\beta}) \geq s\} \{dN_i^*(se^{\boldsymbol{\theta}'\mathbf{Z}_i}) - r_0(s) ds\}.$$

Clearly,

$$\begin{aligned} E\{M_{2i}(t; \boldsymbol{\beta}_0)\} &= E \int_0^t I\{\tilde{X}_i(\boldsymbol{\beta}_0) \geq s\} \\ &\times [E\{dN_i^*(se^{\boldsymbol{\theta}'\mathbf{Z}_i}) | D_i e^{-\boldsymbol{\eta}_0'\mathbf{Z}_i - d_0} \geq s\} - r_0(s) ds] = 0. \end{aligned}$$

Thus, $\mathbf{U}_2(\boldsymbol{\theta}_0; \boldsymbol{\eta}_0)$ is the sum of stochastic integrals with respect to zero-mean processes. Consequently, the use of $\mathbf{U}_2(\boldsymbol{\theta}; \hat{\boldsymbol{\eta}})$ would yield a consistent and asymptotically normal estimator, as demonstrated in the next section.

3.2 Asymptotic Properties of $\hat{\boldsymbol{\beta}}$

Define $N_{1i}(t; \boldsymbol{\eta}) = \delta_i I\{\tilde{X}_i(\boldsymbol{\eta}) \leq t\}$ and

$$M_{1i}(t; \boldsymbol{\eta}) = N_{1i}(t; \boldsymbol{\eta}) - \int_0^t I\{\tilde{X}_i(\boldsymbol{\eta}) \geq s\} \lambda_0(s) ds,$$

where $\lambda_0(t)$ is the hazard function of D_0 . Then

$$\begin{aligned} \mathbf{U}_1(\boldsymbol{\eta}) &= \sum_{i=1}^n \int_0^\infty \{\mathbf{Z}_i - \bar{\mathbf{Z}}^{(1)}(t; \boldsymbol{\eta})\} dN_{1i}(t; \boldsymbol{\eta}) \\ &= \sum_{i=1}^n \int_0^\infty \{\mathbf{Z}_i - \bar{\mathbf{Z}}^{(1)}(t; \boldsymbol{\eta})\} dM_{1i}(t; \boldsymbol{\eta}), \end{aligned}$$

where $\bar{\mathbf{Z}}^{(1)}(t; \boldsymbol{\eta}) = \sum_{j=1}^n I\{\tilde{X}_j(\boldsymbol{\eta}) \geq t\} \mathbf{Z}_j / \sum_{j=1}^n I\{\tilde{X}_j(\boldsymbol{\eta}) \geq t\}$. Because $M_{1i}(t) \equiv M_{1i}(t; \boldsymbol{\eta}_0) (i=1, \dots, n)$ are orthogonal martingales with respect to the marginal filtration $\mathcal{F}_t = \sigma\{N_{1i}(s; \boldsymbol{\eta}_0), I\{\tilde{X}_i(\boldsymbol{\eta}_0) \geq s\}, \mathbf{Z}_i; 0 \leq s \leq t, i=1, \dots, n\}$, the martingale central limit theorem (Fleming and Harrington, 1991, Theorem 5.3.5) implies that $n^{-1/2} \mathbf{U}_1(\boldsymbol{\eta}_0)$ is asymptotically zero-mean normal and

$$n^{-1/2} \mathbf{U}_1(\boldsymbol{\eta}_0) = n^{-1/2} \sum_{i=1}^n \int_0^\infty \{\mathbf{Z}_i - \bar{\mathbf{Z}}^{(1)}(t)\} dM_{1i}(t) + o_P(1), \quad (5)$$

where $\bar{\mathbf{Z}}^{(1)}(t)$ is the probability limit of $\bar{\mathbf{Z}}^{(1)}(t; \boldsymbol{\eta}_0)$.

It is more delicate to derive the asymptotic distribution of $n^{-1/2} \mathbf{U}_2(\boldsymbol{\theta}_0; \boldsymbol{\eta}_0)$ because $M_{2i}(t) \equiv M_{2i}(t; \boldsymbol{\beta}_0) (i=1, \dots, n)$ are not martingales. Nevertheless, by applying the empirical process arguments as found in the Appendix of Lin et al. (1998) to (4), we can establish that

$$n^{-1/2} \mathbf{U}_2(\boldsymbol{\theta}_0; \boldsymbol{\eta}_0) = n^{-1/2} \sum_{i=1}^n \int_0^\infty \{\mathbf{Z}_i - \bar{\mathbf{Z}}^{(2)}(t)\} dM_{2i}(t) + o_P(1), \quad (6)$$

where $\bar{\mathbf{Z}}^{(2)}(t)$ is the probability limit of $\bar{\mathbf{Z}}^{(2)}(t; \boldsymbol{\beta}_0)$. The interested reader is referred to Ghosh (2000, Chapter 6) for a proof of this result, as well as the proofs of other results in this section.

Write $\mathbf{U}(\boldsymbol{\beta}) = \{\mathbf{U}'_1(\boldsymbol{\eta}), \mathbf{U}'_2(\boldsymbol{\theta}; \boldsymbol{\eta})\}'$. Because both the right-hand sides of (1) and (2) consist of sums of n independent random vectors, the multivariate central limit theorem implies that $n^{-1/2} \mathbf{U}(\boldsymbol{\beta}_0)$ is asymptotically zero-mean normal with covariance matrix

$$\mathbf{V} \equiv \lim_{n \rightarrow \infty} n^{-1} \sum_{i=1}^n E \begin{bmatrix} \boldsymbol{\psi}_{1i} \boldsymbol{\psi}'_{1i} & \boldsymbol{\psi}_{1i} \boldsymbol{\psi}'_{2i} \\ \boldsymbol{\psi}_{2i} \boldsymbol{\psi}'_{1i} & \boldsymbol{\psi}_{2i} \boldsymbol{\psi}'_{2i} \end{bmatrix},$$

where $\boldsymbol{\psi}_{1i} = \int_0^\infty \{\mathbf{Z}_i - \bar{\mathbf{Z}}^{(1)}(t)\} dM_{1i}(t)$ and $\boldsymbol{\psi}_{2i} = \int_0^\infty \{\mathbf{Z}_i - \bar{\mathbf{Z}}^{(2)}(t)\} dM_{2i}(t)$. By extending the asymptotic linearity arguments of Ying (1993), we can show that, for $\boldsymbol{\beta}$ in a small neighborhood of $\boldsymbol{\beta}_0$,

$$n^{-1/2} \mathbf{U}(\boldsymbol{\beta}) = n^{-1/2} \mathbf{U}(\boldsymbol{\beta}_0) + \mathbf{A} n^{1/2} (\boldsymbol{\beta} - \boldsymbol{\beta}_0) + o_P(1), \quad (7)$$

where \mathbf{A} is the asymptotic slope matrix of $n^{-1} \mathbf{U}(\boldsymbol{\beta}_0)$. It follows that $n^{1/2} (\hat{\boldsymbol{\beta}} - \boldsymbol{\beta}_0)$ is asymptotically zero-mean normal with covariance matrix $\mathbf{A}^{-1} \mathbf{V} \mathbf{A}^{-1}$.

A direct evaluation of \mathbf{A} would require estimation of $\lambda_0(\cdot)$ and $r_0(\cdot)$, which cannot be done reliably in practical applications. Thus, we appeal to the resampling approach of Parzen, Wei, and Ying (1994). By replacing $\Lambda_0(t) \equiv \int_0^t \lambda_0(s) ds$ and

$R_0(t) \equiv \int_0^t r_0(s) ds$ in $M_{1i}(t; \boldsymbol{\eta})$ and $M_{2i}(t; \boldsymbol{\beta})$ with the Aalen-Breslow type estimators

$$\widehat{\Lambda}_0(t; \boldsymbol{\eta}) \equiv \sum_{i=1}^n \int_0^t \frac{dN_{1i}(s; \boldsymbol{\eta})}{\sum_{j=1}^n I\{\widetilde{X}_j(\boldsymbol{\eta}) \geq s\}}$$

and

$$\widehat{R}_0(t; \boldsymbol{\beta}) \equiv \sum_{i=1}^n \int_0^t \frac{dN_{2i}(s; \boldsymbol{\beta})}{\sum_{j=1}^n I\{\widetilde{X}_j(\boldsymbol{\beta}) \geq s\}},$$

we obtain

$$\widehat{M}_{1i}(t; \boldsymbol{\eta}) = N_{1i}(t; \boldsymbol{\eta}) - \int_0^t I\{\widetilde{X}_i(\boldsymbol{\eta}) \geq s\} d\widehat{\Lambda}_0(s; \boldsymbol{\eta}),$$

$$\widehat{M}_{2i}(t; \boldsymbol{\beta}) = N_{2i}(t; \boldsymbol{\beta}) - \int_0^t I\{\widetilde{X}_i(\boldsymbol{\beta}) \geq s\} d\widehat{R}_0(s; \boldsymbol{\beta}).$$

Let $\widehat{\boldsymbol{\psi}}_{1i} = \int_0^\infty \{\mathbf{Z}_i - \overline{\mathbf{Z}}^{(1)}(t; \widehat{\boldsymbol{\eta}})\} d\widehat{M}_{1i}(t; \widehat{\boldsymbol{\eta}})$ and $\widehat{\boldsymbol{\psi}}_{2i} = \int_0^\infty \{\mathbf{Z}_i - \overline{\mathbf{Z}}^{(2)}(t; \widehat{\boldsymbol{\beta}})\} d\widehat{M}_{2i}(t; \widehat{\boldsymbol{\beta}})$. We construct the pair of equations

$$\mathbf{U}_1(\boldsymbol{\eta}) = \sum_{i=1}^n \widehat{\boldsymbol{\psi}}_{1i} G_i, \tag{8}$$

$$\mathbf{U}_2(\boldsymbol{\theta}; \boldsymbol{\eta}) = \sum_{i=1}^n \widehat{\boldsymbol{\psi}}_{2i} G_i, \tag{9}$$

where (G_1, \dots, G_n) are independent standard normal random variables that are independent of the data $\{N_i(\cdot), X_i, \delta_i, \mathbf{Z}_i\}$ ($i = 1, \dots, n$). Denote the solution to equations (8) and (9) by $\boldsymbol{\beta}^* = (\boldsymbol{\eta}^*, \boldsymbol{\theta}^*)'$.

Obviously, $\mathbf{U}(\boldsymbol{\beta}^*) = \sum_{i=1}^n \widehat{\boldsymbol{\psi}}_i G_i$, where $\widehat{\boldsymbol{\psi}}_i = (\widehat{\boldsymbol{\psi}}_{1i}, \widehat{\boldsymbol{\psi}}_{2i})'$. It then follows from (7) that

$$n^{1/2}(\boldsymbol{\beta}^* - \widehat{\boldsymbol{\beta}}) = \mathbf{A}^{-1} n^{-1/2} \sum_{i=1}^n \widehat{\boldsymbol{\psi}}_i G_i + o_P(1). \tag{10}$$

The conditional distribution of $n^{-1/2} \sum_{i=1}^n \widehat{\boldsymbol{\psi}}_i G_i$ given the data $\{N_i(\cdot), X_i, \delta_i, \mathbf{Z}_i\}$ ($i = 1, \dots, n$) is zero-mean normal with covariance matrix $n^{-1} \sum_{i=1}^n \widehat{\boldsymbol{\psi}}_i \widehat{\boldsymbol{\psi}}_i'$, which converges in probability to \mathbf{V} . Therefore, the conditional distribution of $n^{1/2}(\boldsymbol{\beta}^* - \widehat{\boldsymbol{\beta}})$ is asymptotically equivalent to the unconditional distribution of $n^{1/2}(\widehat{\boldsymbol{\beta}} - \boldsymbol{\beta}_0)$. To approximate the distribution of $\widehat{\boldsymbol{\beta}}$, we obtain a large number of realizations of $\boldsymbol{\beta}^*$ by repeatedly generating the random samples (G_1, \dots, G_n) while fixing the data $\{N_i(\cdot), X_i, \delta_i, \mathbf{Z}_i\}$ ($i = 1, \dots, n$) at their observed values and solving equations (8) and (9). The empirical distribution of $\boldsymbol{\beta}^*$ can then be used to perform hypothesis testing or interval estimation on individual components of $\boldsymbol{\eta}_0$ and $\boldsymbol{\theta}_0$.

4. Model Checking

Model (1) implies that (i) $D_i e^{-\boldsymbol{\eta}_0' \mathbf{Z}_i}$ ($i = 1, \dots, n$) have a common marginal distribution and (ii) $E\{dN_i^*(te^{\boldsymbol{\theta}_0' \mathbf{Z}_i}) | D_i e^{-\boldsymbol{\eta}_0' \mathbf{Z}_i} \geq t\}$ ($i = 1, \dots, n$) have a common value. The inference procedures developed in Section 3 rely on model (1) only

through these two conditions. To verify these two conditions, we consider the “residual” processes

$$\mathbf{U}_1(t; \boldsymbol{\eta}) = \sum_{i=1}^n \mathbf{Z}_i \widehat{M}_{1i}(t; \boldsymbol{\eta}), \tag{11}$$

$$\mathbf{U}_2(t; \boldsymbol{\beta}) = \sum_{i=1}^n \mathbf{Z}_i \widehat{M}_{2i}(t; \boldsymbol{\beta}). \tag{12}$$

It is easy to see that $\mathbf{U}_1(\infty; \boldsymbol{\eta}) = \mathbf{U}_1(\boldsymbol{\eta})$ and $\mathbf{U}_2(\infty; \boldsymbol{\beta}) = \mathbf{U}_2(\boldsymbol{\theta}; \boldsymbol{\eta})$. By extension of the arguments of Wei (1984), the Kolmogorov-type test $\sup_t \|n^{-1/2} \mathbf{U}_1(t; \widehat{\boldsymbol{\eta}})\|$ is consistent against the general alternative that condition (i) fails; given condition (i), the $\sup_t \|n^{-1/2} \mathbf{U}_2(t; \widehat{\boldsymbol{\beta}})\|$ test is consistent against the general alternative that condition (ii) fails.

To carry out the Kolmogorov-type tests, we need to ascertain the distributions of the residual processes $\mathbf{U}_1(\cdot; \widehat{\boldsymbol{\eta}})$ and $\mathbf{U}_2(\cdot; \widehat{\boldsymbol{\beta}})$ under the null hypothesis that model (1) holds. The null distribution of $n^{-1/2} \mathbf{U}(t, t^\dagger) \equiv \{n^{-1/2} \mathbf{U}_1^*(t; \widehat{\boldsymbol{\eta}}), n^{-1/2} \mathbf{U}_2^*(t^\dagger; \widehat{\boldsymbol{\beta}})\}'$ is asymptotically equivalent to the conditional distribution of $n^{-1/2} \mathbf{U}^*(t, t^\dagger) \equiv \{n^{-1/2} \mathbf{U}_1^*(t), n^{-1/2} \mathbf{U}_2^*(t^\dagger)\}'$ given the data $\{N_i(\cdot), X_i, \delta_i, \mathbf{Z}_i\}$ ($i = 1, \dots, n$), where

$$\begin{aligned} \mathbf{U}_1^*(t) &= \sum_{i=1}^n \int_0^t \{\mathbf{Z}_i - \overline{\mathbf{Z}}^{(1)}(s; \widehat{\boldsymbol{\eta}})\} d\widehat{M}_{1i}(s; \widehat{\boldsymbol{\eta}}) G_i \\ &\quad + \mathbf{U}_1(t; \boldsymbol{\eta}^*) - \mathbf{U}_1(t; \widehat{\boldsymbol{\eta}}), \end{aligned}$$

and

$$\begin{aligned} \mathbf{U}_2^*(t) &= \sum_{i=1}^n \int_0^t \{\mathbf{Z}_i - \overline{\mathbf{Z}}^{(2)}(s; \widehat{\boldsymbol{\beta}})\} d\widehat{M}_{2i}(s; \widehat{\boldsymbol{\beta}}) G_i \\ &\quad + \mathbf{U}_2(t; \boldsymbol{\beta}^*) - \mathbf{U}_2(t; \widehat{\boldsymbol{\beta}}), \end{aligned}$$

because it follows from the arguments used for showing (5)–(7) and (10) that

$$\begin{aligned} n^{-1/2} \mathbf{U}(t, t^\dagger) &= n^{-1/2} \sum_{i=1}^n \left[\int_0^t \{\mathbf{Z}_i - \overline{\mathbf{Z}}^{(1)}(s)\} dM_{1i}(s) \right. \\ &\quad \left. + \int_0^{t^\dagger} \{\mathbf{Z}_i - \overline{\mathbf{Z}}^{(2)}(s)\} dM_{2i}(s) \right] \\ &\quad + \mathbf{A}(t, t^\dagger) n^{1/2} (\widehat{\boldsymbol{\beta}} - \boldsymbol{\beta}_0) + o_P(1), \end{aligned}$$

and

$$\begin{aligned} n^{-1/2} \mathbf{U}^*(t, t^\dagger) &= n^{-1/2} \sum_{i=1}^n \left[\int_0^t \{\mathbf{Z}_i - \overline{\mathbf{Z}}^{(1)}(s; \widehat{\boldsymbol{\eta}})\} d\widehat{M}_{1i}(s; \widehat{\boldsymbol{\eta}}) \right. \\ &\quad \left. + \int_0^{t^\dagger} \{\mathbf{Z}_i - \overline{\mathbf{Z}}^{(2)}(s; \widehat{\boldsymbol{\beta}})\} d\widehat{M}_{2i}(s; \widehat{\boldsymbol{\beta}}) \right] \\ &\quad \times G_i + \mathbf{A}(t, t^\dagger) n^{1/2} (\boldsymbol{\beta}^* - \widehat{\boldsymbol{\beta}}) + o_P(1), \end{aligned}$$

where $\mathbf{A}(t, t^\dagger)$ is the asymptotic slope matrix of $n^{-1} \mathbf{U}(t, t^\dagger)$; see Ghosh (2000, Chapter 6) for a detailed proof.

In view of the above result, we approximate the null distribution of $\mathbf{U}(t, t^\dagger)$ by generating a large number of realizations of $\mathbf{U}^*(t, t^\dagger)$ using simulation. The empirical distribution

of $\mathbf{U}^*(t, t^\dagger)$ can then be used to perform the Kolmogorov-type tests. Furthermore, graphical model checking can be carried out by comparing the observed patterns of $\mathbf{U}_1(\cdot; \hat{\boldsymbol{\eta}})$ and $\mathbf{U}_2(\cdot; \hat{\boldsymbol{\beta}})$, with a few realizations from $\mathbf{U}_1^*(\cdot)$ and $\mathbf{U}_2^*(\cdot)$.

5. Simulation Studies

We conducted extensive simulation studies to assess the finite-sample properties of the proposed methods. Because $\boldsymbol{\eta}_0$ is a nuisance parameter, we focused our studies on the estimation of $\boldsymbol{\theta}_0$. We compared the proposed estimator with that of Lin et al. (1998). The latter is referred to as the naive estimator because it does not adjust for dependent censoring.

We generated $\{D_0, N_0^*(\cdot)\}$ in two stages: a gamma random variable v , with mean 1 and variance σ^2 , was first generated; given v , the dependent censoring time D_0 and the gap time between every two successive recurrent events were generated as independent exponential random variables with hazard rates v and $4v$, respectively. We multiply D_0 and the gap times so generated by $e^{\eta_0 Z}$ and $e^{\theta_0 Z}$ to produce the dependent censoring time and gap times associated with Z , where Z is Bernoulli (0.5), uniform (0, 0.5) or uniform (0, 2). Because of the common random effect, D is correlated with $N^*(\cdot)$, unless $\sigma^2 = 0$. Independent censoring time was generated from the uniform (0, τ) distribution.

Table 1 summarizes the results for the following combinations of simulation parameters: $\boldsymbol{\beta}_0 = [0.25, \log(3)]'$, $\sigma^2 = 0, 1$ or 4 , $\tau = 5$ or 20 , and $n = 100$. The choice of $\tau = 5$ or 20 yielded an average of about 2.4 or 4.7 observed recurrent events per subject, respectively. For each configuration of simulation parameters, we generated 1000 data sets. For each data set, the resampling distribution was based on 1000 realizations.

The results in Table 1 as well as additional results reported in Ghosh (2000) demonstrate that the proposed methods have satisfactory small-sample behavior. The proposed parameter estimator is virtually unbiased. The standard error estimator reflects well the true variation of the parameter estimator. The confidence intervals have proper coverage probabilities. When $\sigma^2 = 0$, the naive estimator performs well and is more efficient than the proposed estimator. For nonzero values of σ^2 , however, the naive estimator is biased, and the coverage probabilities of the corresponding confidence intervals are poor.

6. The AIDS Link to Intravenous Experiences Study

We now apply the proposed methods to the AIDS Links to Intravenous Experiences (ALIVE) cohort study (Vlahov et al., 1991), which collected data on HIV infection, in-patient admissions and other variables from a group of intravenous drug users in the city of Baltimore, Maryland, U.S.A. Following Wang et al. (2001), we consider hospitalization data collected between August 1, 1993, and December 31, 1997, from 715 subjects. In this study, the dependent censoring is caused by death. We focus on the comparison between the subjects who are HIV-positive (coded as $Z = 0$) at baseline (i.e., August 1, 1993) versus those who are HIV-negative (coded as $Z = 1$).

Table 2 summarizes the hospitalization and survival experiences for the subjects in the two groups. The Kaplan-Meier estimates of the 4-year survival probabilities are approximately 0.95 and 0.72 for the HIV-negative and HIV-positive subjects, respectively. The log-rank statistic for comparing the two survival distributions is 71.7, yielding to a p -value less than 0.0001. Figure 1 displays the estimated cause-specific

Table 1
Summary of simulation results

Z	σ^2	τ	Naive method					Proposed method					
			Bias	SE	SEE	CP	CP*	Bias	SE	SEE	CP	CP*	AC
Bernoulli (0.5)	0	5	-0.01	0.21	0.20	0.94	0.95	0.01	0.24	0.22	0.94	0.95	0.27
		20	-0.01	0.15	0.14	0.95	0.95	0.00	0.18	0.17	0.93	0.94	0.26
	1	5	-0.28	0.16	0.14	0.75	0.77	-0.01	0.20	0.19	0.94	0.95	0.27
		20	-0.32	0.13	0.11	0.71	0.72	-0.01	0.16	0.15	0.94	0.94	0.28
	4	5	-0.39	0.14	0.13	0.60	0.62	0.00	0.17	0.16	0.94	0.94	0.27
		20	-0.43	0.11	0.10	0.56	0.57	0.01	0.13	0.12	0.94	0.95	0.26
Uniform (0, 0.5)	0	5	0.01	0.14	0.12	0.95	0.95	0.00	0.17	0.17	0.95	0.95	0.12
		20	-0.01	0.11	0.09	0.94	0.94	0.00	0.15	0.14	0.95	0.94	0.13
	1	5	-0.10	0.10	0.08	0.85	0.83	0.00	0.12	0.11	0.94	0.94	0.13
		20	-0.12	0.09	0.08	0.82	0.80	-0.01	0.09	0.08	0.94	0.94	0.12
	4	5	-0.17	0.11	0.10	0.79	0.78	0.01	0.12	0.11	0.95	0.95	0.12
		20	-0.19	0.07	0.08	0.77	0.75	-0.01	0.08	0.07	0.94	0.94	0.13
Uniform (0, 2)	0	5	-0.01	0.21	0.23	0.95	0.94	0.01	0.27	0.26	0.95	0.95	0.29
		20	0.01	0.18	0.17	0.95	0.95	0.00	0.19	0.18	0.94	0.94	0.30
	1	5	-0.15	0.20	0.19	0.85	0.83	-0.01	0.22	0.20	0.94	0.95	0.30
		20	-0.18	0.16	0.16	0.80	0.78	0.01	0.15	0.14	0.93	0.94	0.29
	4	5	-0.22	0.19	0.19	0.77	0.75	-0.01	0.20	0.20	0.95	0.94	0.31
		20	-0.26	0.14	0.13	0.73	0.71	0.01	0.17	0.16	0.95	0.95	0.30

Note: Bias is the sampling mean of the estimator of θ_0 minus θ_0 ; SE is the sampling standard error of the estimator of θ_0 ; SEE is the sampling mean of the standard error estimator; CP is the coverage probability of the Wald 95% confidence interval, and CP* is the coverage probability of the 95% confidence interval based on the empirical percentiles; AC is the percentage of artificially censored observations.

Table 2
Hospitalization and survival experiences for the HIV-negative and HIV-positive subjects in the ALIVE study

HIV status	Number of subjects	Number of hospitalizations							Number of deaths
		0	1	2	3	4	5	≥ 6	
Negative	430	209	95	47	20	25	7	21	22
Positive	285	109	61	35	36	15	3	26	78

hospitalization rates (i.e., the rates of hospitalization given being alive) obtained from the Nelson-Aalen estimates by monotone splines (Ramsay, 1988). These results show that the HIV-positive subjects tend to have higher hospitalization rates than the HIV-negative subjects.

Table 3 provides the results of fitting model (1) to the ALIVE data. As expected, the baseline HIV status is highly predictive of survival. On average, the subjects who are HIV-negative at baseline live 4.76 times as long as their HIV-positive counterparts. The proposed method yields a much larger estimate for the effect of the baseline HIV status on hospitalizations than the naive method. This difference is attributed primarily to the strong positive correlation between hospitalizations and death. Approximately 23.3% of the hospitalizations in the HIV-negative subjects are artificially censored. The proposed estimate indicates that, on average, the times to hospitalizations for an HIV-negative drug user are 1.78 times those of an HIV-positive drug user. Using a random-effect proportional rates model, Wang et al. (2001) found that the hospitalization occurrence rate of the

HIV-positive drug users is about 1.64 times that of the HIV-negative drug users.

Figure 2 displays the observed curves of $U_1(\cdot; \hat{\eta})$ and $U_2(\cdot; \hat{\beta})$ along with 20 realizations from the simulated null distributions. The observed patterns are not extreme compared with the simulated realizations. Based on 1000 realizations, the P -values for the Kolmogorov-type tests $\sup_t |U_1(t; \hat{\eta})|$ and $\sup_t |U_2(t; \hat{\beta})|$ are estimated at 0.63 and 0.48, respectively. Thus, there is no evidence of model misspecification.

7. Discussion

As evident from Sections 5 and 6, failure to adjust for dependent censoring can result in serious bias in the analysis of recurrent events data, especially when the dependence of censoring time on recurrent event times is strong. The proposed method enables one to make valid inference under mild assumptions. There is a trade-off between the bias of the naive estimator and the loss of efficiency caused by the artificial censoring of the proposed method. The efficiency loss is a more important concern in smaller samples, whereas the bias

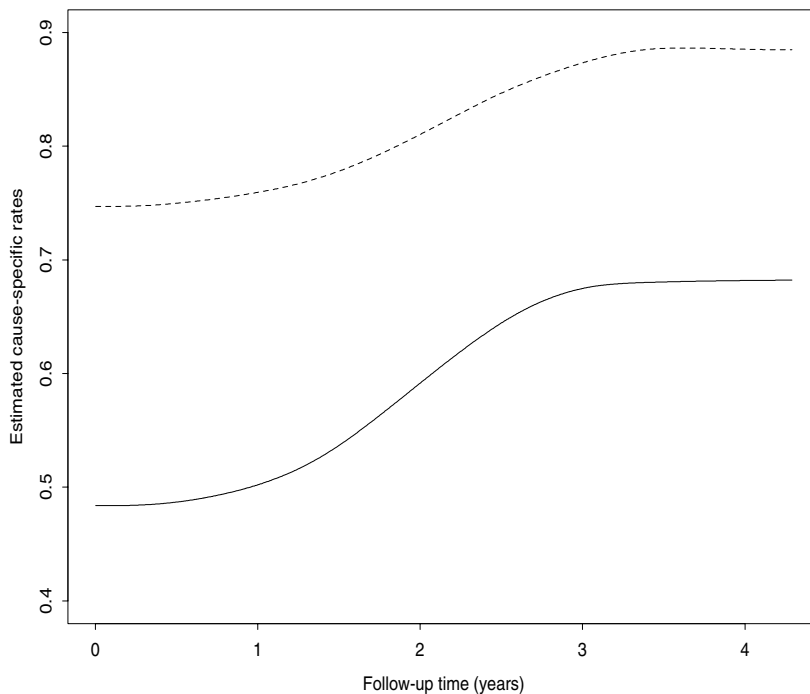


Figure 1. Cause-specific hospitalization rates separated by the baseline HIV status: the solid and dotted curves correspond to the HIV-negative and HIV-positive subjects, respectively.

Table 3
Regression estimates for the effects of the baseline HIV status on death and hospitalizations in the ALIVE study

	Estimate	Estimated standard error	95% confidence intervals	
			Percential	Wald
Death	1.56	0.28	(1.31, 1.91)	(1.01, 2.11)
Hospitalizations				
Naive	0.27	0.07	(0.15, 0.43)	(0.14, 0.40)
Proposed	0.58	0.17	(0.24, 0.79)	(0.24, 0.92)

Note: The estimated standard error and confidence intervals are based on 1000 realizations.

becomes a more prominent issue in larger samples. In this sense, the proposed method is most useful in large samples. The proposed estimator would be inefficient and thus of limited value if the range of $(\boldsymbol{\theta}_0 - \boldsymbol{\eta}_0)' \mathbf{Z}$ is very large, because then there would be excessive artificial censoring. Unfortunately, that is also the situation where the naive estimator would yield severe bias. In that kind of situation, the method of Wang et al. (2001) may be used instead, especially if the sample size is small. An alternative approach would be to develop methods for sensitivity analysis along the lines of Scharfstein, Rotnitzky, and Robins (1999).

The equation $N_i^*(te^{\boldsymbol{\theta}_0' \mathbf{Z}_i}) \stackrel{d}{=} N_0^*(t)$ in model (1) essentially specifies accelerated failure time models for the recurrent event times: $\log T_{ik} = -\boldsymbol{\theta}'_0 \mathbf{Z}_i + \epsilon_{ik}$ ($i = 1, \dots, n$; $k = 1, 2, \dots$), where the error terms $(\epsilon_{i1}, \epsilon_{i2}, \dots)$ ($i = 1, \dots, n$)

are independent random vectors with a common, but completely unspecified, joint distribution. By contrast, Chang (2000) postulated accelerated failure time models for the gap times: $\log(T_{ik} - T_{i,k-1}) = -\theta_k Z_i + \epsilon_{ik}^*$ ($i = 1, \dots, n$; $k = 1, \dots, K$), where $T_{i0} = 0$, Z_i is the treatment indicator, θ_k is the regression parameter for the k th gap time, and ϵ_{ik}^* are the corresponding error terms. Because the proposed model formulates the effects of covariates on the entire recurrent event process through a single (possibly vector-valued) regression parameter, whereas Chang's (2000) model pertains to a fixed number of gap times with separate regression parameters for each gap time, our approach is more parsimonious and more efficient than Chang's approach. Wang et al. (2001) also used the data on the entire recurrent event process, but they relied on the proportional means model (Lawless et al., 1997),

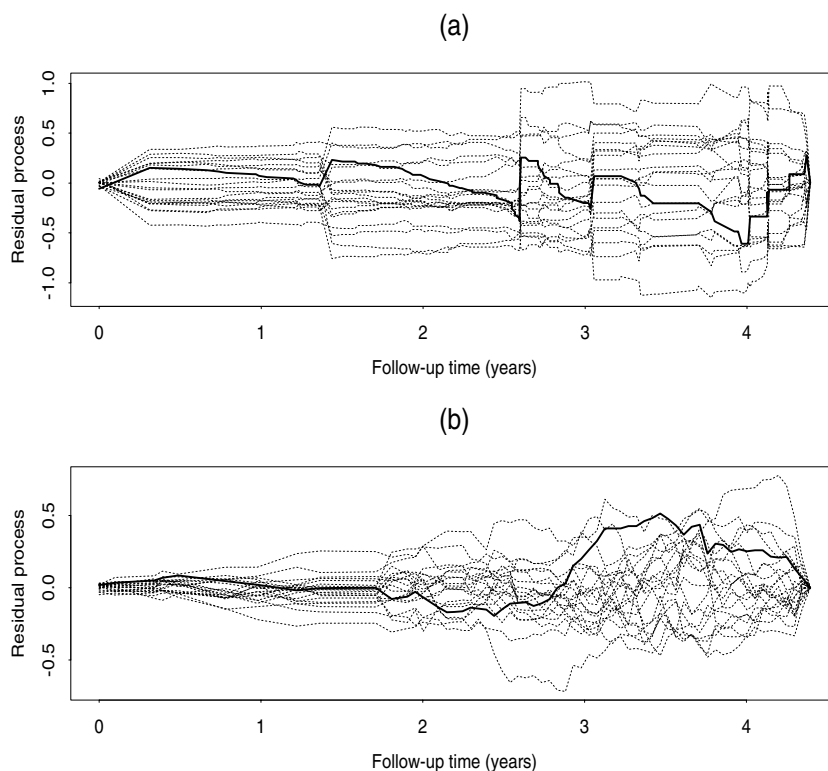


Figure 2. Plots of residual processes $U_1(\cdot; \hat{\boldsymbol{\eta}})$ and $U_2(\cdot; \hat{\boldsymbol{\beta}})$, shown in (a) and (b), respectively. The observed process is shown by the solid curve, and 20 simulated realizations by the dotted curves.

rather than the accelerated time model. Their joint model is more restrictive than ours, in that the dependence between D and $N^*(\cdot)$, as well as the dependence among the recurrent event times themselves, are characterized by a common random effect. The Wang et al.'s (2001) approach should be more efficient than ours and Chang's, because it is more restrictive and does not involve artificial censoring.

For simplicity, we have focused on the estimating functions based on unweighted log-rank statistics. We may incorporate weight functions into $\mathbf{U}_1(\boldsymbol{\eta})$ and $\mathbf{U}_2(\boldsymbol{\theta}; \boldsymbol{\eta})$, and all the asymptotic results will continue to hold. Optimal weight functions for $\mathbf{U}_1(\boldsymbol{\eta})$ have been well studied, and involve the unknown density function of the error term (Tsiatis, 1990). It is unclear what the optimal weight functions would be for $\mathbf{U}_2(\boldsymbol{\theta}; \boldsymbol{\eta})$, which is based on complex correlated data. Furthermore, it would be difficult to use the optimal weight functions in applications, because they will certainly involve the density functions of the underlying failure time distributions.

As mentioned in Section 2, if censoring is due to death, then $N^*(\cdot)$ corresponds to the latent or potential recurrent event process. An alternative approach is to consider the marginal mean function for the actual number of recurrent events, acknowledging the fact that death precludes further recurrences. This approach was taken by Ghosh and Lin (2000, 2002).

In many applications, the exact times of the recurrent events cannot be observed. What is known instead is the number of recurrences since the subject's last visit. In addition, the subject's visit schedule can be a random process. The resulting data are commonly referred to as panel count data (e.g., Lawless and Zhan, 1998; Sun and Wei, 2000). It would be worthwhile to generalize the methods proposed here to the panel count data.

ACKNOWLEDGEMENTS

This research was supported by the National Institutes of Health. The authors are grateful to the referees for their helpful comments, and to Drs David Vlahov and Steffanie Stratthdee for providing the data from the ALIVE study, which was supported by the NIDA grants DA 04334 and DA 08009.

RÉSUMÉ

La censure dépendante apparaît dans les études longitudinales d'événements récurrents quand le temps de censure dépend des temps d'événements récurrents potentiellement non observés. Pour réaliser une analyse de régression dans cette situation, nous proposons un modèle semi-paramétrique joint qui spécifie les distributions marginales du processus des événements récurrents et du temps de censure dépendant au moyen de modèles à changement d'échelle tout en laissant la forme distributionnelle et la structure de dépendance non spécifiées. Nous déduisons des estimateurs consistants et asymptotiquement normaux pour les paramètres de la régression. Nous développons également des méthodes graphiques et numériques pour évaluer la pertinence du modèle proposé. Le comportement de la nouvelle procédure pour des échantillons finis est évalué par des études de simulations. Une application de données d'hospitalisations récurrentes dans une étude chez les usagers de drogue par voie intraveineuse est fournie.

REFERENCES

- Albert, P. S. (1991). A two-state Markov mixture model for a time series of epileptic seizure counts. *Biometrics* **47**, 1371–1381.
- Andersen, P. K. and Gill, R. D. (1982). Cox's regression model for counting processes: A large sample study. *Annals of Statistics* **10**, 1100–1120.
- Byar, D. P. (1980). The Veterans Administration study of chemoprophylaxis for recurrent stage I bladder tumors: Comparisons of placebo, pyridoxine and topical thiotepa. In *Bladder Tumors and Other Topics in Urological Oncology*, M. Pavone-Macaluso, P. H. Smith, and F. Edsmyr (eds), 363–370. New York: Plenum.
- Chang, S.-H. (2000). A two-sample comparison for multiple ordered event data. *Biometrics* **56**, 183–189.
- Cox, D. R. and Oakes, D. (1984). *Analysis of Survival Data*. London: Chapman and Hall.
- Fleming, T. R. and Harrington, D. P. (1991). *Counting Processes and Survival Analysis*. New York: Wiley.
- Ghosh, D. (2000). Nonparametric and semiparametric analysis of recurrent events in the presence of terminal events and dependent censoring. Ph.D. thesis, University of Washington.
- 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 events. *Statistica Sinica* **12**, 663–688.
- Kalbfleisch, J. D. and Prentice, R. L. (1980). *The Statistical Analysis of Failure Time Data*. New York: Wiley.
- Lawless, J. F. and Zhan, M. (1998). Analysis of interval-grouped recurrent-event data using piecewise constant rate functions. *Canadian Journal of Statistics* **26**, 549–565.
- Lawless, J. F., Nadeau, C., and Cook, R. J. (1997). Analysis of mean and rate functions for recurrent events. In *Proceedings of the First Seattle Survival Analysis Symposium*, D. Y. Lin and T. R. Fleming (eds), 137–149. New York: Springer.
- Lee, J. J., Li, K.-C., and Elashoff, R. M. (1993). On censoring for censored paired data. *Journal of the American Statistical Association* **88**, 104–118.
- 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., Robins, J. M., and Wei, L. J. (1996). Comparing two failure time distributions in the presence of dependent censoring. *Biometrika* **83**, 381–393.
- Lin, D. Y., Wei, L. J., and Ying, Z. (1998). Accelerated failure time models for counting processes. *Biometrika* **85**, 605–618.
- Parzen, M. I., Wei, L. J., and Ying, Z. (1994). A resampling method based on pivotal estimating functions. *Biometrika* **81**, 341–350.
- 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.

- Prentice, R. L., Williams, B. J., and Peterson, A. V. (1981). On the regression analysis of multivariate failure time data. *Biometrika* **68**, 373–379.
- Ramsay, J. O. (1988). Monotone regression splines in action (with discussion). *Statistical Science* **3**, 425–461.
- Robins, J. and Rotnitzky, A. (1992). Recovery of information and adjustment for dependent censoring using surrogate markers. In *Aids Epidemiology—Methodological Issues*, N. Jewell, K. Dietz, and V. Farewell (eds), 297–331. Boston: Birkhäuser.
- Scharfstein, D. O., Rotnitzky, A., and Robins, J. M. (1999). Adjusting for nonignorable drop-out using semiparametric nonresponse models (with comments and rejoinder). *Journal of the American Statistical Association* **94**, 1096–1146.
- Sun, J. and Wei, L. J. (2000). Regression analysis of panel count data with covariate-dependent observation and censoring times. *Journal of the Royal Statistical Society, Series B* **62**, 293–302.
- Tsiatis, A. A. (1990). Estimating regression parameters using linear rank tests for censored data. *Annals of Statistics* **18**, 354–372.
- Vlahov, D., Anthony, J. C., Muñoz, A., Margolick, J., Nelson, K. E., Celentano, D. D., Solomon, L., and Polk, B. F. (1991). The ALIVE study: A longitudinal study of HIV-1 infection in intravenous drug users: Description of methods. *Journal of Drug Issues* **21**, 759–776.
- Wang, M. C., Qin, J., and Chiang, C. T. (2001). Analyzing recurrent event data with informative censoring. *Journal of the American Statistical Association* **96**, 1057–1065.
- Wei, L. J. (1984). Testing goodness of fit for proportional hazards model with censored observations. *Journal of the American Statistical Association* **79**, 649–652.
- Wei, L. J., Lin, D. Y., and Weissfeld, L. (1989). Regression analysis of multivariate incomplete failure time data by modeling marginal distributions. *Journal of the American Statistical Association* **84**, 1065–1073.
- Ying, Z. (1993). A large sample study of rank estimation for censored regression data. *Annals of Statistics* **21**, 76–99.

Received April 2002. Revised April 2003.

Accepted April 2003.