Acta Mathematicae Applicatae Sinica, English Series
Vol. 32, No. 4 (2016) 1035–1052
DOI: 10.1007/s10255-016-0624-3
http://www.ApplMath.com.cn & www.SpringerLink.com

Analyzing Longitudinal Data with Informative Observation and Terminal Event Times

Rui MIAO, Xin CHEN, Liu-quan SUN†

Institute of Applied Mathematics, Academy of Mathematics and Systems Science, Chinese Academy of Sciences, Beijing 100190, China (†E-mail: slq@amt.ac.cn)

Abstract Longitudinal data often arise when subjects are followed over a period of time, and in many situations, there may exist informative observation times and a dependent terminal event such as death that stops the follow-up. In this article, we propose joint modeling and analysis of longitudinal data with possibly informative observation times and a dependent terminal event in which a common subject-specific latent variable is used to characterize the correlations. A borrow-strength estimation procedure is developed for parameter estimation, and both large-sample and finite-sample properties of the proposed estimators are established. In addition, some goodness-of-fit methods for assessing the adequacy of the model are provided. An application to a bladder cancer study is illustrated.

Keywords borrow-strength method; frailty model; informative observation times; joint modeling; longitudinal data; terminal event

2000 MR Subject Classification 62N01, 62N02; 62G05

1 Introduction

Longitudinal data occur frequently in many medical follow-up studies and observational investigations, and longitudinal responses are often correlated with observation times. For example, subjects often selectively miss their visits or return at non-scheduled points in time. As a result, the measurement times are irregular, and may be correlated with the longitudinal outcomes. In recent years, studies of longitudinal data with informative observation times have attracted considerable attention[4,9,14,15]. For example, [9] presented a class of inverse intensity-of-visit process-weighted estimators in a typical marginal regression model. [15] considered a joint model for the longitudinal and observation processes via a shared latent variable. [4] proposed a joint model for analysis of the longitudinal outcomes through two latent variables.

In practice, however, there exists a terminal event such as death that stops the follow-up. Further, it is often the case that the terminal event is strongly correlated with the longitudinal and observation processes. For example, patients in a severe disease stage often die in a shorter period, and longitudinal medical costs may be less than that of patients in a mild disease stage[10]. [11] showed that ignoring the dependent terminal event could lead to biased estimates in the intensity model of hospital visits. Also, [12] showed that disregarding the dependent terminal event would lead to biased estimates in modeling the longitudinal medical costs. Hence there is clearly a need for an analytical method that can directly model longitudinal data, which accounts for both informative observation times and a dependent terminal event simultaneously.

However, there exists some limited research on the analysis of longitudinal data in the pres-

Manuscript received December 27, 2012. Revised March 20, 2013.
Supported by the National Natural Science Foundation of China Grants (No. 11231010 and 11171330) and Key Laboratory of RCSDS, CAS (No. 2008DP173182).
†Corresponding author.

ence of informative observation times and a dependent terminal event. Recently, [10] presented a joint random effects model for longitudinal data with informative observation times and a dependent terminal event, where the distributions of the random effects are specified. [2] proposed some shared frailty models to analyze panel count data with correlated observation and follow-up times, where one random effect is required to be normally distributed. In applications, it is difficult to determine an appropriate frailty distribution for a specific data set, and miss-specifying the frailty distribution may lead to bias in the aforementioned two approaches. In addition, [17] presented a joint model to analyze longitudinal data via two latent variables for subjects who are currently alive.

In this article, we propose a new class of joint models to analyze longitudinal data with informative observation times and a dependent terminal event for all subjects. The association among the longitudinal response, observation times and terminal event is modeled via a latent variable. To be specific, a class of semiparametric random effect models is used for the longitudinal response, and a subject specific nonhomogeneous Poisson process is used for the observation time process, and a proportional frailty model is used for the terminal event time. The proposed joint models are flexible in that no parametric assumptions on the distributions of the latent variable and censoring times are made, and informative censoring is allowed for the longitudinal response, observation times and terminal event. A borrow-strength estimation procedure is proposed by first estimating the value of the latent variable from recurrent event data, then using the estimated value in the proportional frailty model and the semiparametric random effect models.

The rest of the paper is organized as follows. Section 2 introduces notation and model specification. Section 3 presents the borrow-strength estimation procedure about regression parameters of interest with the focus on the effect of covariates on the longitudinal response. The asymptotic properties of the resulting estimates are established. In Section 4, goodness-of-fit methods for assessing the adequacy of the models are presented. In Section 5, simulation studies are conducted to assess the performance of the proposed methods. Section 6 applies the methods to the bladder cancer study, and concluding remarks are given in Section 7.

2 Notation and Joint Models

Consider a longitudinal study, and let $Y(t)$ denote the longitudinal response variable of interest. Also let X be the $p \times 1$ vector of covariates. In addition, let D be the time of the terminal event and C be the follow-up or censoring time. Write $T = C \wedge D$ and $\delta = I\ (D \leq C)$, where $a \wedge b = \min(a, b)$, and $I(\cdot)$ is the indicator function. Then the observed counting process for the terminal event is represented by $N^D(t) = I(T \leq t, \delta = 1)$. Similarly, let $N(t)$ be the counting process denoting the number of the observation times before or at time t. The longitudinal process $Y(t)$ is observed only at the time points where $N(t)$ jumps for $t \leq T$.

Let V be a nonnegative valued latent variable with $E(V|X) = 1$. We will assume that given X and V, $Y(t)$ follows the marginal model

$$E\{Y(t)|X, V\} = \mu_0(t) + \beta_0' X + V, \tag{1}$$

where $\mu_0(t)$ is an unspecified smooth function of t, and β_0 is a vector of unknown regression parameters[4,15]. Model (1) characterizes the marginal mean of the process $Y(t)$ while leaving its dependence structure and distributional form completely unspecified. More precisely, Model (1) can be written as

$$Y(t) = \mu_0(t) + \beta_0' X + V + \varepsilon(t),$$

which is a semiparametric random effect model, where $\varepsilon(t)$ is a mean-zero measurement error process[1,4]. The above Model (1) without V has been considered by many authors such as [1] and [8].

For the observation process, we assume that, conditioning on X and V, $N(t)$ is a nonhomogeneous Poisson process with intensity function

$$\lambda(t|X,V) = V\lambda_0(t)\exp(\gamma_0'X), \tag{2}$$

where γ_0 is a vector of unknown regression parameters, and $\lambda_0(t)$ is an unknown continuous baseline intensity function. Model (2) has been used by [3,18] for the analysis of recurrent-event data. Note that by Model (2), conditioning on X, the rate function of $N(t)$ is $\lambda_0(t)\exp(\gamma_0'X)$.

For the terminal event time, we specify the proportional frailty model for D as

$$h(t|X,V) = Vh_0(t)\exp(\alpha_0'X), \tag{3}$$

where α_0 is a vector of unknown regression parameters, and $h_0(t)$ is an unspecified baseline hazard function. Clearly, a large value of V inflates the marginal mean of the longitudinal process and the intensity of recurrent events as well as the hazard of the terminal event.

For the joint models, the role of the latent variable V is to correlate the longitudinal process $Y(t)$, the observation process $N(t)$ and the terminal event time D. In particular, they specify a positive correlation of the three processes. We can similarly develop models and a corresponding inference procedure for the case in which the three processes are negatively correlated. We comment on this further in Section 7. Since the distribution of the latent variable is left unspecified, the proposed models are not able to measure the degree of association between the longitudinal outcome and the observation time process, and the association between the longitudinal outcome and the terminal event. [15] studied Models (1) and (2) in the presence of informative censoring time, where the censoring time was treated as a nuisance, and the modeling of the censoring time was not considered. However, it does not seem to exist research on Models (1), (2) and (3) together in the presence of a dependent terminal event, in which the association among the three processes is modeled explicitly.

Note that we do not make any assumption on the distributions of $V$and C. Here our main interest is to assess covariate effects on the longitudinal process, i.e., to estimate parameter β_0. Note that Model (1) implies that

$$E\{Y(t)|X\} = \{\mu_0(t) + 1\} + \beta_0'X.$$

Hence, the parameters β_0 can serve as marginal effects of covariates on the longitudinal process of interest. In the following, we assume that the censoring time C may depend on X and V in an arbitrary way, but $Y(\cdot)$, $N(\cdot)$, D and C are mutually independent conditional on (X, V).

Remark 1. For notational convenience, Models (1), (2) and (3) assume the same set of covariates X. The proposed estimation procedure can be extended in a straightforward manner to deal with different sets of covariates for these three models.

3 Borrow-Strength Estimation Procedure

For a random sample of n subjects, the observed data consist of $\{Y_i(t)dN_i(t), N_i(t), T_i, \delta_i, X_i, 0 \le t \le T_i,\ i = 1, \cdots, n\}$. Also let m_i denote the total number of observations and $t_{i1}, \cdots, t_{i,m_i}$ the observation times on subject i. If V_i is known, we can use Lin and Ying's method[8] and the inverse probability weighting technique to estimate β_0. In reality, we are not able to observe the value of V_i, and thus cannot directly use the above approach. For this, consider Model (2), and conditioning on $\{m_i, T_i, V_i, X_i\}$, the event times $(t_{i1}, \cdots, t_{i,m_i})$ are the order statistics of a set of i.i.d. random variables with the density function $\lambda_0(t)I(0 \le t \le T_i)/\Lambda_0(T_i)$ (see [18]). Let $\Delta_i(t) = I(T_i \ge t)$ and $\Lambda_i(t) = \Lambda_0(t)\exp(\gamma_0'X_i)$, where $\Lambda_0(t) = \int_0^t \lambda_0(u)du$ is the cumulative baseline intensity function. Then

$$E\{\Delta_i(t)dN_i(t)|V_i,T_i,X_i,m_i\} = \Delta_i(t)m_i\frac{d\Lambda_i(t)}{\Lambda_i(T_i)}.$$

Also note that for any constant ρ,

$$E\{m_i\rho^{m_i}|V_i,T_i,X_i\} = \rho V_i\Lambda_i(T_i)\exp\{V_i\Lambda_i(T_i)(\rho-1)\},$$

and

$$E\{m_i(m_i-1)\rho^{m_i}|V_i,T_i,X_i\} = \rho^2V_i^2\Lambda_i(T_i)^2\exp\{V_i\Lambda_i(T_i)(\rho-1)\}.$$

Let $H_0(t)=\int_0^t h_0(u)du$ be the cumulative baseline hazard function. For notational convenience, define $\mathcal{A}_0(t)=\int_0^t \mu_0(u)d\Lambda_0(u)$, $S_i(t)=\exp\{-m_iH_0(t)\exp(\alpha_0'X_i)/\Lambda_i(T_i)\}$,

$$\Omega_i(t) = \frac{m_i-1}{\Lambda_i(T_i)}\exp\Big\{-\frac{H_0(t)\exp(\alpha_0'X_i)}{\Lambda_i(T_i)}\Big\},$$

and

$$M_i(t) = \int_0^t \Big[\big\{Y_i(u)-\beta_0'X_i-\Omega_i(u)\big\}\frac{\Delta_i(u)}{S_i(u)}dN_i(u) - \frac{m_i\Delta_i(u)}{\Lambda_0(T_i)S_i(u)}d\mathcal{A}_0(u)\Big].$$

Note that under Models (1), (2) and (3), $M_i(t)$ $(i=1,\cdots,n)$ are zero-mean stochastic processes. Thus, for given $(\beta_0,\gamma_0,\alpha_0,\Lambda_0,H_0)$, a reasonable estimator for $\mathcal{A}_0(t)$ is the solution to

$$\sum_{i=1}^n\int_0^t \Big[\big\{Y_i(u)-\beta_0'X_i-\Omega_i(u)\big\}\frac{\Delta_i(u)}{S_i(u)}dN_i(u) - \frac{m_i\Delta_i(u)}{\Lambda_0(T_i)S_i(u)}d\mathcal{A}(u)\Big] = 0.$$

Denote this estimator by $\widehat{\mathcal{A}}(t;\beta_0,\gamma_0,\alpha_0,\Lambda_0,H_0)$, which can be expressed as

$$\widehat{\mathcal{A}}(t;\beta_0,\gamma_0,\alpha_0,\Lambda_0,H_0) = \sum_{i=1}^n\int_0^t \frac{\big\{Y_i(u)-\beta_0'X_i-\Omega_i(u)\big\}\Delta_i(u)S_i(u)^{-1}dN_i(u)}{\sum\limits_{j=1}^n \Delta_j(u)S_j(u)^{-1}m_j\Lambda_0(T_j)^{-1}}.$$

Note that for given $\mathcal{A}_0(t)$ and $(\gamma_0,\alpha_0,\Lambda_0,H_0)$, applying the generalized estimating equation approach[5], we can estimate β_0 using the following unbiased estimating equation

$$n^{-1}\sum_{i=1}^n\int_0^\tau Q(t)X_i\Big[\big\{Y_i(t)-\beta'X_i-\Omega_i(t)\big\}\frac{\Delta_i(t)}{S_i(t)}dN_i(t) - \frac{m_i\Delta_i(t)}{\Lambda_0(T_i)S_i(t)}d\mathcal{A}_0(t)\Big] = 0,$$

where τ is a pre-specified constant, and $Q(t)$ is a possibly data-dependent weight function. Then for given $(\gamma_0,\alpha_0,\Lambda_0,H_0)$, replacing $\mathcal{A}_0(t)$ with the estimator $\widehat{\mathcal{A}}(t;\beta,\gamma_0,\alpha_0,\Lambda_0,H_0)$ in the above estimating function, we can estimate β_0 using the following estimating equation

$$n^{-1}\sum_{i=1}^n\int_0^\tau Q(t)\{X_i-\overline{X}^*(t)\}\big\{Y_i(t)-\beta'X_i-\Omega_i(t)\big\}\frac{\Delta_i(t)}{S_i(t)}dN_i(t) = 0, \tag{4}$$

where

$$\overline{X}^*(t) = \frac{\sum\limits_{i=1}^n X_i\Delta_i(t)S_i(t)^{-1}m_i\Lambda_0(T_i)^{-1}}{\sum\limits_{i=1}^n \Delta_i(t)S_i(t)^{-1}m_i\Lambda_0(T_i)^{-1}}.$$

To guarantee the limit of the denominator of $\overline{X}^*(t)$ to be bounded away from zero, and also to avoid lengthy technical discussion of the tail behavior of the limiting distributions of $\widehat{S}_i(t)$, here τ is chosen such that $P(T_i \geq \tau) > 0$. This means that the supports of the terminating event time and the censoring time is greater than τ. Thus, τ should not be greater than all the observed T_i's.

Since $(\gamma_0, \alpha_0, \Lambda_0, H_0)$ are unknown in practice, we can estimate them by the fit of Models (2) and (3). Specifically, consider Model (2) and define $F(t) = \Lambda_0(t)/\Lambda_0(\tau)$ for $0 \leq t \leq \tau$, which is the corresponding cumulative distribution function of $\Lambda_0(t)$. Then $F(t)$ can be estimated by the nonparametric maximum likelihood estimator $\widehat{F}(t)$, where

$$\widehat{F}(t) = \prod_{t<s\leq\tau} \Big(1 - \frac{\sum_{i=1}^{n} dN_i(s)}{\sum_{i=1}^{n} \Delta_i(s) N_i(s)}\Big)$$

(see [18]) and 0/0 is defined to be 0. Let $Z_i = (1, X_i')'$, $\theta_1 = \log \Lambda_0(\tau)$ and $\theta = (\theta_1, \gamma')'$. Then under Model (2), [18] proposed the following estimating equation for θ:

$$n^{-1} \sum_{i=1}^{n} W_i Z_i \big\{ m_i \widehat{F}^{-1}(T_i) - \exp(\theta' Z_i) \big\} = 0, \tag{5}$$

where W_i is a weight function that could depend on $(X_i, \theta, \widehat{F})$. Let $\widehat{\theta} = (\widehat{\theta}_1, \widehat{\gamma}')'$ denote the solution to the above estimating equation. Note that $\Lambda_0(t) = F(t)\exp(\theta_1)$. Thus, $\Lambda_0(t)$ can be consistently estimated by $\widehat{\Lambda}_0(t) = \widehat{F}(t)\exp(\widehat{\theta}_1)$. Let

$$\widehat{V}_i = \frac{m_i}{\widehat{\Lambda}_0(T_i)\exp(\widehat{\gamma}' X_i)}.$$

Using the method proposed by [3] for Model (3), α_0 can be consistently estimated from the following estimating equation:

$$n^{-1} \sum_{i=1}^{n} \int_0^{\tau} \big\{ X_i - \overline{X}^D(t;\alpha) \big\} dN_i^D(t) = 0, \tag{6}$$

where $N_i^D(t) = I(T_i \leq t, \delta_i = 1)$, and

$$\overline{X}^D(t;\alpha) = \frac{\sum_{i=1}^{n} X_i \widehat{V}_i \Delta_i(t) \exp(\alpha' X_i)}{\sum_{i=1}^{n} \widehat{V}_i \Delta_i(t) \exp(\alpha' X_i)}.$$

Let $\widehat{\alpha}$ denote the solution to the above estimating equation, and $\widehat{H}_0(t)$ be the Breslow-type estimator of $H_0(t)$, where

$$\widehat{H}_0(t) = \int_0^t \frac{\sum_{i=1}^{n} dN_i^D(u)}{\sum_{j=1}^{n} \Delta_j(u) \widehat{V}_j \exp(\widehat{\alpha}' X_j)}.$$

Define

$$\widehat{S}_i(t) = \exp\{-\widehat{V}_i \widehat{H}_0(t) \exp(\widehat{\alpha}' X_i)\},$$

$$\widehat{\Omega}_i(t) = \frac{m_i - 1}{\widehat{\Lambda}_0(T_i)\exp(\widehat{\gamma}' X_i)} \exp\Big\{ - \frac{\widehat{H}_0(t)\exp(\widehat{\alpha}' X_i)}{\widehat{\Lambda}_0(T_i)\exp(\widehat{\gamma}' X_i)} \Big\},$$

and

$$\overline{X}(t) = \frac{\sum_{i=1}^{n} X_i \Delta_i(t) \widehat{S}_i(t)^{-1} m_i \widehat{\Lambda}_0(T_i)^{-1}}{\sum_{i=1}^{n} \Delta_i(t) \widehat{S}_i(t)^{-1} m_i \widehat{\Lambda}_0(T_i)^{-1}}.$$

By replacing $\Omega_i(t)$ and $S_i(t)$ with $\widehat{\Omega}_i(t)$ and $\widehat{S}_i(t)$ in (4), we propose to estimate β_0 using the solution to $U(\beta) = 0$, where

$$U(\beta) = n^{-1} \sum_{i=1}^{n} \int_0^{\tau} Q(t)\{X_i - \overline{X}(t)\}\big[Y_i(t) - \beta' X_i - \widehat{\Omega}_i(t)\big] \frac{\Delta_i(t)}{\widehat{S}_i(t)} dN_i(t) = 0. \tag{7}$$

Let $\widehat{\beta}$ denote the solution to $U(\beta) = 0$. By using the law of large numbers and the consistency of $\widehat{\gamma}$, $\widehat{\alpha}$, $\widehat{\Lambda}_0(t)$ and $\widehat{H}_0(t)$ (see [3, 18]), one can show that $\widehat{\beta}$ is consistent. To establish the asymptotic normality of $\widehat{\beta}$, set

$$\begin{aligned}
\widehat{\Phi}(t) &= n^{-1} \sum_{i=1}^{n} \sum_{j=1}^{m_i} I(t_{ij} \le t), \\
\widehat{R}(t) &= n^{-1} \sum_{i=1}^{n} \sum_{j=1}^{m_i} I(t_{ij} \le t \le T_i), \\
\widehat{b}_i(t) &= \sum_{j=1}^{m_i} \Big\{ \int_t^{\tau} \frac{I(t_{ij} \le u \le T_i) d\widehat{\Phi}(u)}{\widehat{R}^2(u)} - \frac{I(t < t_{ij} \le \tau)}{\widehat{R}(t_{ij})} \Big\}, \\
\widehat{e}_i(\theta) &= -\int \frac{wzm\widehat{b}_i(c) dP_{1n}(w,z,c,m)}{\widehat{F}(c)} + W_i Z_i \Big\{ \frac{m_i}{\widehat{F}(C_i)} - \exp(\theta' X_i) \Big\},
\end{aligned}$$

where $P_{1n}(w,z,c,m)$ is the empirical measure of $\{(W_i, Z_i, T_i, m_i),\ i = 1, \cdots, n\}$. Let $\widehat{\eta}_i(\theta)$ to be the vector function $\big[E\{-\partial \widehat{e}_i(\theta)/\partial\theta\}\big]^{-1} \widehat{e}_i(\theta)$ without its first element, $\widehat{\varsigma}_i(\theta)$ the first element of $(E\{-\partial \widehat{e}_i(\theta)/\partial\theta\})^{-1}\widehat{e}_i(\theta)$ and $\widehat{\phi}_i(t) = \widehat{b}_i(t) + \widehat{\varsigma}_i(\widehat{\theta})$. Define

$$\begin{aligned}
\widehat{B} =& n^{-1} \sum_{i=1}^{n} \int_0^{\tau} \Bigg\{ \frac{\sum_{j=1}^{n} \Delta_j(t) \widehat{V}_j \exp(\widehat{\alpha}' X_j) X_j^{\otimes 2}}{\sum_{j=1}^{n} \Delta_j(t) \widehat{V}_j \exp(\widehat{\alpha}' X_j)} - \overline{X}^D(t; \widehat{\alpha})^{\otimes 2} \Bigg\} dN_i^D(t), \\
\widehat{\sigma}_i =& \widehat{B}^{-1} \int_0^{\tau} \{X_i - \overline{X}^D(t;\widehat{\alpha})\}\big[dN_i^D(t) - \Delta_i(t)\widehat{V}_i \exp(\widehat{\alpha}' X_i) d\widehat{H}_0(t)\big] \\
&+ \int_0^{\tau} \Big[\int \{x - \overline{X}^D(t;\widehat{\alpha})\} \frac{I(c \ge t) m \exp(\widehat{\alpha}' x)}{\widehat{\Lambda}_0(c) \exp(\widehat{\gamma}' x)} \{\widehat{\phi}_i(c) + x' \widehat{\eta}_i\} \widehat{P}_{2n}(x,c,m) \Big] d\widehat{H}_0(t), \\
\widehat{\psi}_i(t) =& \int_0^t \frac{dN_i^D(u) - \Delta_i(u) \widehat{V}_i \exp(\widehat{\alpha}' X_i) d\widehat{H}_0(u)}{n^{-1} \sum_{j=1}^{n} \Delta_j(u) \widehat{V}_j \exp(\widehat{\alpha}' X_j)\}} - \int_0^t \overline{X}^D(u;\widehat{\alpha})' d\widehat{H}_0(u) \widehat{\sigma}_i \\
&- \int_0^t \Big[\int \frac{I(c \ge u) m \exp(\widehat{\alpha}' x)}{\widehat{\Lambda}_0(c)\ \exp(\widehat{\gamma}' x)} \{\widehat{\phi}_i(c) + x' \widehat{\eta}_i\} P_{2n}(x,c,m) \Big] \frac{d\widehat{H}_0(u)}{n^{-1} \sum_{j=1}^{n} \Delta_j(u) \widehat{V}_j \exp(\widehat{\alpha}' X_j)\}}, \\
\widehat{M}_i(t) =& \int_0^t \Big[\big\{ Y_i(u) - \widehat{\beta}' X_i - \widehat{\Omega}_i(u) \big\} \frac{\Delta_i(u)}{\widehat{S}_i(u)} dN_i(u) - \frac{\Delta_i(u)}{\widehat{S}_i(u)} \widehat{V}_i \exp(\widehat{\gamma}' X_i) d\widehat{\mathcal{A}}(u) \Big],
\end{aligned}$$

$$\widehat{\mathcal{A}}(t) = \sum_{i=1}^{n} \int_0^t \frac{\{Y_i(u) - \widehat{\beta}' X_i - \widehat{\Omega}_i(u)\} \Delta_i(u) \widehat{S}_i(u)^{-1} dN_i(u)}{\sum\limits_{i=1}^{n} \Delta_i(t) \widehat{S}_i(u)^{-1} \widehat{V}_i \exp(\widehat{\gamma}' X_i)}.$$

where $a^{\otimes 2} = aa'$ for a vector a, and $P_{2n}(x, c, m)$ is the empirical measure of $\{(X_i, T_i, m_i), i = 1, \cdots, n\}$.

We show in the Appendix I that under some regularity conditions, $n^{1/2}(\widehat{\beta} - \beta_0)$ is asymptotically normal with mean zero and a covariance matrix that can be consistently estimated by $\widehat{A}^{-1} \widehat{\Sigma} \widehat{A}^{-1}$, where $\widehat{\Sigma} = n^{-1} \sum\limits_{i=1}^{n} \widehat{\xi}_i^{\otimes 2}$,

$$\begin{aligned}
\widehat{\xi}_i = & \int_0^\tau Q(t)\{X_i - \overline{X}(t)\} d\widehat{M}_i(t) + \int \sum_{j=1}^{m} Q(t_j)\{x - \overline{X}(t_j)\}\{Y(t_j) - \widehat{\beta}' x\} \\
& \times \exp\left\{\frac{m\widehat{H}_0(t_j)\exp(\widehat{\alpha}' x)}{\widehat{\Lambda}_0(c)\exp(\widehat{\gamma}' x)}\right\} \frac{mI(c \geq t_j)\exp(\widehat{\alpha}' x)}{\widehat{\Lambda}_0(c)\exp(\widehat{\gamma}' x)} \\
& \times \left[\widehat{\psi}_i(t_j) - \widehat{\phi}_i(c)\widehat{H}_0(t_j) + x'(\widehat{\sigma}_i - \widehat{\eta}_i)\widehat{H}_0(t_j)\right] dP_{3n}(x, c, m, y, t_1, \cdots, t_m) \\
& - \int \sum_{j=1}^{m} Q(t_j)\{x - \overline{X}(t_j)\} \exp\left\{\frac{(m-1)\widehat{H}_0(t_j)\exp(\widehat{\alpha}' x)}{\widehat{\Lambda}_0(c)\exp(\widehat{\gamma}' x)}\right\} \frac{(m-1)I(c \geq t_j)}{\widehat{\Lambda}_0(c)\exp(\widehat{\gamma}' x)} \\
& \times \left[\frac{(m-1)\exp(\widehat{\alpha}' x)}{\widehat{\Lambda}_0(c)\exp(\widehat{\gamma}' x)}\left\{\widehat{\psi}_i(t_j) - \widehat{\phi}_i(c)\widehat{H}_0(t_j) + x'(\widehat{\sigma}_i - \widehat{\eta}_i)\widehat{H}_0(t_j)\right\} - \left\{\widehat{\phi}_i(c) + x'\widehat{\eta}_i\right\}\right] \\
& \times dP_{4n}(x, c, m, t_1, \cdots, t_m) \\
& - \int_0^\tau Q(t)\Big[\int \{x - \overline{X}(t)\} \exp\left\{\frac{m\widehat{H}_0(t)\exp(\widehat{\alpha}' x)}{\widehat{\Lambda}_0(c)\exp(\widehat{\gamma}' x)}\right\} \frac{mI(c \geq t)}{\widehat{\Lambda}_0(c)}\Big(\frac{m\exp(\widehat{\alpha}' x)}{\widehat{\Lambda}_0(c)\exp(\widehat{\gamma}' x)} \\
& \times \left\{\widehat{\psi}_i(t) - \widehat{\phi}_i(c)\widehat{H}_0(t) + x'(\widehat{\sigma}_i - \widehat{\eta}_i)\widehat{H}_0(t)\right\} - \widehat{\phi}_i(c)\Big) dP_{2n}(x, c, m)\Big] d\widehat{\mathcal{A}}(t)
\end{aligned}$$

and

$$\widehat{A} = n^{-1} \sum_{i=1}^{n} \int_0^\tau Q(t)\{X_i - \overline{X}(t)\}^{\otimes 2} \frac{\Delta_i(t)}{\widehat{S}_i(t)} dN_i(t).$$

In the above, $P_{3n}(x, c, m, y, t_1, \cdots, t_m)$ and $P_{4n}(x, c, m, t_1, \cdots, t_m)$ denote the empirical measures of $\{(X_i, T_i, m_i, Y_i, t_{i1}, \cdots, t_{i,m_i}), i = 1, \cdots, n\}$ and $\{(X_i, T_i, m_i, t_{i1}, \cdots, t_{i,m_i}), i = 1, \cdots, n\}$, respectively.

Remark 2. The weight functions $Q(t)$ and W_i play a role in finding the estimator which has a relatively small variance. Ideally, we would choose $Q(t)$ and W_i to minimize the variances of $\widehat{\beta}$ and $\widehat{\theta}$. However, it does not appear possible to derive an optimal weight without specification of dependence structures on the longitudinal response process and the observation precess, and the selection of weight functions is usually a complicated problem[6]. Thus, our choice of $Q(t)$ is somewhat ad hoc, such as $Q(t) = 1$ and $W_i = 1$ in the simulation studies of Section 5. Other examples are $Q(t) = n^{-1} \sum\limits_{i=1}^{n} \Delta_i(t)$, $Q(t) = n^{-1} \sum\limits_{i=1}^{n} N_i(t)$, $Q(t) = n^{-1} \sum\limits_{i=1}^{n} I(T_i \leq t)$ and $W_i = Q(\tau)$.

4 Goodness-of-fit Methods

In this section, we propose some simple graphical and numerical procedures for assessing the adequacy of the proposed models. First, by replacing V_i with $\widehat{V}_i$, the standardized score process[7,13] can be used to check the adequacy of Model (3). To check Model (2), we can use some discussion and simple approaches of [3] and [19] for recurrent event data with a terminal event. Here we propose some simple graphical and numerical procedures for assessing the adequacy of Model (1). Following [7] and [15], we consider the following cumulative sums of residuals:

$$\Gamma(t,z)=n^{-1/2}\sum_{i=1}^{n}\int_0^t I\{X_i\le z\}d\widehat{M}_i(u), \tag{8}$$

where $I\{X_i \le z\}$ means that each component of X_i is no larger than the corresponding component of z.

It can be shown that the null distribution of $\Gamma(t,z)$ can be approximated by the zero-mean Gaussian process

$$\begin{aligned}\widetilde{\Gamma}(t,z)=&n^{-1/2}\sum_{i=1}^{n}\int_0^t \Big\{I(X_i\le z)-\overline{X}_\gamma(u,z)\}d\widehat{M}_i(u)+n^{-1/2}\sum_{i=1}^{n}\int\sum_{j=1}^{m}\{I(x\le z)-\overline{X}_\gamma(t_j,z)\Big\}\\
&\times\{y(t_j)-\widehat{\beta}'x\}\exp\Big\{\frac{m\widehat{H}_0(t_j)\exp(\widehat{\alpha}'x)}{\widehat{\Lambda}_0(c)\exp(\widehat{\gamma}'x)}\Big\}\frac{mI(c\wedge t\ge t_j)\exp(\widehat{\alpha}'x)}{\widehat{\Lambda}_0(c)\exp(\widehat{\gamma}'x)}\\
&\times\big[\widehat{\psi}_i(t_j)-\widehat{\phi}_i(c)\widehat{H}_0(t_j)+x'(\widehat{\sigma}_i-\widehat{\eta}_i)\widehat{H}_0(t_j)\big]dP_{3n}(x,c,m,y,t_1,\cdots,t_m)\\
&-n^{-1/2}\sum_{i=1}^{n}\int\sum_{j=1}^{m}\{I(x\le z)-\overline{X}_\gamma(t_j,z)\}\exp\Big\{\frac{(m-1)\widehat{H}_0(t_j)\exp(\widehat{\alpha}'x)}{\widehat{\Lambda}_0(c)\exp(\widehat{\gamma}'x)}\Big\}\\
&\times\frac{(m-1)I(c\wedge t\ge t_j)}{\widehat{\Lambda}_0(c)\exp(\widehat{\gamma}'x)}\Big[\frac{(m-1)\exp(\widehat{\alpha}'x)}{\widehat{\Lambda}_0(c)\exp(\widehat{\gamma}'x)}\{\widehat{\psi}_i(t_j)-\widehat{\phi}_i(c)\widehat{H}_0(t_j)\\
&+x'(\widehat{\sigma}_i-\widehat{\eta}_i)H_0(t_j)\}-\big\{\widehat{\phi}_i(c)+x'\widehat{\eta}_i\big\}\Big]dP_{4n}(x,c,m,t_1,\cdots,t_m)\\
&-n^{-1/2}\sum_{i=1}^{n}\int_0^t\Big[\int\{I(x\le z)-\overline{X}_\gamma(u,z)\}\frac{mI(c\wedge t\ge u)}{\widehat{\Lambda}_0(c)}\exp\Big\{\frac{m\widehat{H}_0(u)\exp(\widehat{\alpha}'x)}{\widehat{\Lambda}_0(c)\exp(\widehat{\gamma}'x)}\Big\}\\
&\times\Big(\frac{m\exp(\widehat{\alpha}'x)}{\widehat{\Lambda}_0(c)\exp(\widehat{\gamma}'x)}\{\widehat{\psi}_i(u)-\widehat{\phi}_i(c)\widehat{H}_0(u)+x'(\widehat{\sigma}_i-\widehat{\eta}_i)\widehat{H}_0(u)\}\\
&-\widehat{\phi}_i(c)\Big)dP_{2n}(x,c,m)\Big]d\widehat{\mathcal{A}}(u)-\widehat{B}(t)'\widehat{A}^{-1}n^{-1/2}\sum_{i=1}^{n}\widehat{\xi}_i+o_p(1),\end{aligned} \tag{9}$$

where

$$\overline{X}_\gamma(u,z)=\frac{\sum_{i=1}^{n}I\{X_i(u)\le z\}\Delta_i(u)\widehat{S}_i(u)^{-1}m_i\widehat{\Lambda}_0(T_i)^{-1}}{\sum_{i=1}^{n}\Delta_i(u)\widehat{S}_i(u)^{-1}m_i\widehat{\Lambda}_0(T_i)^{-1}},$$

$$\widehat{B}(t)=n^{-1}\sum_{i=1}^{n}\int_0^t\big[I\{X_i(u)\le z\}-\overline{X}_\gamma(u,z)\big]X_i\frac{\Delta_i(u)}{\widehat{S}_i(u)}dN_i(u),$$

and all other variables are defined in the last and second-last paragraphs of Section 3.

It is difficult to estimate the asymptotic covariance function of $\Gamma(t,z)$ analytically because the limiting process of $\Gamma(t,z)$ does not have an independent increments structure. To handle this problem, we can appeal to the resampling approach[7] and show that the null distribution of $\Gamma(t,z)$ can be approximated by the conditional distribution of $\widehat{\Gamma}(t,z)$, where

$$\widehat{\Gamma}(t,z)=n^{-1/2}\sum_{i=1}^{n}\int_0^t\{I(X_i\le z)-\overline{X}_\gamma(u,z)\}d\widehat{M}_i(u)\mathcal{Z}_i+n^{-1/2}\sum_{i=1}^{n}\int\sum_{j=1}^{m}\{I(x\le z)-\overline{X}_\gamma(t_j,z)\}$$

$$
\begin{aligned}
&\times \{y(t_j) - \widehat{\beta}' x\} \exp\Big\{ \frac{m\widehat{H}_0(t_j)\exp(\widehat{\alpha}' x)}{\widehat{\Lambda}_0(c)\exp(\widehat{\gamma}' x)} \Big\} \frac{mI(c \wedge t \geq t_j)\exp(\widehat{\alpha}' x)}{\widehat{\Lambda}_0(c)\exp(\widehat{\gamma}' x)} \\
&\times \big[\widehat{\psi}_i(t_j) - \widehat{\phi}_i(c)\widehat{H}_0(t_j) + x'(\widehat{\sigma}_i - \widehat{\eta}_i)\widehat{H}_0(t_j)\big] dP_{3n}(x, c, m, y, t_1, \cdots, t_m)\mathcal{Z}_i \\
&- n^{-1/2}\sum_{i=1}^{n}\int\sum_{j=1}^{m}\{I(x \leq z) - \overline{X}_\gamma(t_j, z)\}\exp\Big\{ \frac{(m-1)\widehat{H}_0(t_j)\exp(\widehat{\alpha}' x)}{\widehat{\Lambda}_0(c)\exp(\widehat{\gamma}' x)} \Big\} \\
&\times \frac{(m-1)I(c \wedge t \geq t_j)}{\widehat{\Lambda}_0(c)\exp(\widehat{\gamma}' x)} \Big[\frac{(m-1)\exp(\widehat{\alpha}' x)}{\widehat{\Lambda}_0(c)\exp(\widehat{\gamma}' x)} \{\widehat{\psi}_i(t_j) - \widehat{\phi}_i(c)\widehat{H}_0(t_j) \\
&+ x'(\widehat{\sigma}_i - \widehat{\eta}_i)H_0(t_j)\} - \big\{\widehat{\phi}_i(c) + x'\widehat{\eta}_i\big\} \Big] dP_{4n}(x, c, m, t_1, \cdots, t_m)\mathcal{Z}_i \\
&- n^{-1/2}\sum_{i=1}^{n}\int_0^t \Big[\int \{I(x \leq z) - \overline{X}_\gamma(u, z)\} \frac{mI(c \wedge t \geq u)}{\widehat{\Lambda}_0(c)} \exp\Big\{ \frac{m\widehat{H}_0(u)\exp(\widehat{\alpha}' x)}{\widehat{\Lambda}_0(c)\exp(\widehat{\gamma}' x)} \Big\} \\
&\times \Big(\frac{m\exp(\widehat{\alpha}' x)}{\widehat{\Lambda}_0(c)\exp(\widehat{\gamma}' x)} \{\widehat{\psi}_i(u) - \widehat{\phi}_i(c)\widehat{H}_0(u) + x'(\widehat{\sigma}_i - \widehat{\eta}_i)\widehat{H}_0(u)\} \\
&- \widehat{\phi}_i(c)\Big) dP_{2n}(x, c, m)\Big] d\widehat{\mathcal{A}}(u)\mathcal{Z}_i - \widehat{B}(t)'\widehat{A}^{-1} n^{-1/2}\sum_{i=1}^{n}\widehat{\xi}_i\mathcal{Z}_i,
\end{aligned}
$$

where $(\mathcal{Z}_1, \cdots, \mathcal{Z}_n)$ are independent standard normal variables which are independent of the observed data, and all other variables are the same as in (9). Thus, one can obtain a large number of realizations from $\widehat{\Gamma}(t, z)$ by repeatedly generating the standard normal random sample $(\mathcal{Z}_1, \cdots, \mathcal{Z}_n)$ while fixing the observed data, and may plot $\Gamma(t, z)$ along with a few realizations of $\widehat{\Gamma}(t, z)$. Since the validity of approximating $\Gamma(t, z)$ by $\widehat{\Gamma}(t, z)$ depends on the correct specification of Model (1), an unusual pattern of $\Gamma(t, z)$ compared to the realizations of $\widehat{\Gamma}(t, z)$ would suggest a lack-of-fit of Model (1). Because $\Gamma(t, z)$ is expected to fluctuate randomly around 0 under Model (1), a formal goodness-of-fit test may be constructed based on the supremum statistic $\sup_{0 \leq t \leq \tau, z} |\Gamma(t, z)|$, with which the p-value can be obtained by comparing the observed value of $\sup_{0 \leq t \leq \tau, z} |\Gamma(t, z)|$ to a large number of realizations from $\sup_{0 \leq t \leq \tau, z} |\widehat{\Gamma}(t, z)|$.

5 Numerical Results

In this section, we conducted simulation studies to assess the performance of the proposed estimators with the focus on estimating β_0. In the study, for subject i, we considered two situations for the X_i's:

(i) X_i was generated from a normal distribution $N(0.5, 0.5)$;

(ii) X_i was generated from a Bernoulli distribution with success probability 0.5. For given X_i from the normal distribution $N(0.5, 0.5)$, following [15], we generated the latent variable V_i by setting $V_i = \exp\{-\ln(2.75)I(X_i \geq 0.5)\}V_i^*$ with V_i^* generated from the density function

$$f(v^*|X_i) = I(X_i \leq 0.5)I(0.5 \leq v^* \leq 1.5) + I(X_i \geq 0.5)I(1.5 \leq v^* \leq 4)/2.5.$$

For given X_i from the Bernoulli distribution, following [18], we used $V_i = \exp\{-\ln(2.75)X_i\}V_i^*$ with V_i^* generated from the density function

$$f(v^*|X_i) = (1 - X_i)I(0.5 \leq v^* \leq 1.5) + X_iI(1.5 \leq v^* \leq 4)/2.5.$$

It can be verified that $E(V_i|X_i) = 1$. The censoring times C_i's were generated from a uniform distribution $U(2, 5)$.

For given X_i and V_i, the observation times T_{ij}'s were generated from Model (2) with $\gamma_0 = 0.5$ and $\lambda_0(t) = 2$ (homogeneous Poisson process) or $\lambda_0(t) = 2t + (t-3)^2/360$ (nonhomogeneous Poisson process). The terminal event times were generated from Model (3) with $h_0(t) = 1$ and $\alpha_0 = 0.5$ or -0.5. When X_i was generated from the normal distribution $N(0.5, 0.5)$ and $\alpha_0 = 0.5$ or -0.5, the censoring rates for the terminal event were about 27% and 42%, respectively. When X_i was generated from the Bernoulli distribution and $\alpha_0 = 0.5$ or -0.5, the censoring rates were about 25% and 40%, respectively. Here the censoring rates are observed in the simulation.

For the longitudinal response, we assumed that $Y_i(t)$ was given by

$$Y_i(t) = V_i + \mu_0(t) + \beta_0 X_i + \varepsilon_i,$$

where the ε_i's were normal with mean 0 and standard deviation 0.25 for all $t,\ i = 1, \cdots, n$. For $\mu_0(t)$, we considered two situations. One is $\mu_0(t) = (t+1)/5$ and the other is $\mu_0(t) = 2t^{-1/2}+1$. The regression parameter β_0 was chosen to be $-0.5, 0$ and 0.5. For each simulation study, we took $W_i = 1$, $Q(t) = 1$ and $\tau = 3$. Here we chose $\tau = 3$ to guarantee $P(T_i \geq \tau) > 0$. The results presented below are based on 1000 replications and the sample size $n = 200$.

Table 1. Simulation Results for Estimation of β_0 when the Observation Times Follow a Homogeneous Poisson Process

			$X\sim$Normal				$X\sim$Bernoulli			
$\mu_0(t)$	α_0	β_0	Bias	SEE	SE	CP	Bias	SEE	SE	CP
$(t+1)/5$	-0.5	-0.5	-0.0014	0.0957	0.0881	0.932	-0.0069	0.0946	0.0896	0.931
		0	0.0004	0.0936	0.0880	0.928	-0.0007	0.0912	0.0899	0.949
		0.5	-0.0071	0.0942	0.0887	0.931	-0.0027	0.0938	0.0903	0.931
	0.5	-0.5	-0.0009	0.1058	0.0985	0.943	0.0028	0.1051	0.1004	0.939
		0	-0.0019	0.1098	0.0998	0.924	0.0003	0.1027	0.1001	0.939
		0.5	0.0031	0.1042	0.0984	0.929	-0.0013	0.1080	0.1004	0.942
$2t^{-1/2}+1$	-0.5	-0.5	-0.0057	0.2183	0.2172	0.951	-0.0073	0.2177	0.2140	0.944
		0	-0.0039	0.2420	0.2308	0.938	-0.0127	0.2559	0.2586	0.959
		0.5	-0.0087	0.2038	0.2026	0.946	-0.0047	0.2226	0.2113	0.939
	0.5	-0.5	0.0147	0.1946	0.1946	0.954	0.0098	0.2024	0.2048	0.952
		0	0.0023	0.2386	0.2425	0.946	0.0068	0.2690	0.2689	0.942
		0.5	-0.0040	0.2204	0.2146	0.936	-0.0171	0.3358	0.3288	0.944

Table 2. Simulation Results for Estimation of β_0 when the Observation Times Follow a Nonhomogeneous Poisson Process

			$X\sim$Normal				$X\sim$Bernoulli			
$\mu_0(t)$	α_0	β_0	Bias	SEE	SE	CP	Bias	SEE	SE	CP
$(t+1)/5$	-0.5	-0.5	0.0007	0.1317	0.1328	0.935	0.0026	0.1257	0.1146	0.939
		0	-0.0032	0.1315	0.1311	0.936	0.0014	0.1202	0.1143	0.951
		0.5	-0.0016	0.1214	0.1136	0.945	0.0047	0.1159	0.1129	0.942
	0.5	-0.5	0.0046	0.1555	0.1406	0.929	0.0056	0.1470	0.1344	0.933
		0	0.0028	0.2247	0.2153	0.944	0.0044	0.1427	0.1324	0.934
		0.5	-0.0092	0.2102	0.2208	0.939	-0.0053	0.1481	0.1344	0.933
$2t^{-1/2}+1$	-0.5	-0.5	-0.0011	0.1280	0.1173	0.922	0.0007	0.1258	0.1141	0.937
		0	-0.0032	0.1191	0.1143	0.947	-0.0029	0.1249	0.1165	0.936
		0.5	0.0073	0.2061	0.2227	0.946	-0.0022	0.1296	0.1170	0.935
	0.5	-0.5	0.0037	0.1543	0.1454	0.945	-0.0001	0.1479	0.1357	0.931
		0	0.0005	0.1678	0.1674	0.931	0.0027	0.1513	0.1376	0.932
		0.5	0.0042	0.1474	0.1360	0.930	0.0019	0.1498	0.1366	0.945

Table 1 presents the simulation results for estimation of β_0 when the observation times follow a homogeneous Poisson process. Table 2 gives the simulation results for the same setup as in Table 1, except that the observation times follow a nonhomogeneous Poisson process. In these tables, Bias is the sample mean of the estimate minus the true value, SSE is the sampling standard error of the estimate, SE is the sampling mean of the standard error estimate and CP is the 95% empirical coverage probability for β_0 based on a normal approximation. Tables 1 and 2 show that the proposed estimators are virtually unbiased, the estimated standard errors are practically close to the empirical standard errors, and the empirical 95% confidence intervals have reasonable coverage rates.

For comparison, we also considered the method of [10] (denoted by LHO), using the same setup as in Tables 1 and 2 with $\mu_0(t) = (t+1)/5$ and $n = 200$. Note that in these situations, the random effects were misspecified for the LHO's method. Thus, we considered another model that was correctly specified for the LHO's method, in which

$$h(t|X_i, \mu_i) = 0.5t \exp(X_i + \mu_i), \qquad Y_i(t) = 0.2t + \beta_0 X_i + \gamma_1 \mu_i + \nu_i + \varepsilon_i$$

and

$$\lambda(t|X_i, \mu_i, \nu_i) = \lambda_0(t) \exp(0.5X_i + \gamma_2 \mu_i + \gamma_3 \nu_i),$$

where ε_i was the standard normal random variable, μ_i and ν_i were independent and identically distributed $N(0, \sigma^2)$ with $\sigma^2 = 0.5$ or 0.25, and $(\gamma_1, \gamma_2, \gamma_3)'$ were taken as $(\exp(\sigma), 1, 0)'$. The other setups were the same as in Tables 1 and 2 with $\beta_0 = 1$ and X_i from the Bernoulli distribution, except that $\lambda_0(t) = 1$ (homogeneous Poisson process) or $\lambda_0(t) = 2t$ (nonhomogeneous Poisson process). Note that the LHO's models were misspecified for our proposed method. Table 3 gives the comparison results on estimation of β_0 with all weight functions taken to be 1 and $n = 200$. It can be seen from the table that the LHO's method works well when the model is correctly specified. However, it fails when the model is misspecified. Under the LHO's models, the proposed estimators are still unbiased when the observation process is a homogeneous Poisson process, but when the observation process is a nonhomogeneous Poisson process, the proposed estimators are biased. Thus, these two methods are not comparable. We also considered other setups and the results were similar to those given above.

Table 3. Simulation Results for Comparing our Method with the LHO's Method when $n = 200$

		$\lambda_0(t) = 1$				$\lambda_0(t) = 2t$			
		Ours		LHO		Ours		LHO	
α_0	β_0	Bias	SEE	Bias	SEE	Bias	SEE	Bias	SEE
−0.5	−0.5	0.0023	0.1783	0.2316	0.0649	0.0126	0.2088	0.2262	0.0683
	0	0.0067	0.1816	0.2310	0.0616	−0.0108	0.2188	0.2244	0.0662
	0.5	−0.0105	0.1858	0.2320	0.0627	0.0039	0.2094	0.2260	0.0666
0.5	−0.5	−0.0188	0.1317	0.2216	0.0662	0.0006	0.1719	0.1766	0.0589
	0	−0.0086	0.1319	0.2245	0.0656	−0.0157	0.1749	0.1806	0.0574
	0.5	−0.0127	0.1381	0.2221	0.0661	−0.0079	0.1771	0.1772	0.0550
M1		−0.0138	0.3933	0.0055	0.1615	−0.1943	0.4869	0.0063	0.1385
M2		−0.0111	0.4409	0.0059	0.2052	−0.1585	0.6677	−0.0035	0.1990

Note: M1 and M2 stand for the specific models of Liu, Huang and O'Quigley (2008) with $\sigma^2 = 0.5$ and 0.25, respectively.

To examine the performance of the model checking method, we conducted additional simulation studies to assess the size and power of the test based on $\sup_{0 \le t \le \tau, z} |\widehat{\Gamma}(t, z)|$ with $n = 100$. We first generated the covariate X_i taking values $0, 1, 2, 3$ and 4 with equal probabilities. The censoring time, the terminal event and the observation times were generated in the same way as in Tables 1 and 2 with $\gamma_0 = -0.2,\ \ \alpha_0 = 0.3,\ \ h_0(t) = t/5$, and $\lambda_0(t) = 2$ or $\lambda_0(t) = 2 + t/5$. We assumed that the longitudinal response $Y_i(t)$ was given by

$$Y_i(t) = \mu_0(t) + \beta_0 X_i + 0.1k^2 X_i^2 + V_i + \varepsilon_i$$

with $k = 0, 1, 2, 3$ and 4, where ε_i was the standard normal random variable, $\mu_0(t) = t$, and $\beta_0 = -0.5, 0$ or 0.5. We considered the null hypothesis H_0 as $k = 0$. Table 4 shows the empirical sizes and powers of the proposed test at the significance level of 0.05. The results indicate that the empirical sizes are close to the nominal size, and the test has a reasonable power to detect deviations from the null hypothesis. As anticipated, the power increases as k increases.

Table 4. The Empirical Sizes and Powers of the Model Checking Method with $n = 100$

$\lambda_0(t)$	β_0	$k=0$	$k=1$	$k=2$	$k=3$	$k=4$
2	-0.5	0.042	0.098	0.707	0.947	0.960
	0	0.047	0.098	0.661	0.948	0.963
	0.5	0.054	0.091	0.674	0.949	0.957
$2+t/5$	-0.5	0.060	0.106	0.736	0.963	0.965
	0	0.054	0.124	0.743	0.947	0.971
	0.5	0.060	0.091	0.741	0.950	0.969

6 An Application

For illustration purpose, we applied the proposed method to a longitudinal bladder cancer data[4,14,15]. In the study, the patients were randomly assigned to placebo and thiotepa treatment groups. The data include 85 bladder cancer patients, 47 in the placebo group and 38 in the thiotepa treatment group. For each patient, the observed information includes the clinical visit times (in month) and the number of bladder tumors that occurred between clinical visits. The frequency of visits ranges from 1 to 38, and the total follow-up is 53 months. About 25.88% of patients died during follow-up, others were censored. Two baseline covariates were measured and they are the number of initial tumors before entering the study and the size of the largest initial tumor. Since patients with different visiting times and frequencies seem to have different tumor recurrence rates, we focus on the effects of thiotepa treatment and the number of initial tumors on the recurrence rate of bladder tumor with informative observation times and a dependent terminal event (death).

Table 5. Analysis of the Bladder Cancer Data

		Est	SE	p-value	Est	SE	p-value
Our method	Survival time						
	α_1	0.3377	2.5054	0.8928			
	α_2	0.1629	0.0771	0.0348			
	Visiting process						
	γ_1	0.4587	0.1215	0.0002			
	γ_2	-0.0364	0.0348	0.2956			
	Tumor recurrence		$Q_1(t)$			$Q_2(t)$	
	β_1	-0.4888	0.1236	0.0001	-0.5722	0.1239	0.0000
	β_2	0.0215	0.0349	0.5388	0.0224	0.0321	0.4852
SSL's method	Tumor recurrence		$Q_1(t)$			$Q_2(t)$	
	β_1	-0.4656	0.1137	0.0000	-0.5242	0.1134	0.0000
	β_2	0.0145	0.0331	0.6613	0.0093	0.0322	0.7727

Note: Est is the estimate of the parameter, SE is the standard error estimate, and SSL's method stands for the method of [15].

For the analysis, we defined $Y_i(t)$ as the natural logarithm of the number of observed tumors at time t plus 1 to avoid 0, $i = 1, \cdots, 85$. For covariates, let $X_{i1} = 1$ if the patient was in the thiotepa group and 0 if the patient was in the placebo group, and X_{i2} as the number of initial tumors. We took the weight functions $W_i = 1$, $Q(t) = Q_1(t) = 1$ or $Q(t) = Q_2(t) = n^{-1} \sum_{i=1}^{n} \Delta_i(t)$. The results are summarized in Table 5. It can be seen that the thiotepa treatment significantly reduces the occurrence rate of the bladder tumors, but the number of initial tumors seems to have no significant effect on the tumor occurrence process. Also, the thiotepa treatment has a significant positive effect on the visiting process, whereas the initial number of tumors shows no effect on the visiting process. In addition, both the initial number of tumor and the thiotepa treatment seem to have no significant effect on the death rate, although there is slight evidence that the initial number of tumor might be predictive of survival. This is because in the bladder cancer study, most of the deaths are caused by other reasons rather than the bladder cancer. For comparison, Table 5 also gives the results of [15]. It can be seen that our results are consistent with those obtained by [15].

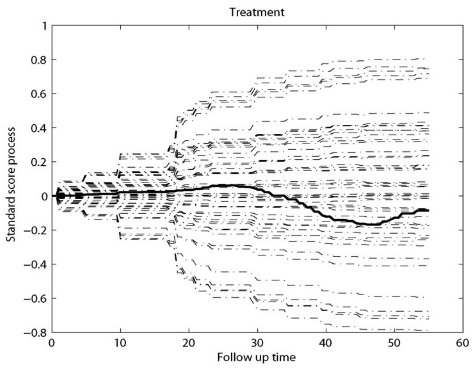

Figure 1. Plot of the standardized score process versus follow-up time for X_{i1} in Model (3). Bold line: observed process; dash-dotted lines: 50 simulated processes

For model checking, we first used the standardized score process[7] to check the adequacy of Model (3). Figure 1 presents the observed score process for X_{i1} along with 50 simulated processes. The plot for X_{i2} is similar and thus ignored. These plots shows that the observed score processes appear to be within the normal ranges. Thus, there is no evidence against Model (3). To check Model (2), note that conditioning on X, the rate function of $N(t)$ is $\lambda_0(t)\exp(\gamma_0' X)$. Following [19], we examined the total summation of the residuals for each subject $\int_0^{T_i} [dN_i(t) - \exp(\widehat{\gamma}' X_i) d\widehat{\Lambda}(t)]$, which has an approximate mean zero and should be independent of X_i under Model (2). Thus, a simple graphical procedure for assessing the adequacy of the assumed model is to plot the residuals against the covariate X_i's. As an example, Figure 2 displays the residuals for each subject versus X_{i2} with the placebo group. Other residual plots are similar and thus ignored. These results show that the residuals fluctuate around zero and seem to be random, indicating little evidence against Model (2). Finally, we apply the goodness-of-fit method to assess the adequacy of Model (1) for the bladder cancer data. We calculated the statistic $\Gamma(t, z)$, and obtained $\sup_{0 \le t \le \tau, z} |\Gamma(t, z)| = 5.1535$ and 3.9674 with p-values of 0.8137 and 0.8912, under $Q_1(t)$ and $Q_2(t)$, respectively, based on 10000 realizations of the corresponding statistic $\sup_{0\le t\le\tau, z} |\widehat{\Gamma}(t,z)|$. These results suggest that Model (1) fits the data well.

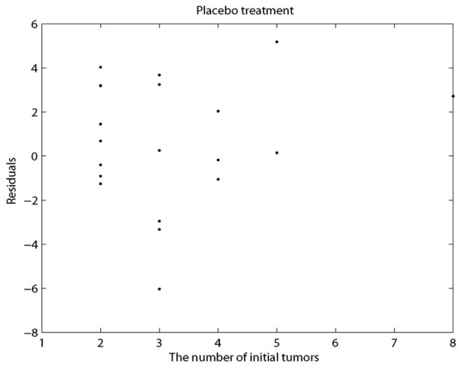

Figure 2. The residual plot of the visiting process versus the number of initial tumors for the subjects with the placebo group

7 Concluding Remarks

In this article, we proposed a joint model for analyzing longitudinal data via a latent variable when there exist informative observation times and a dependent terminal event. A borrow-strength estimation procedure was developed for parameter estimation, which yields consistent and asymptotically normal estimators. The simulation results suggested that the proposed method works well for the situations considered. An application to a bladder cancer study was illustrated.

Note that Models (1), (2) and (3) only allow a positive correlation between the longitudinal process $Y(t)$, the observation process $N(t)$ and the terminal event time D. However, the negative correlation may exist in some applications. A more general approach is to generalize Model (1) to

$$E\{Y(t)|X,V\} = \mu_0(t) + \beta_0' X + \sigma V,$$

where σ is an unknown parameter. Note that for identifiability reasons we assume $E(V|X) = 1$. Thus, the additional parameter σ is estimated, and is not confounded with the latent variable V. The proposed estimation procedure can be extended in a straightforward manner to deal with this model. The resulting inference procedure, however, would be much more complicated[16].

Furthermore, it would be more general to assume different but correlated frailty variables for each outcome. For example, let (V_1, V_2, V_3) be three latent variables with $V_2 > 0$ and $V_3 > 0$. Then Models (1), (2) and (3) can be generalized to

$$E\{Y(t)|X, V_1, V_2, V_3\} = \mu_0(t) + \beta_0' X + V_1,$$
$$\lambda(t|X, V_1, V_2, V_3) = V_2 \lambda_0(t) \exp(\gamma_0' X)$$

and

$$h(t|X, V_1, V_2, V_3) = V_3 h_0(t) \exp(\alpha_0' X).$$

When there are linear relationships among V_1, V_2 and V_3 (see [4]), under some identifiability conditions, our proposed estimation procedure can be extended in a straightforward manner to deal with these models using a borrow-strength estimation procedure. When the dependence structure of the three latent variables is not linear or is left unspecified, the proposed estimation procedure cannot be extended in a straightforward manner to deal with this situation. This is a challenging problem and requires further research efforts.

Here, we have used the proportional frailty model for the terminal event. Other competing models, such as the additive hazards model and the accelerated failure time model with frailty may be used as well. In addition, Model (3) can be generalized to

$$h(t|X,V) = V^{\rho} h_0(t)\exp(\alpha_0' X),$$

where ρ is an unknown parameter. It seems not to be straightforward to generalize the proposed approach to this situation. Further research is needed to address this issue.

8 Appendix: Asymptotic Normality of $\widehat{\beta}$

In order to study the asymptotic distribution of $\widehat{\beta}$, we need the following regularity conditions:

(C1) $P(T \geq \tau, V > 0) > 0$, and $E(V^2) < \infty$.

(C2) X is bounded, and $G(t) = E\{V\exp(\gamma_0' X)I(T \geq t)\}$ is a continuous function for $t \in [0,\tau]$.

(C3) The weight function $Q(t)$ has bounded variation and converges to a deterministic function $q(t)$ in probability uniformly in $t \in [0,\tau]$.

(C4) A is nonsingular, where

$$A = E\Big\{\int_0^{\tau} q(t)\{X_i - \overline{x}(t)\}^{\otimes 2}\frac{\Delta_i(t)}{S_i(t)}dN_i(t)\Big\}$$

and $\overline{x}(t)$ is the limit of $\overline{X}(t)$ defined in (7).

Define $R(t) = G(t)\Lambda_0(t)$, $\Phi(t) = \int_0^t G(u)d\Lambda_0(u)$,

$$b_i(t) = \sum_{j=1}^{m_i}\Big\{\int_t^{\tau}\frac{I(t_{ij} \leq u \leq T_i)d\Phi(u)}{R^2(u)} - \frac{I(t < t_{ij} \leq \tau)}{R(t_{ij})}\Big\}$$

and

$$e_i(\theta) = -\int\frac{wzmb_i(c)dP_1(w,z,c,m)}{F(c)} + W_i Z_i\Big\{\frac{m_i}{F(T_i)} - \exp(\theta' Z_i)\Big\},$$

where $P_1(w,z,c,m)$ is the joint probability measure of (W_i, Z_i, T_i, m_i). Note that $\widehat{R}(t)$ and $\widehat{\Phi}(t)$ are unbiased estimators for $R(t)$ and $\Phi(t)$ (see [18]), where $\widehat{R}(t)$ and $\widehat{\Phi}(t)$ are defined in the second last paragraph of Section 3. Denote $\eta_i(\theta)$ the vector function $(-\partial e_i(\theta)/\partial\theta)^{-1}e_i(\theta)$ without the first entry, $\varsigma_i(\theta)$ the first entry of $(-\partial e_i(\theta)/\partial\theta)^{-1}\ e_i(\theta)$, and $\phi_i(t;\theta) = b_i(t) + \varsigma_i(\theta)$. Under Conditions (C1) and (C2), it follows from [18] that

$$n^{1/2}(\widehat{\gamma} - \gamma_0) = n^{-1/2}\sum_{i=1}^{n}\eta_i + o_p(1), \tag{8.1}$$

and

$$n^{1/2}\{\widehat{\Lambda}_0(t) - \Lambda_0(t)\} = n^{-1/2}\Lambda_0(t)\sum_{i=1}^{n}\phi_i(t) + o_p(1), \tag{8.2}$$

where $\eta_i \equiv \eta_i(\theta_0)$, $\phi_i(t) \equiv \phi_i(t;\theta_0)$, and θ_0 is the true value of θ. Let B be the limit of $\widehat{B}$ and $\overline{x}^D(t)$ be the limit of $\overline{X}^D(t;\alpha_0)$, where $\overline{X}^D(t;\alpha_0)$ and $\widehat{B}$ are defined in (6) and the second last paragraph of Section 3, respectively. Then by [3],

$$n^{1/2}(\widehat{\alpha}-\alpha_0)=n^{-1/2}\sum_{i=1}^{n}\sigma_i+o_p(1), \tag{8.3}$$

and

$$n^{1/2}\{\widehat{H}_0(t)-H_0(t)\}=n^{-1/2}\sum_{i=1}^{n}\psi_i(t)+o_p(1), \tag{8.4}$$

where

$$\begin{aligned}
\sigma_i =&B^{-1}\int_0^{\tau}\{X_i-\overline{x}^D(t)\}\Big[dN_i^D(t)-\frac{m_i\exp(\alpha_0'X_i)}{\Lambda_0(T_i)\exp(\gamma_0'X_i)}\Delta_i(t)dH_0(t)\Big]\\
&+\int_0^{\tau}\Big[\int\{x-\overline{x}^D(t)\}\frac{I(c\geq t)m\exp(\alpha_0'x)}{\Lambda_0(c)\exp(\gamma_0'x)}\{\phi_i(c)+x'\eta_i\}P_2(x,c,m)\Big]dH_0(t),\\
\psi_i(t) =&\int_0^{t}\frac{1}{E\{V_i\exp(\alpha_0'X_i)\Delta_i(u)\}}\Big[dN_i^D(u)-\frac{\Delta_i(u)m_i\exp(\alpha_0'X_i)}{\Lambda_0(T_i)\exp(\gamma_0'X_i)}dH_0(u)\Big]\\
&-\int_0^{t}\Big[\int\frac{I(c\geq u)m\exp(\alpha_0'x)}{\Lambda_0(c)\exp(\gamma_0'x)}\{\phi_i(c)+x'\eta_i\}P_2(x,c,m)\Big]\frac{dH_0(u)}{E\{V_i\exp(\alpha_0'X_i)\Delta_i(u)\}}\\
&-\int_0^{t}\overline{x}^D(u)'dH_0(u)\sigma_i,
\end{aligned}$$

$P_2(x,c,m)$ is the joint probability measure of (X_i,T_i,m_i), and η_i and $\phi_i(c)$ are defined in (A.1) and (A.2), respectively. By the functional delta method,

$$\begin{aligned}
\frac{1}{\widehat{S}_i(t)}-\frac{1}{S_i(t)}=&\exp\Big\{\frac{m_iH_0(t)\exp(\alpha_0'X_i)}{\Lambda_0(T_i)\exp(\gamma_0'X_i)}\Big\}\frac{m_i\exp(\alpha_0'X_i)}{\Lambda_0(T_i)\exp(\gamma_0'X_i)}\Big[\{\widehat{H}_0(t)-H_0(t)\}-\frac{H_0(t)}{\Lambda_0(T_i)}\\
&\times\{\widehat{\Lambda}_0(T_i)-\Lambda_0(T_i)\}+X_i'\{(\widehat{\alpha}-\alpha_0)-(\widehat{\gamma}-\gamma_0)\}H_0(t)\Big]+o_p(n^{-1/2}).
\end{aligned} \tag{8.5}$$

Thus, it follows from (A.1)–(A.5) that

$$\begin{aligned}
&n^{-1/2}\sum_{i=1}^{n}\int_0^{\tau}q(t)\{X_i-\overline{x}(t)\}\{Y_i(t)-\beta_0'X_i\}\Big[\frac{1}{\widehat{S}_i(t)}-\frac{1}{S_i(t)}\Big]\Delta_i(t)dN_i(t)\\
=&n^{-1/2}\sum_{i=1}^{n}\int\sum_{j=1}^{m}q(t_j)\{x-\overline{x}(t_j)\}\{y(t_j)-\beta_0'x\}\\
&\times\exp\Big\{\frac{mH_0(t_j)\exp(\alpha_0'x)}{\Lambda_0(c)\exp(\gamma_0'x)}\Big\}\frac{mI(c\geq t_j)\exp(\alpha_0'x)}{\Lambda_0(c)\exp(\gamma_0'x)}\\
&\times\Big[\psi_i(t_j)-\phi_i(c)H_0(t_j)+x'(\sigma_i-\eta_i)H_0(t_j)\Big]dP_3(x,c,m,y,t_1,\cdots,t_m)+o_p(1),
\end{aligned} \tag{8.6}$$

where $P_3(x,c,m,y,t_1,\cdots,t_m)$ is the joint probability measure of $(X_i,T_i,m_i,Y_i,t_{i1},\cdots,t_{i,m_i})$.

In a similar manner, together with (8.1)–(8.5), we obtain

$$\begin{aligned}
&n^{-1/2}\sum_{i=1}^{n}\int_0^{\tau} q(t)\{X_i-\overline{x}(t)\}\Big[\frac{\widehat{\Omega}_i(t)}{\widehat{S}_i(t)}-\frac{\Omega_i(t)}{S_i(t)}\Big]\Delta_i(t)dN_i(t)\\
=&n^{-1/2}\sum_{i=1}^{n}\int\sum_{j=1}^{m} q(t_j)\{x-\overline{x}(t_j)\}\exp\Big\{\frac{(m-1)H_0(t_j)\exp(\alpha_0'x)}{\Lambda_0(c)\exp(\gamma_0'x)}\Big\}\\
&\times\frac{(m-1)I(c\ge t_j)}{\Lambda_0(c)\exp(\gamma_0'x)}\Big[\frac{(m-1)\exp(\alpha_0'x)}{\Lambda_0(c)\exp(\gamma_0'x)}\{\psi_i(t_j)-\phi_i(c)H_0(t_j)\\
&+x'(\sigma_i-\eta_i)H_0(t_j)\big\}-\big\{\phi_i(c)+x'\eta_i\big\}\Big]dP_4(x,c,m,t_1,\cdots,t_m)+o_p(1)
\end{aligned}\tag{8.7}$$

and

$$\begin{aligned}
&n^{-1/2}\sum_{i=1}^{n}\int_0^{\tau} q(t)\{X_i-\overline{x}(t)\}\Big[\frac{1}{\widehat{\Lambda}_0(T_i)\widehat{S}_i(t)}-\frac{1}{\Lambda_0(T_i)S_i(t)}\Big]\Delta_i(t)m_i d\mathcal{A}_0(t)\\
=&n^{-1/2}\sum_{i=1}^{n}\int_0^{\tau} q(t)\Big[\int\{x-\overline{x}(t)\}\exp\Big\{\frac{mH_0(t)\exp(\alpha_0'x)}{\Lambda_0(c)\exp(\gamma_0'x)}\Big\}\frac{mI(c\ge t)}{\Lambda_0(c)}\Big(\frac{m\exp(\alpha_0'x)}{\Lambda_0(c)\exp(\gamma_0'x)}\\
&\times\big\{\psi_i(t)-\phi_i(c)H_0(t)+x'(\sigma_i-\eta_i)H_0(t)\big\}-\phi_i(c)\Big)dP_2(x,c,m)\Big]d\mathcal{A}_0(t)+o_p(1),
\end{aligned}\tag{8.8}$$

where $\mathcal{A}_0(t)$ is defined in the first paragraph of Section 3. Thus, it follows from (8.6)–(8.8) that

$$n^{1/2}U(\beta_0)=n^{-1/2}\sum_{i=1}^{n}\xi_i+o_p(1),\tag{8.9}$$

where

$$\begin{aligned}
\xi_i=&\int_0^{\tau} q(t)\{X_i-\overline{x}(t)\}dM_i(t)+\int\sum_{j=1}^{m} q(t_j)\{x-\overline{x}(t_j)\}\{y(t_j)-\beta_0'x\}\exp\Big\{\frac{mH_0(t_j)\exp(\alpha_0'x)}{\Lambda_0(c)\exp(\gamma_0'x)}\Big\}\\
&\times\frac{mI(c\ge t_j)\exp(\alpha_0'x)}{\Lambda_0(c)\exp(\gamma_0'x)}\Big[\psi_i(t_j)-\phi_i(c)H_0(t_j)+x'(\sigma_i-\eta_i)H_0(t_j)\Big]dP_3(x,c,m,y,t_1,\cdots,t_m)\\
&-\int\sum_{j=1}^{m} q(t_j)\{x-\overline{x}(t_j)\}\exp\Big\{\frac{(m-1)H_0(t_j)\exp(\alpha_0'x)}{\Lambda_0(c)\exp(\gamma_0'x)}\Big\}\frac{(m-1)I(c\ge t_j)}{\Lambda_0(c)\exp(\gamma_0'x)}\Big[\frac{(m-1)\exp(\alpha_0'x)}{\Lambda_0(c)\exp(\gamma_0'x)}\\
&\times\big\{\psi_i(t_j)-\phi_i(c)H_0(t_j)+x'(\sigma_i-\eta_i)H_0(t_j)\big\}-\big\{\phi_i(c)+x'\eta_i\big\}\Big]dP_4(x,c,m,t_1,\cdots,t_m)\\
&-\int_0^{\tau} q(t)\Big[\int\{x-\overline{x}(t)\}\exp\Big\{\frac{mH_0(t)\exp(\alpha_0'x)}{\Lambda_0(c)\exp(\gamma_0'x)}\Big\}\frac{mI(c\ge t)}{\Lambda_0(c)}\Big(\frac{m\exp(\alpha_0'x)}{\Lambda_0(c)\exp(\gamma_0'x)}\\
&\times\big\{\psi_i(t)-\phi_i(c)H_0(t)+x'(\sigma_i-\eta_i)H_0(t)\big\}-\phi_i(c)\Big)dP_2(x,c,m)\Big]d\mathcal{A}_0(t).
\end{aligned}$$

Note that $-\partial U(\beta_0)/\partial\beta'$ converges in probability to A, which is defined in Condition (C4). The Taylor series expansion of $U(\widehat{\beta})$ at β_0 yields

$$n^{1/2}(\widehat{\beta}-\beta_0)=A^{-1}n^{1/2}U(\beta_0)+o_p(1).\tag{8.10}$$

It then follows from (8.9) and (8.10) that $n^{1/2}(\widehat{\beta}-\beta_0)$ is asymptotically zero-mean normal with covariance matrix $A^{-1}\Sigma A^{-1}$, which can be consistently estimated by $\widehat{A}^{-1}\widehat{\Sigma}\widehat{A}^{-1}$, where $\Sigma=E\{\xi_i^{\otimes 2}\}$ and $\widehat{A}^{-1}\widehat{\Sigma}\widehat{A}^{-1}$ is defined in the last paragraph of Section 3.

References

[1] Fan, J., Li, R. New estimation and model selection procedures for semiparametric modeling in longitudinal data analysis. *Journal of the American Statistical Association*, 99: 710–723 (2004)

[2] He, X., Tong, X., Sun, J. Semiparametric analysis of panel count data with correlated observation and follow-up times. *Lifetime Data Analysis*, 15: 177–196 (2009)

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

[4] Liang, Y., Lu, W., Ying, Z. Joint modeling and analysis of longitudinal data with informative observation times. *Biometrics*, 65: 377–384 (2009)

[5] Liang, K.Y., Zeger, S.L. Longitudinal data analysis using generalized linear models. *Biometrika*, 73: 13–22 (1986)

[6] Lin, D.Y., Wei, L.J., Ying, Z. Semiparametric transformation models for point processes. *Journal of the American Statistical Association*, 96: 620–628 (2001)

[7] Lin, D.Y., Wei, L.J., Ying, Z. Checking the Cox model with cumulative sums of martingale-based residuals. *Biometrika*, 85: 605–619 (1993)

[8] Lin, D.Y., Ying, Z. Semiparametric and nonparametric regression analysis of longitudinal data. *Journal of the American Statistical Association*, 96: 103–126 (2001)

[9] Lin, H., Scharfstein, D.O., Rosenheck, D.O. Analysis of longitudinal data with irregular outcome-dependent follow-up. *Journal of Royal Statistical Society B*, 66: 791–813 (2004)

[10] Liu, L., Huang, X., O'Quigley, J. Analysis of longitudinal data in the presence of informative observational times and a dependent terminal event, with application to medical cost data. *Biometrics*, 64: 950–958 (2008)

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

[12] Liu, L., Wolfe, R.A., Kalbfleisch, J.D. A shared random effects model for censored medical costs and mortality. *Statistics in Medicine*, 26: 139–155 (2007)

[13] Schoenfeld, D. Partial residuals for the proportional hazards regression model. *Biometrika*, 69: 239–241 (1982)

[14] Sun, J., Park, D.-H., Sun, L., Zhao, X. Semiparametric regression analysis of longitudinal data with informative observation times. *Journal of the American Statistical Association*, 100: 882–889 (2005)

[15] Sun, J., Sun, L., Liu, D. Regression analysis of longitudinal data in the presence of informative observation and censoring times. *Journal of the American Statistical Association*, 102: 1397–1406 (2007)

[16] Sun, L., Song, X., Zhou, J. Regression analysis of longitudinal data with time-dependent covariates in the presence of informative observation and censoring times. *Journal of Statistical Planning and Inference*, 141: 2902–2919 (2011)

[17] Sun, L., Song, X., Zhou, J., Liu, L. Joint analysis of longitudinal data with informative observation times and a dependent terminal event. *Journal of the American Statistical Association*, 107: 688–700 (2012)

[18] Wang, M.C., Qin, J., Chiang, C.T. Analyzing recurrent event data with informative censoring. *Journal of the American Statistical Association*, 96: 1057–1065 (2001)

[19] Zeng, D., Cai, J. A semiparametric additive rate model for recurrent events with an informative terminal event. *Biometrika*, 97: 699–712 (2010)