Regression Analysis of Longitudinal Data with Correlated Censoring and Observation Times

Yang Li[1], Xin He[2], Haiying Wang[3] and Jianguo Sun[4]

University of North Carolina-Charlotte[1]
University of Maryland[2]
University of New Hampshire[3]
University of Missouri[4]

Abstract

Longitudinal data occur in many fields such as the medical follow-up studies that involve repeated measurements. For their analysis, most existing approaches assume that the observation or follow-up times are independent of the response process either completely or given some covariates. In practice, it is apparent that this may not be true. In this paper, we present a joint analysis approach that allows the possible mutual correlations that can be characterized by time-dependent random effects. Estimating equations are developed for the parameter estimation and the resulted estimators are shown to be consistent and asymptotically normal. The finite sample performance of the proposed estimators is assessed through a simulation study and an illustrative example from a skin cancer study is provided.

Keywords: Estimating equation; Informative censoring; Informative observation process; Longitudinal data.

1 Introduction

Longitudinal data occur in many fields such as the medical follow-up studies that involve repeated measurements. In these situations, study subjects are generally observed only at discrete times. Therefore, for the analysis of longitudinal data, two processes need to be considered: one is the response process, which is usually of the primary interest but not continuously observable; the other one is the observation process, which is nuisance but gives rise to the discrete times when the responses are observed.

An extensive literature exists for the analysis of longitudinal data. Sun and Kalbfleisch (1995) and Wellner and Zhang (2000) investigated nonparametric estimation of the mean function when the response process is a counting process. Cheng and Wei (2000), Sun and Wei (2000), Zhang (2002) and Wellner and Zhang (2007) developed some semiparametric approaches for regression analysis under the proportional means models. However, with respect to the observation process, most existing approaches assume that the observation times are independent of the underlying response process either completely or given some covariates. For the analysis with a correlated observation process, there is limited work and most of them assume independent censoring or require some restrictive conditions such as the Poisson assumption or other distribution assumptions on the correlation structure (Huang et al. 2006; Sun et al. 2007; He et al. 2009; Zhao and Tong 2011; Li et al. 2013; Zhao et al. 2013b).

In many situations, however, the response process, the observation and censoring times may be mutually correlated. In addition, such correlations may be time-dependent. For instance, both the observation times and longitudinal responses may depend on the stage of disease progression. Their correlation may change over time and so are their correlations with the follow-up times. He et al. (2009) considered such correlations in shared frailty models. However, their method requires the assumptions that the underlying random effect is normally distributed and the observation process is a nonhomogeneous Poisson process. Also all correlations between the three processes are assumed to be fixed over time. Zhao et al. (2013b) proposed a robust estimation procedure and relaxed the Poisson assumption required in He et al. (2009). However, the follow-up times are assumed to be independent from covariates, responses and observation times; and the possible correlations between responses and observation times are time-independent. More recently, Sun et al. (2012) presented a joint model with time-dependent correlations between the response process, the observation times and a terminal event, where the random effect associated with the terminal event is fixed over time and follow a specified distribution. In practice, however, such conditions may not hold or be difficult to check when informative censoring involves.

In this paper, we consider regression analysis of longitudinal data when the underlying response process, the observation and censoring times are mutually correlated and none of the correlations is restricted by specified forms or distributions. A general estimation approach is proposed. The remainder of this paper is organized as follows: In Section 2, we introduce the notation and present the model. Sections 3 presents the estimation procedure and establishes the asymptotic properties of the resulted estimators, and Section 4 discusses a simplified estimation procedure for a special case. In Section 5, a simulation study is performed to evaluate the finite sample properties of the proposed estimators and an illustrative example is provided in Section 6. Some concluding remarks are given in Section 7.

2 Notation and Models

Consider a longitudinal study in which the response process of interest is observed only at some discrete sampling time points. For each subject i, $i = 1, \cdots, n$, let $N_i(t)$ be the observation process, which gives the cumulative number of observation times up to time t. In practice, one observes $\widetilde{N}_i(t) = N_i(t \wedge C_i)$ where $a \wedge b = min(a, b)$ and C_i denotes the censoring or follow-up time. Let $Y_i(t)$ denote the response process, which is observed only at discrete times $\{T_{i,1}, \cdots, T_{i,m_i}\}$ when $N_i(t)$ has jumps. Suppose that there exists a p-dimensional vector of covariates denoted by $\mathbf{Z}_i$, which will be assumed to be time-independent.

In the following, we model the correlation between $Y_i(t)$, $N_i(t)$ and C_i through an unobserved random vector $\mathbf{b}_i(t) = (b_{1i}(t), b_{2i}(t), b_{3i}(t))'$, which could be time-dependent. Define $\mathcal{B}_{it} = \{\mathbf{b}_i(s), s \leq t\}$. It will be assumed that the $\mathbf{b}_i(t)$'s are independent and identically distributed, $\mathcal{B}_{it}$ is independent of $\mathbf{Z}_i$, and given $\mathbf{Z}_i$ and $\mathcal{B}_{it}$, C_i, $N_i(t)$ and $Y_i(t)$ are mutually independent. To be specific, the mean function of $Y_i(t)$ has the form

$$E\{Y_i(t)|\mathbf{Z}_i, \mathbf{b}_i(t)\} = \Lambda_0(t) \exp\{\beta'\mathbf{Z}_i + b_{1i}(t)\}, \tag{1}$$

where $\Lambda_0(t)$ is an unknown baseline mean function and β denotes a vector of p-dimensional regression coefficients. Also the observation process $N_i(t)$ follows the proportional rates model

$$E\{dN_i(t)|\mathbf{Z}_i, \mathbf{b}_i(t)\} = \exp\{\gamma'\mathbf{Z}_i + b_{2i}(t)\}d\mu_0(t)\,, \tag{2}$$

where γ is a vector of unknown parameters and $d\mu_0(t)$ is an unknown baseline rate function. For the $C_i's$, motivated by the additive hazards models that have been commonly used in survival analysis (Lin and Ying, 2001; Kalbfleisch and Prentice, 2002; Zhang et al. 2005), we consider the following model

$$\lambda_i(t|\mathbf{Z}_i, \mathbf{b}_i(t)) = \lambda_0(t) + \xi'\mathbf{Z}_i + b_{3i}(t)\,. \tag{3}$$

Here $\lambda_0(t)$ is an unknown baseline hazard function and ξ denotes the effect of covariates on the hazard function of $C_i's$. Note that instead of model (3), one may consider the proportional hazards model. As pointed out by Lin et al. (1998) and others, the additive model (3) can be more plausible than the proportional hazards model in many applications. More comments on this are given in Section 7.

In the above, models (1) - (3) can be viewed as natural generalizations of some existing and commonly used models. For example, when $b_{1i}(t) = 0$, model (1) is equivalent to the proportional means models considered by Cheng and Wei (2000), Sun and Wei (2000), Zhang (2002) and Hu et al. (2003) among others. When $b_{1i}(t)$ is time-independent, model (1) is equivalent to model (3) considered in Zhao et al. (2013b). In fact, when any of the $b_{ki}(t)$'s ($k = 1, 2, 3$) is zero or independent from other $b_{ji}(t)$'s ($j = 1, 2, 3$ and $j \neq k$), the corresponding process is independent from the others. Therefore, the proposed joint model also applies to special cases when either the observation or censoring times are noninformative. In general, since the form or distribution of $\mathbf{b}_i(t)$ is arbitrary and completely unspecified, the joint model described above is quite flexible compared to many existing procedures.

Note that in models (1) - (3), for simplicity, we have assumed that the set of covariates that may affect $Y_i(t)$, $N_i(t)$ and C_i is the same. In practice, it is apparent that this may not be the case and actually the estimation procedure proposed below still applies as long as one replaces $\mathbf{Z}_i$ by appropriate covariates. As an alternative, one can define a single and big covariate vector by combining all different covariates together. In the following, we will focus on estimation of regression parameters β along with γ and ξ. For this, it is easy to see that the use of the existing procedures that assume independence could give biased or even misleading results.

3 General Estimation Procedure (GEP)

In this section, we will present an inference procedure for estimation of β which is usually of the primary interest. For this, first note that the counting process $\widetilde{N}_i(t) = N_i(t \wedge C_i)$ jumps by one at time t if and only if $C_i \geq t$ and $dN_i(t) = 1$. Also we have

$$
\begin{aligned}
& E\{d\widetilde{N}_i(t)|\mathbf{Z}_i\} = E\{I(t \le C_i)dN_i(t)|\mathbf{Z}_i\} \\
= \quad & E\left[E\{I(t \le C_i)dN_i(t)|\mathbf{Z}_i, \mathcal{B}_{it}\}\bigg|\mathbf{Z}_i\right] \\
= \quad & E\left[E\{I(t \le C_i)|\mathbf{Z}_i, \mathcal{B}_{it}\}E\{dN_i(t)|\mathbf{Z}_i, \mathcal{B}_{it}\}\bigg|\mathbf{Z}_i\right] \\
= \quad & E\left[exp\{-\Lambda_0^*(t) - B_i(t) - \xi'\mathbf{Z}_i t\}\exp\{\gamma'\mathbf{Z}_i + b_{2i}(t)\}d\mu_0(t)\bigg|\mathbf{Z}_i\right] \\
= \quad & \exp\{\gamma'\mathbf{Z}_i - \xi'\mathbf{Z}_i t\}d\Lambda_1^*(t), \qquad (4)
\end{aligned}
$$

where

$$
\Lambda_0^*(t) = \int_0^t \lambda_0(s)ds, \quad B_i(t) = \int_0^t b_{3i}(s)ds
$$

and

$$
d\Lambda_1^*(t) = \exp\{-\Lambda_0^*(t)\}E[exp\{b_{2i}(t) - B_i(t)\}]d\mu_0(t).
$$

Define

$$
dM_i^*(t;\eta) = d\widetilde{N}_i(t) - e^{\eta'\mathbf{X}_i(t)}d\Lambda_1^*(t)
$$

and $dM_i^*(t) = dM_i^*(t;\eta_0)$, where $\eta = (\gamma, \quad \xi)'$, $\mathbf{X}_i(t) = (\mathbf{Z}_i, \quad -\mathbf{Z}_i t)'$ and η_0 denotes the true value of η. It can be shown that $M_i^*(t)$ is a mean-zero stochastic process. It follows that the estimators of η and $d\Lambda_1^*(t)$ can be obtained by solving the following two estimating equations

$$
U_\eta(\eta) = \sum_{i=1}^n \int_0^\tau \left\{\mathbf{X}_i(t) - \bar{X}(t;\eta)\right\}d\widetilde{N}_i(t) = 0 \qquad (5)
$$

and

$$
\sum_{i=1}^n \left[d\widetilde{N}_i(t) - e^{\eta'\mathbf{X}_i(t)}d\Lambda_1^*(t)\right] = 0. \qquad (6)
$$

In the above, τ is the longest follow-up time, $\bar{X}(t;\eta) = S^{(1)}(t;\eta)/S^{(0)}(t;\eta)$ and $S^{(k)}(t;\eta) = n^{-1}\sum_{i=1}^n e^{\eta'\mathbf{X}_i(t)}\mathbf{X}_i(t)^{\otimes k}$ with $a^{\otimes 0} = 1$, $a^{\otimes 1} = a$, $\bar{x}(t) = lim_{n\to\infty}\bar{X}(t;\eta_0)$ and $s^{(k)}(t) = lim_{n\to\infty}S^{(k)}(t;\eta_0),\ k = 0, 1$.

To estimate β, consider

$$
\begin{aligned}
& E\{Y_i(t)d\widetilde{N}_i(t)|\mathbf{Z}_i, \mathcal{B}_{it}\} \\
= \quad & E\{I(t \le C_i)Y_i(t)dN_i(t)|\mathbf{Z}_i, \mathcal{B}_{it}\} \\
= \quad & E\{I(t \le C_i)|\mathbf{Z}_i, \mathcal{B}_{it}\}E\{Y_i(t)|\mathbf{Z}_i, \mathcal{B}_{it}\}E\{dN_i(t)|\mathbf{Z}_i, \mathcal{B}_{it}\} \\
= \quad & exp\{-\Lambda_0^*(t) - B_i(t) - \xi'\mathbf{Z}_i t\} \\
& \Lambda_0(t)\exp\{\beta'\mathbf{Z}_i + b_{1i}(t)\}\exp\{\gamma'\mathbf{Z}_i + b_{2i}(t)\}d\mu_0(t) \\
= \quad & \exp\{(\beta+\gamma)'\mathbf{Z}_i - \xi'\mathbf{Z}_i t\} \\
& \exp\{-\Lambda_0^*(t) + b_{1i}(t) + b_{2i}(t) - B_i(t)\}\Lambda_0(t)d\mu_0(t),
\end{aligned}
$$

and therefore

$$E\{Y_i(t)d\widetilde{N}_i(t)|\mathbf{Z}_i\} = \exp\{\beta'\mathbf{Z}_i + \eta'\mathbf{X}_i(t)\}d\Lambda_2^*(t), \tag{7}$$

where

$$d\Lambda_2^*(t) = \exp\{-\Lambda_0^*(t)\}\Lambda_0(t)E[exp\{b_{1i}(t) + b_{2i}(t) - B_i(t)\}]d\mu_0(t).$$

Define

$$dM_i(t;\beta,\eta) = Y_i(t)d\widetilde{N}_i(t) - \exp\{\beta'\mathbf{Z}_i + \eta'\mathbf{X}_i(t)\}d\Lambda_2^*(t)$$

and $dM_i(t) = dM_i(t;\beta_0,\eta_0)$, where β_0 denotes the true value of β. Then $M_i(t)$ is a mean-zero stochastic process. This naturally suggests the following estimating equations to estimate β and $d\Lambda_2^*(t)$:

$$U_\beta(\beta;\hat{\eta}) = \sum_{i=1}^n \int_0^\tau W(t)\mathbf{Z}_i\Big[Y_i(t)d\widetilde{N}_i(t) - e^{\beta'\mathbf{Z}_i+\hat{\eta}'\mathbf{X}_i(t)}d\Lambda_2^*(t)\Big] = 0, \tag{8}$$

and

$$\sum_{i=1}^n \Big[Y_i(t)d\widetilde{N}_i(t) - e^{\beta'\mathbf{Z}_i+\hat{\eta}'\mathbf{X}_i(t)}d\Lambda_2^*(t)\Big] = 0,\ \ 0 \le t \le \tau, \tag{9}$$

where $\hat{\eta} = (\hat{\gamma},\ \hat{\xi})'$ and $d\widehat{\Lambda}_1^*(t)$ are the estimators of η and $d\Lambda_1^*(t)$, respectively, solved from (5) and (6), and $W(t)$ is a possibly data-dependent weight function. We denote the estimates of β and $d\Lambda_2^*(t)$ by $\hat{\beta}$ and $d\widehat{\Lambda}_2^*(t)$, respectively, solved from (8) and (9).

To establish the asymptotic properties of $\hat{\beta}$ and $\hat{\eta}$, define

$$\widehat{M}_i^*(t) = \widetilde{N}_i(t) - \int_0^t e^{\hat{\eta}'\mathbf{X}_i(s)}d\widehat{\Lambda}_1^*(s;\hat{\eta}),$$

$$\widehat{M}_i(t) = \int_0^t Y_i(s)d\widetilde{N}_i(s) - \int_0^t e^{\hat{\beta}'\mathbf{Z}_i+\hat{\eta}'\mathbf{X}_i(s)}d\widehat{\Lambda}_2^*(s;\hat{\beta},\hat{\eta}),$$

$$\widehat{E}_Z(t;\beta,\eta) = \frac{\sum_{i=1}^n \mathbf{Z}_i e^{\beta'\mathbf{Z}_i+\eta'\mathbf{X}_i(t)}}{\sum_{i=1}^n e^{\beta'\mathbf{Z}_i+\eta'\mathbf{X}_i(t)}} \text{ and } e_z(t) = lim_{n\to\infty}\widehat{E}_Z(t;\beta_0,\eta_0).$$

The following theorem gives the consistency and asymptotic normality of $\hat{\beta}$ and $\hat{\eta}$.

Theorem 1. Assume that the conditions (C1)-(C5) given in the Appendix hold. Then $\hat{\eta}$ and $\hat{\beta}$ are consistent estimators of η_0 and β_0, respectively. The distributions of $n^{1/2}(\hat{\eta}-\eta_0)$ and $n^{1/2}(\hat{\beta}-\beta_0)$ can be asymptotically approximated by the normal distributions with mean zero and covariance matrices $\widehat{\Sigma}_\eta = \widehat{\Omega}_\eta^{-1}\widehat{\Psi}\widehat{\Omega}_\eta^{-1}$ and $\widehat{\Sigma}_\beta = \widehat{A}_\beta^{-1}\widehat{\Sigma}\widehat{A}_\beta^{-1}$, respectively, where $a^{\otimes 2} = aa'$, $\widehat{\Psi} = n^{-1}\sum_{i=1}^n \hat{u}_i^{\otimes 2}$, $\widehat{\Sigma} = n^{-1}\sum_{i=1}^n(\hat{v}_{1i} - \hat{v}_{2i})^{\otimes 2}$,

$$\hat{u}_i = \int_0^\tau \Big(\mathbf{X}_i(t) - \bar{X}(t;\hat{\eta})\Big)d\widehat{M}_i^*(t)\,,$$

$$\hat{v}_{1i} = \int_0^\tau W(t)\Big(\mathbf{Z}_i - \widehat{E}_Z(t;\hat{\beta},\hat{\eta})\Big)d\widehat{M}_i(t)\,,$$

$$\hat{v}_{2i} = \int_0^\tau \widehat{A}_\eta\widehat{\Omega}_\eta^{-1}\Big(\mathbf{X}_i(t) - \bar{X}(t;\hat{\eta})\Big)d\widehat{M}_i^*(t)\,,$$

$$\widehat{A}_\beta = n^{-1}\sum_{i=1}^n \int_0^\tau W(t)e^{\hat{\beta}'\mathbf{Z}_i+\hat{\eta}'\mathbf{X}_i(t)}\Big(\mathbf{Z}_i - \widehat{E}_Z(t;\hat{\beta},\hat{\eta})\Big)^{\otimes 2}d\widehat{\Lambda}_2^*(t;\hat{\beta},\hat{\eta}),$$

$$\widehat{A}_\eta = n^{-1}\sum_{i=1}^{n}\int_0^\tau W(t)e^{\hat\beta'\mathbf{Z}_i+\hat\eta'\mathbf{X}_i(t)}\Big(\mathbf{Z}_i - \widehat{E}_Z(t;\hat\beta,\hat\eta)\Big)X_i'(t)d\widehat{\Lambda}_2^*(t;\hat\beta,\hat\eta)$$

and

$$\widehat{\Omega}_\eta = n^{-1}\sum_{i=1}^{n}\int_0^\tau \{\mathbf{X}_i(t) - \bar{X}(t;\hat\eta)\}^{\otimes 2}e^{\hat\eta'\mathbf{X}_i(t)}d\widehat{\Lambda}_1^*(t;\hat\eta).$$

For the implementation of the estimation procedure described above, one question of interest is the model-checking on models (1) - (3). Note that for both models (2) and (3), one observes complete data and there exist several procedures to check their goodness-of-fit (Schoenfeld, 1982; Lin et al. 1993; Lin et al. 2000; Ghosh and Lin, 2002). Thus here we will focus on model (1). For this, by following Lin et al. (1993, 2000), a general approach is to employ the following supremum statistic

$$\mathcal{F}(t,z) = n^{-1/2}\sum_{i=1}^{n}\int_0^t I(\mathbf{Z}_i \le z)d\widehat{M}_i(s)\,,$$

where the event $\{\mathbf{Z}_i \le z\}$ means that each component of $\mathbf{Z}_i$ is not larger than the corresponding component of z. In Appendix 2, we will show that the null distribution of $\mathcal{F}(t,z)$ converges weakly to a mean-zero Gaussian process that can be approximated by

$$\widehat{\mathcal{F}}(t,z) = n^{-1/2}\sum_{i=1}^{n}\Big\{\hat u_{1i}(t,z) - \widehat{\Phi}_\eta(t,z)\widehat{\Omega}_\eta^{-1}\hat u_{2i} - \widehat{\Phi}_\beta(t,z)\widehat{A}_\beta^{-1}(\hat v_{1i} - \hat v_{2i})\Big\}e_i\,. \qquad (10)$$

Here $e_1,\ldots,e_n$ are independent standard normal variables independent of the observed data,

$$\hat u_{1i}(t,z) = \int_0^t \{I(\mathbf{Z}_i \le z) - \widehat{E}_I(s,z;\hat\beta,\hat\eta)\}d\widehat{M}_i(s),$$

$$\widehat{\Phi}_\eta(t,z) = n^{-1}\sum_{i=1}^{n}\int_0^t \{I(\mathbf{Z}_i \le z) - \widehat{E}_I(s,z;\hat\beta,\hat\eta)\}e^{\hat\beta'\mathbf{Z}_i+\hat\eta'\mathbf{X}_i(s)}\mathbf{X}_i'(s)d\widehat{\Lambda}_2^*(s;\hat\beta,\hat\eta),$$

$$\widehat{\Phi}_\beta(t,z) = n^{-1}\sum_{i=1}^{n}\int_0^t \{I(\mathbf{Z}_i \le z) - \widehat{E}_I(s,z;\hat\beta,\hat\eta)\}e^{\hat\beta'\mathbf{Z}_i+\hat\eta'\mathbf{X}_i(s)}\mathbf{Z}_i'\,d\widehat{\Lambda}_2^*(s;\hat\beta,\hat\eta),$$

$$\widehat{E}_I(t,z;\beta,\eta) = \frac{\sum_{i=1}^{n} I(\mathbf{Z}_i \le z)e^{\beta'\mathbf{Z}_i+\eta'\mathbf{X}_i(t)}}{\sum_{i=1}^{n} e^{\beta'\mathbf{Z}_i+\eta'\mathbf{X}_i(t)}},\ e_I(t,z) = lim_{n\to\infty}\widehat{E}_I(t,z;\beta_0,\eta_0)$$

and $\hat u_{2i} = \hat u_i$, where $\hat u_i$, $\hat v_{1i}$ and $\hat v_{2i}$ are as defined earlier in this section. Therefore, one could obtain a large number of realizations from $\widehat{\mathcal{F}}(t,z)$ by repeatedly generating the standard normal random sample $\{e_1,\ldots,e_N\}$ while fixing the observation data. Because $\mathcal{F}(t,z)$ is expected to fluctuate randomly around 0 under model (1), a formal lack-of-fit test can be constructed based on the statistic $sup_{0\le t\le\tau,z}|\mathcal{F}(t,z)|$. The corresponding p-value can be obtained by comparing the observed value of $sup_{0\le t\le\tau,z}|\mathcal{F}(t,z)|$ to a large number of realizations from $sup_{0\le t\le\tau,z}|\widehat{\mathcal{F}}(t,z)|$.

4 Simplified Estimation Procedure (SEP)

In this section, we consider a simplification of the GEP presented in Section 3 for a special case, when C_i is independent of the covariate $\mathbf{Z}_i$. In such situation, for the same models considered in Section 2, $\xi = 0$. Following (4) and (7), we have

$$E\{d\widetilde{N}_i(t)|\mathbf{Z}_i\} = \exp\{\gamma'\mathbf{Z}_i\}d\Lambda_1^*(t), \tag{11}$$

and

$$E\{Y_i(t)d\widetilde{N}_i(t)|\mathbf{Z}_i\} = \exp\{(\beta+\gamma)'\mathbf{Z}_i\}d\Lambda_2^*(t), \tag{12}$$

where $\Lambda_0^*(t)$, $B_i(t)$, $d\Lambda_1^*(t)$ and $d\Lambda_2^*(t)$ are defined the same as in the previous section. Let

$$m_i = \int_0^\tau d\widetilde{N}_i(t)$$

be the total number of observation times associated with subject i and

$$\bar{Y}_i = \int_0^\tau Y_i(t)d\widetilde{N}_i(t),$$

where τ represents the largest follow-up time in the study. Then we have

$$E\{m_i|\mathbf{Z}_i\} = \exp\{\gamma'\mathbf{Z}_i\}\Lambda_1^*(\tau)$$

and

$$E\{\bar{Y}_i|\mathbf{Z}_i\} = \exp\{(\beta+\gamma)'\mathbf{Z}_i\}\Lambda_2^*(\tau) = E(m_i|\mathbf{Z}_i)\exp\{\beta'\mathbf{Z}_i + \alpha\},$$

where $\alpha = \log\{\Lambda_2^*(\tau)/\Lambda_1^*(\tau)\}$ is an unknown parameter.

Define $\mathbf{Z}_{1i} = (\mathbf{Z}_i',\ 1)'$ and $\phi = (\beta',\ \alpha')'$. Motivated by Zhao et al. (2013b), the following class of estimating equations can be used for the estimation of ϕ

$$U(\phi) = \sum_{i=1}^n W_i\mathbf{Z}_{1i}\{\bar{Y}_i - m_i\exp(\phi'\mathbf{Z}_{1i})\} = 0,$$

where W_i's are some possibly covariate-dependent weights. Under conditions (C1)-(C4) given in Appendix 1, it can be shown that the resulted estimator $\hat{\phi}$ is consistent and $\sqrt{n}(\hat{\phi} - \phi_0)$ can be asymptotically approximated by a normal distribution with mean zero and covariance matrix $\widehat{\Sigma}_\phi = \widehat{\Gamma}^{-1}\widehat{V}\widehat{\Gamma}^{-1}$, where ϕ_0 denotes the true value of ϕ,

$$\widehat{\Gamma} = \frac{1}{n}\sum_{i=1}^n\{W_i m_i \mathbf{Z}_{1i}^{\otimes 2}\exp(\hat{\phi}'\mathbf{Z}_{1i})\}$$

and

$$\widehat{V} = \frac{1}{n}\sum_{i=1}^n \hat{v}_i^{\otimes 2} \text{ with } \hat{v}_i = W_i\mathbf{Z}_{1i}\{\bar{Y}_i - m_i\exp(\hat{\phi}'\mathbf{Z}_{1i})\}.$$

In practice, one question of interest is when one may want to apply the SEP. For this, it is apparent that a simple way is first to directly check on the observed data to see if the censoring times are covariate-dependent. Otherwise, one only needs to fit the data to model (3) without $b_{3i}(t)$ or employ some nonparametric test procedures such as the log-rank test.

5 A Simulation Study

In this section, we report some results obtained from a simulation study conducted to assess the finite sample behavior of the estimation procedure proposed in the previous sections. For each subject i, the covariate $\mathbf{Z}_i$ was assumed to be a Bernoulli random variable with the probability of success being 0.5. Given $\mathbf{Z}_i$ and some unobserved random effects $\mathbf{b}_i(t) = (b_{1i}(t), b_{2i}(t), b_{3i}(t))'$, the hazard function of the censoring time C_i was assumed to have the form

$$\lambda_i(t|\mathbf{Z}_i, \mathcal{B}_{it}) = \lambda_0 + \xi \mathbf{Z}_i + b_{3i}(t), \tag{13}$$

with the largest follow-up time $\tau = 1$. The number of observations $\widetilde{N}_i(t)$ was assumed to follow a Poisson process on $(0, C_i)$ with the mean function

$$E\{N_i(t)|\mathbf{Z}_i, \mathcal{B}_{it}\} = \int_0^t \exp\{\gamma \mathbf{Z}_i + b_{2i}(s)\} d\mu_0(s)\,. \tag{14}$$

In practice, the exact time of C_i may not be observable and $d\widetilde{N}_i(t)$ is observed instead of $dN_i(t)$, thus we considered $E\{\widetilde{N}_i(t)|\mathcal{B}_{it}\}$ for the observation process. From (13) and (14),

$$E\{d\widetilde{N}_i(t)|\mathbf{Z}_i, \mathcal{B}_{it}\} = \exp\{\gamma \mathbf{Z}_i - \xi \mathbf{Z}_i t\} d\Lambda_1^*(t),$$

where $d\Lambda_1^*(t) = \exp\{-\lambda_0 t + b_{2i}(t) - B_i(t)\} d\mu_0(t)$ and $B_i(t) = \int_0^t b_{3i}(s) ds$. Given $\mathbf{Z}_i$ and $\mathcal{B}_{it}$, $\widetilde{N}_i(t)$ was assumed to follow a nonhomogeneous Poisson process and the total number of observation times m_i was generated with mean $E\{m_i\} = E\{\widetilde{N}_i(\tau)|Z_i, \mathcal{B}_{i\tau}\}$. Then the observation times $\{T_{i,1}, \ldots, T_{i,m_i}\}$ were taken as m_i order statistics from the density function

$$f_{\widetilde{N}}(t) = \frac{\exp\{\gamma \mathbf{Z}_i - \xi \mathbf{Z}_i t\} d\Lambda_1^*(t)}{\int_0^\tau \exp\{\gamma \mathbf{Z}_i - \xi \mathbf{Z}_i t\} d\Lambda_1^*(t)}.$$

The longitudinal response $Y_i(t)$ was generated from a mixed Poisson process with the mean function

$$E\{Y_i(t)|\mathbf{Z}_i, \mathcal{B}_{it}\} = Q_i \Lambda_0(t) \exp\{-\beta \mathbf{Z}_i + b_{1i}(t)\}, \tag{15}$$

where Q_i was generated independently from a gamma distribution with mean 1 and variance 0.5. The results given below are based on the sample size of 100 or 200 with 1000 replications and $W(t) = W_i = 1$.

Tables 1 and 2 show the estimation results on β for the situation when b_{1i}, b_{2i} and b_{3i} are time-independent. Note that here $\xi_0 = 0$ or $\gamma_0 = 0$ represents the cases when either censoring or the observation times is independent of covariates, respectively. For the random effects, we took $b_{1i} = b_{2i} = b_{3i} = b_i$, where the $b_i's$ were generated from the uniform distribution over $(-0.5, 0.5)$. It can be seen that the proposed estimates seem unbiased and the estimated standard errors (SEE) are close to the sample standard errors (SSE). Also the empirical 95% coverage probabilities (CP) are quite accurate. In addition, it seems that both estimation procedures perform comparably well, but SEE agrees slightly better with the corresponding SSE under SEP as compared to GEP when $\xi_0 = 0$. The same conclusions are also obtained for the situation when b_{1i}, b_{2i} and b_{3i} are time-dependent, for which the results are presented in Tables 3 and 4. Here we took $b_{1i}(t) = b_i\, t^{1/3}$, $b_{2i}(t) = b_i\, t^{1/2}$ and $b_{3i} = b_i$ with the same b_i generated as for Tables 1 and 2.

Note that all results above were obtained with a binary covariate and are on estimation of β. We also considered other set-ups and estimation of other regression parameters and obtained similar results. For example, Table 5 presents the estimation results on $\eta = (\gamma, \xi)'$ given by GEP with the same set-ups as in Tables 2 and 4 and Table 6 gives results with $\mathbf{Z}_i$ generated from the normal distribution with mean 0 and standard deviation 0.5. One can see that all results are similar to those described above.

To further investigate the performance of the proposed estimators of β in comparison with those proposed by He et al. (2009) and Sun et al. (2012), we carried out a simulation study and estimated β using all four methods. Note that unlike the proposed estimation procedures, the latter two methods require observing the exact time of a censoring or terminal event C_i. For this, we used the subjects' last observation times as commonly done in practice. With respect to the method given by Sun et al. (2012), we applied it by using C_i as its original terminal event time D_i and τ as its C_i. Note that as mentioned earlier, both He et al. (2009) and Sun et al. (2012) considered the distribution-based random effects for possible correlations. For the comparison, we focus on the performances of their procedures when the random effects follow various distributions besides those assumed. However, since both of them involve covariate effects in forms different from those considered by our proposed models, we fix $\beta_0 = 0$ and $\xi_0 = 0$ in order to avoid unfair comparisons caused by the misspecification of covariate effects. The estimation results are given in Table 7 with three set-ups. In the first set-up, referred to as M_1, we considered the situation as used for Table 1 except $\mu_0(t) = 10t$ and $b_{1i} = -b_{2i} = b_{3i}$. In the second and third set-ups called M_2 and M_3, we generated $b_{1i}(t)$, $b_{2i}(t)$ and $b_{3i}(t)$ from various distributions such that the assumptions required by either Sun et al. (2012) or He et al. (2009) are satisfied. For example, we took $\lambda_0(t) = 0$ and generated $b_{3i}(t)$ from an extreme-value distribution as assumed by Sun et al. (2012). We also generated $b_{1i}(t)$, $b_{2i}(t)$ and $b_{3i}(t)$ from the assumed distributions required by He et al. (2009).

Note that in all set-ups considered above, our proposed models are correctly specified because there are no assumed distributions on $b_{1i}(t)$, $b_{2i}(t)$ or $b_{3i}(t)$. In contrast, the models from either of He et al. (2009) or Sun et al. (2012) are only correctly specified in one of the set-ups. On the other hand, since there are no covariate effects in all set-ups, we do not expect that the point estimates of β given by He et al. (2009) or Sun et al. (2012) are much biased even if the imposed distributions are misspecifyed in the estimation. For their variance estimates, we expect that SEE and SSE agree for both, because the former applied bootstrap resampling and the latter did not involve any assumed distribution of random effects in their variance estimation. Therefore, we only compare bias and SSE. It can be seen that all estimation procedures gave comparably small bias as expected. However, it appears that the proposed estimators are more efficient for all cases in general. In comparison, the method given by He et al. (2009) is comparably efficient to the proposed estimators only under M_3 when all its distribution assumptions are satisfied. For the method given by Sun et al. (2012), it is worth noting that when D_i is substituted by the last observation time C_i from subject i, it gives relatively large SSE, especially when C_i's vary much, regardless of whether the assumption about $b_{3i}(t)$ is satisfied (for M_2) or not (for M_3).

6 An Application

In this section, we applied the proposed methodology described in the previous sections to the data from a skin cancer study conducted by the University of Wisconsin Comprehensive Cancer Center in Madison, Wisconsin (Li et al. 2013; Sun and Zhao 2013). It is a double-blinded and placebo-controlled randomized Phase III clinical trial on the patients with a history of nonmelanoma skin cancers. The study consists of 291 patients randomly assigned to the placebo or 0.5g/m^2/day PO diuoromethylornithine (DFMO) and all subjects were scheduled to be assessed every six months, but the actual observation times differ from patient to patient. Thus only longitudinal data are available on the recurrences of skin cancers, and one objective is to evaluate the overall effectiveness of DFMO in reducing the recurrences of basal cell carcinoma (BCC). Following Li et al. (2013), we will focus on the 290 patients with at least one observation. Among these patients, 143 of them were assigned to the DFMO group and the rest were assigned to the placebo group. For each patient, the number of prior skin cancer occurrences was also reported at the beginning of the study, ranging from 1 to 35 and the median was 2. With respect to the new occurrences of BCC, the numbers ranged from 0 to 16. It was found by Zhang et al. (2013) that the cumulative numbers of BCC and observation numbers were positively correlated with varying magnitude over time. However, none of the existing literature has incorporated such information to analyze the effectiveness of DFMO.

For the analysis, we define $\mathbf{Z}_i = (Z_{i1},\ Z_{i2})'$, where $Z_{i1} = 1$ if the patient was given the DFMO treatment and $Z_{i1} = 0$ otherwise, and $Z_{i2} = 1$ if the patient had more than two skin tumors and $Z_{i2} = 0$ if not. The longest follow-up time was scaled to be $\tau = 1$, which corresponds to 1879 days in the original data set. For patient i, let $Y_i(t)$ be the total number of BCC tumors observed up to time t. The follow-up times C_i's were taken as the subjects' last observation times and vary from 180 to 1879 days. Assume that the recurrence process of skin tumors $Y_i(t)$, the observation process $N_i(t)$ and the follow-up time C_i can be described by models (1)-(3), respectively. A direct check of the observed data indicates that most early drop-outs ($C_i \leq 700$) occur in the placebo group and thus it seems that the censoring is covariate-dependent. This suggest that we should use the GEP described in Section 3.

Table 8 presents the analysis results given by the GEP with $W(t) = 1$ and for comparison, we also obtained and included in the table the analysis results by applying the methods given by He et al. (2009) and Sun et al. (2012). First one can see from the table that both ξ_1 and ξ_2 seem to be significant based on the GEP, indicating that the censoring times are indeed dependent of covariates. All three methods imply that the number of prior skin tumors could increase the recurrence rate of new BCC tumors since β_2 appears to be significantly positive. With respect to the effectiveness of DFMO treatment, however, the GEP suggests that DFMO may have significantly reduced the recurrence rate of new BCC tumors which is similar to the result given by Li et al. (2014), but neither of the competing procedures agree. One possible reason for this is that as mentioned before, both He et al. (2009) and Sun et al. (2012) considered different models for the association of covariate effects and the treatment effect can be insignificant from a different aspect of view. Another possible reason is that the assumed distributions of random effects were misspecified in both He et al. (2009) and Sun et al. (2012), and as shown in the simulation study, such misspecification effects may affect the variance estimates and mask the true significance. For the analysis, we also applied the model-checking procedure described in Section 3 and obtained the p-value of 0.632, which suggests that model (1) seems to be appropriate for the skin cancer data.

7 Concluding Remarks

In this paper, we proposed a joint model for analyzing longitudinal data with informative censoring and observation times. The mutual correlations are characterized via a shared vector of time-dependent random effects. As mentioned earlier, several procedures have been developed in the literature for longitudinal data when either censoring or observation process is informative.

However when both of them are informative, there is limited work that can apply except those given in He et al. (2009) and Sun et al. (2012). In addition, all the existing procedures assumed time-independent or specifically distributed correlation structures. The proposed joint model is flexible in that the shared vector of random effects can be time-dependent and neither of its structure nor distribution are specified. For the parameter estimation, the proposed procedure is simple and easy to implement. It can be further simplified under special situations when the censoring does not depend on covariates.

Note that as mentioned above, instead of the additive model (3), one may consider the frailty proportional hazards model for the censoring times. It is well-known that both models are commonly used in survival analysis and they describe different types of covariate effects. In other words, one may want to decide which model to be used based on the question of interest. Another factor for this is that one may also need to see if there exist some established inference procedures available. As one can see, under model (3), we have developed a simple procedure for the estimation of regression parameters. Although one can apply the same idea for the proportional hazards model situation, the specific development and implementation of similar procedures would be quite different and difficult. The same can be said on the modeling of $Y_i(t)$ (Zhao et al. 2011; Zhao et al. 2013a). Another assumption used in the proposed method is the existence of time-dependent random effects $\mathbf{b}_i(t)$, a commonly used technique in the analysis of longitudinal data to characterize the correlation between some related variables or processes. In most methods involving random effects, they are usually supposed to be time-independent, meaning that the correlation is constant. On the other hand, it is apparent that the correlation may change with time and time-dependent random effects include time-independent random effects as a special case.

There exist several directions for future research. One is that as mentioned above, one may want to consider other models rather than models (1) - (3) and develop similar estimation procedures. Of course, a related problem is model selection and one may want to develop some model selection techniques to choose the optimal model among several candidate models (Tong et al. 2009; Wang et al. 2014). Note that in the proposed method, we have employed a weight function $W(t)$ and it would be desirable to develop some procedures for the selection of an optimal $W(t)$. As in most similar situations, this is clearly a difficult problem as it requires the specification of the covariance function of $Y_i(t)$ and $\widetilde{N}_i(t)$ (Sun et al. 2012). Finally in the above, we have focused on regression analysis of $Y_i(t)$ with time-independent covariates. Sometimes one may face time-dependent covariates and thus it would be helpful to generalize the proposed method to this latter situation. Also sometimes nonparametric estimation of $Y_i(t)$ may be of interest. For this purpose, some constraints should be imposed on $\mathbf{b}_i(t)$ for identifiability, for example, $E\{\mathbf{b}_i(t)\} = \mathbf{0}$. When panel count data arise (Sun and Zhao 2013), the generalization of existing nonparametric estimation procedures to cases with informative observation or censoring times is a challenging direction for future work too.

Appendices

Appendix 1: Proof of Theorem 1

To derive the asymptotic properties of the proposed estimators $\hat{\beta}$ and $\hat{\eta}$, we need the following regularity conditions analogous to those given by Lin et al. (2000) (Section 2):

(C1). $\{\widetilde{N}_i(\cdot), Y_i(\cdot), C_i, \mathbf{Z}_i\}_{i=1}^n$ are independent and identically distributed.

(C2). There exists a $\tau > 0$ such that $P(C_i \geq \tau) > 0$.

(C3). Both $\widetilde{N}_i(t)$ and $Y_i(t)$ ($0 \leq t \leq \tau$, $i = 1, \ldots, n$) are bounded.

(C4). $W(t)$ and $\mathbf{Z}_i$, $i = 1, \ldots, n$, have bounded variations and $W(t)$ converges almost surely to a deterministic function $w(t)$ uniformly in $t \in [0, \tau]$.

(C5). $A_\beta = E\{\int_0^\tau w(t) e^{\beta_0' \mathbf{Z}_i + \eta_0' \mathbf{X}_i(t)} [\mathbf{Z}_i - e_z(t)]^{\otimes 2} d\Lambda_2^*(t)\}$ and $\Omega_\eta = E\Big[\int_0^\tau \big\{\mathbf{X}_i(t) - \bar{x}(t)\big\}^{\otimes 2} e^{\eta_0' \mathbf{X}_i(t)} d\Lambda_1^*(t)\Big]$ are both positive definite.

Under condition (C2), we define

$$U_1(\beta; \hat{\eta}) = \sum_{i=1}^n \int_0^\tau W(t) \mathbf{Z}_i \Big[Y_i(t) d\widetilde{N}_i(t) - e^{\beta' \mathbf{Z}_i + \hat{\eta}' \mathbf{X}_i(t)} d\widehat{\Lambda}_2^*(t)\Big],$$

which is integrable under conditions (C3) and (C4). Also note that $d\widehat{\Lambda}_2^*(t)$ satisfies

$$\sum_{i=1}^n \Big[Y_i(t) d\widetilde{N}_i(t) - e^{\beta' \mathbf{Z}_i + \hat{\eta}' \mathbf{X}_i(t)} d\widehat{\Lambda}_2^*(t)\Big] = 0, \ \ 0 \leq t \leq \tau. \qquad (A.1)$$

Let

$$\widehat{A}_\beta(\beta) = -n^{-1} \partial U_1(\beta, \hat{\eta}) / \partial \beta', \widehat{A}_\eta(\eta) = -n^{-1} \partial U_1(\beta_0, \eta) / \partial \eta',$$

and under (C1), let

$$A_\beta = \lim_{n \to \infty} \widehat{A}_\beta(\beta_0), \ A_\eta = \lim_{n \to \infty} \widehat{A}_\eta(\eta_0).$$

The consistency of $\hat{\beta}$ and $\hat{\eta}$ follows from the facts that $U_1(\beta_0; \hat{\eta})$ and $U_\eta(\eta_0)$ both tend to 0 in probability as $n \to \infty$, and that under condition (C5), $\widehat{A}_\beta(\beta)$ and $-n^{-1} \partial U_\eta(\eta) / \partial \eta'$ both converge uniformly to the positive definite matrices A_β and Ω_η over β and η, respectively, in neighborhoods around the true values β_0 and η_0. Then the Taylor series expansions of $U_1(\hat{\beta}; \hat{\eta})$ at $(\beta_0; \hat{\eta})$ and (β_0, η_0) yield $n^{1/2}(\hat{\beta} - \beta_0) = A_\beta^{-1} n^{-1/2} U_1(\beta_0; \hat{\eta}) + o_p(1) = A_\beta^{-1} \Big\{n^{-1/2} U_1(\beta_0; \eta_0) - A_\eta n^{1/2} (\hat{\eta} - \eta_0)\Big\} + o_p(1)$. The proof of Theorem 1 is sketched as follows:

(1) First, using some derivation operation to $U_1(\beta; \hat{\eta})$ and (A.1), we can get

$$\widehat{A}_\beta(\beta) = n^{-1} \sum_{i=1}^n \int_0^\tau W(t) \big\{\mathbf{Z}_i - \widehat{E}_Z(t; \beta, \hat{\eta})\big\}^{\otimes 2} e^{\beta' \mathbf{Z}_i + \hat{\eta}' \mathbf{X}_i(t)} d\widehat{\Lambda}_2^*(t; \beta, \hat{\eta}).$$

(2) Solving $d\widehat{\Lambda}_2^*(t; \beta_0, \eta_0)$ from (A.1) and applying to $U_1(\beta_0; \eta_0)$ yields

$$U_1(\beta_0; \eta_0) = \sum_{i=1}^n \int_0^\tau w(t) \Big(\mathbf{Z}_i - e_z(t)\Big) dM_i(t) + o_p(n^{1/2}),$$

where $e_z(t) = lim_{n \to \infty} \widehat{E}_Z(t; \beta_0, \eta_0)$ as defined earlier in Section 3 and $w(t)$ is a deterministic function defined under (C5).

(3) Differentiation of $U_1(\beta_0, \eta)$ and (A.1) with respect to η yields

$$\widehat{A}_\eta(\eta) = n^{-1} \sum_{i=1}^n \int_0^\tau W(t) \big[\mathbf{Z}_i - \widehat{E}_Z(t; \beta_0, \eta)\big] e^{\beta_0' \mathbf{Z}_i + \eta' \mathbf{X}_i(t)} X_i'(t) d\widehat{\Lambda}_2^*(t; \beta_0, \eta).$$

(4) According to equation (5) and by using the asymptotic results in Lin et al. (2000) (A.5), one can show that

$$n^{1/2}\{\hat{\eta}-\eta_0\} = \Omega_\eta^{-1} n^{-1/2} \sum_{i=1}^{n} \left[\int_0^\tau \left(\mathbf{X}_i(t) - \frac{s^{(1)}(t)}{s^{(0)}(t)} \right) dM_i^*(t) \right] + o_p(1), \qquad (A.2)$$

where $\Omega_\eta = E\left[\int_0^\tau \left\{ \mathbf{X}_i(t) - \bar{x}(t) \right\}^{\otimes 2} e^{\eta_0' \mathbf{X}_i(t)} d\Lambda_1^*(t) \right]$, which is invertible under (C5).

Combining the results in steps (1)-(4), we have

$$\begin{aligned} U_1(\beta_0;\hat{\eta}) &= \sum_{i=1}^{n} \left[\int_0^\tau w(t)\{\mathbf{Z}_i - e_z(t)\} dM_i(t) \right] \\ &\quad - A_\eta \Omega_\eta^{-1} \sum_{i=1}^{n} \left[\int_0^\tau \{\mathbf{X}_i(t) - \bar{x}(t)\} dM_i^*(t) \right] + o_p(n^{1/2}). \end{aligned}$$

Since A_β is also invertible under (C5), it then follows from the multivariate central limit theorem that the conclusions hold.

Appendix 2: Proof of the null distribution of $\mathcal{F}(t,z)$ in Section 3

Let $V(\hat{\beta},\hat{\eta}) = \sum_{i=1}^{n} \int_0^t I(\mathbf{Z}_i \le z) d\widehat{M}_i(s;\hat{\beta},\hat{\eta})$, then by the Taylor expansion,

$$\mathcal{F}(t,z;\hat{\beta},\hat{\eta}) = n^{-1/2} V(\beta_0,\eta_0) + \frac{\partial V(\beta_0,\eta_0)}{n\partial \eta'} \sqrt{n}(\hat{\eta}-\eta_0) + \frac{\partial V(\beta_0,\hat{\eta})}{n\partial \beta'} \sqrt{n}(\hat{\beta}-\beta_0) + o_p(1).$$

Using the arguments and algebra manipulation similar to those in Appendix 1, we have $V(\beta_0,\eta_0) = \sum_{i=1}^{n} u_{1i}(t,z) + o_p(n^{1/2})$, where $u_{1i}(t,z) = \int_0^t \{I(\mathbf{Z}_i \le z) - e_I(s,z)\} dM_i(s)$. Also, $\frac{\partial V(\beta_0,\eta_0)}{n\partial \eta'}$ and $\frac{\partial V(\beta_0,\hat{\eta})}{n\partial \beta'}$ can be estimated by $-\widehat{\Phi}_\eta(t,z)$ and $-\widehat{\Phi}_\beta(t,z)$, respectively.

In addition, it follows from (A.2) and Theorem 1 that

$$\sqrt{n}\{\hat{\eta}-\eta_0\} = \Omega_\eta^{-1} n^{-1/2} \sum_{i=1}^{n} \left[\int_0^\tau \left(\mathbf{X}_i(t) - \frac{s^{(1)}(t)}{s^{(0)}(t)} \right) dM_i^*(t) \right] + o_p(1),$$

and

$$\sqrt{n}\{\hat{\beta}-\beta_0\} = A_\beta^{-1} n^{-1/2} \sum_{i=1}^{n} (v_{1i} - v_{2i}) + o_p(1),$$

where $v_{1i} = \int_0^\tau w(t)\Big(\mathbf{Z}_i - e_z(t)\Big) dM_i(t)$, and $v_{2i} = \int_0^\tau A_\eta \Omega_\eta^{-1} \Big(\mathbf{X}_i(t) - \bar{x}(t)\Big) dM_i^*(t)$. Therefore, $\mathcal{F}(t,z;\hat{\beta},\hat{\eta})$ can be expressed as a sum of i.i.d. mean-zero terms for fixed t. By the multivariate central limit theorem, $\mathcal{F}(t,z)$ converges in finite-dimensional distribution to a mean-zero Gaussian process. Since $\mathcal{F}(t,z)$ is tight based on the empirical process theory, $\mathcal{F}(t,z)$ converges weakly to a mean-zero Gaussian process that can be approximated by $\widehat{\mathcal{F}}(t,z)$ given by equation (10).

Table 1: Estimation results based on SEP with $\lambda_0 = 2$, $\mu_0(t) = 20t$, $\Lambda_0(t) = 5t$, $\xi_0 = 0$, $b_{1i} = b_{2i} = b_{3i}$.

	$n = 100$			$n = 200$		
β_0	0	0.2	0.5	0	0.2	0.5
			$\gamma_0 = 0$			
Bias	-0.003	0.005	0.005	0.002	-0.005	0.001
SEE	0.179	0.182	0.183	0.130	0.130	0.132
SSE	0.179	0.197	0.189	0.135	0.132	0.136
CP	0.946	0.921	0.934	0.942	0.953	0.952
			$\gamma_0 = 0.5$			
Bias	0.004	-0.001	-0.004	-0.002	-0.003	0.003
SEE	0.176	0.175	0.177	0.126	0.126	0.127
SSE	0.191	0.178	0.183	0.132	0.128	0.130
CP	0.929	0.945	0.939	0.938	0.947	0.931

Table 2: Estimation results based on GEP with $\lambda_0 = 2$, $\mu_0(t) = 20t$, $\Lambda_0(t) = 5t$, $b_{1i} = b_{2i} = b_{3i}$.

	$n = 100$			$n = 200$		
β_0	0	0.2	0.5	0	0.2	0.5
			$(\gamma_0, \xi_0) = (0, 0)$			
Bias	0.007	0.012	0.000	-0.009	-0.005	-0.003
SEE	0.177	0.177	0.179	0.127	0.128	0.129
SSE	0.194	0.188	0.199	0.134	0.129	0.132
CP	0.924	0.934	0.905	0.934	0.946	0.934
			$(\gamma_0, \xi_0) = (0, 0.2)$			
Bias	0.036	0.035	0.042	0.036	0.036	0.042
SEE	0.178	0.180	0.182	0.127	0.128	0.130
SSE	0.192	0.186	0.197	0.133	0.134	0.138
CP	0.922	0.937	0.921	0.922	0.932	0.923
			$(\gamma_0, \xi_0) = (0.5, 0)$			
Bias	0.006	-0.005	0.004	0.004	-0.003	0.002
SEE	0.173	0.174	0.174	0.123	0.125	0.125
SSE	0.177	0.179	0.183	0.126	0.130	0.130
CP	0.938	0.939	0.937	0.934	0.943	0.927
			$(\gamma_0, \xi_0) = (0.5, 0.2)$			
Bias	0.047	0.043	0.035	0.042	0.037	0.041
SEE	0.174	0.173	0.176	0.125	0.125	0.126
SSE	0.181	0.184	0.182	0.128	0.131	0.134
CP	0.918	0.922	0.936	0.929	0.931	0.923

Table 3: Estimation results based on SEP with $\lambda_0 = 2$, $\mu_0(t) = 20t$, $\Lambda_0(t) = 5t$, $\xi_0 = 0$, $b_{1i}(t) = b_i\, t^{1/3}$, $b_{2i}(t) = b_i\sqrt{t}$ and $b_{3i}(t) = b_i$.

	$n = 100$			$n = 200$		
β_0	0	0.2	0.5	0	0.2	0.5
			$\gamma_0 = 0$			
Bias	-0.006	-0.008	-0.008	0.000	-0.001	-0.003
SEE	0.176	0.178	0.179	0.127	0.128	0.129
SSE	0.187	0.185	0.186	0.133	0.132	0.134
CP	0.932	0.946	0.933	0.931	0.939	0.929
			$\gamma_0 = 0.5$			
Bias	-0.005	-0.001	-0.007	0.000	-0.002	0.005
SEE	0.172	0.172	0.173	0.123	0.124	0.124
SSE	0.173	0.179	0.177	0.124	0.126	0.122
CP	0.940	0.927	0.945	0.941	0.944	0.958

Table 4: Estimation results based on GEP with $\lambda_0 = 2$, $\mu_0(t) = 20t$, $\Lambda_0(t) = 5t$, $b_{1i}(t) = b_i t^{1/3}$, $b_{2i}(t) = b_i\sqrt{t}$ and $b_{3i}(t) = b_i$.

	$n = 100$			$n = 200$		
β_0	0	0.2	0.5	0	0.2	0.5
			$(\gamma_0, \xi_0) = (0, 0)$			
Bias	0.003	-0.005	-0.006	-0.003	-0.001	-0.004
SEE	0.172	0.171	0.173	0.123	0.123	0.125
SSE	0.182	0.181	0.181	0.127	0.128	0.130
CP	0.940	0.928	0.933	0.940	0.944	0.942
			$(\gamma_0, \xi_0) = (0, 0.2)$			
Bias	0.045	0.038	0.040	0.036	0.044	0.042
SEE	0.173	0.173	0.175	0.123	0.125	0.127
SSE	0.183	0.186	0.185	0.129	0.132	0.133
CP	0.921	0.923	0.927	0.927	0.918	0.926
			$(\gamma_0, \xi_0) = (0.5, 0)$			
Bias	0.006	-0.004	-0.002	-0.006	0.006	0.002
SEE	0.168	0.168	0.169	0.120	0.120	0.121
SSE	0.178	0.181	0.173	0.129	0.127	0.122
CP	0.939	0.933	0.944	0.939	0.928	0.944
			$(\gamma_0, \xi_0) = (0.5, 0.2)$			
Bias	0.051	0.043	0.035	0.037	0.044	0.036
SEE	0.166	0.169	0.171	0.120	0.120	0.122
SSE	0.182	0.179	0.169	0.126	0.123	0.128
CP	0.911	0.921	0.939	0.922	0.914	0.925

Table 5: Estimation results on $\eta = (\gamma, \xi)'$ given by GEP. Set-ups 1 and 2 represent the same set-ups as for Tables 2 and 4, respectively.

n	Set-up 1		Set-up 2		n	Set-up 1		Set-up 2	
100	$\hat{\gamma}$	$\hat{\xi}$	$\hat{\gamma}$	$\hat{\xi}$	200	$\hat{\gamma}$	$\hat{\xi}$	$\hat{\gamma}$	$\hat{\xi}$
	$(\gamma_0, \xi_0) = (0, 0)$					$(\gamma_0, \xi_0) = (0, 0)$			
Bias	-0.001	-0.006	0.007	0.005		-0.001	-0.006	-0.001	-0.006
SEE	0.124	0.264	0.112	0.258		0.088	0.186	0.080	0.183
SSE	0.134	0.297	0.121	0.294		0.089	0.198	0.082	0.198
CP	0.934	0.917	0.942	0.924		0.953	0.932	0.941	0.936
	$(\gamma_0, \xi_0) = (0, 0.2)$					$(\gamma_0, \xi_0) = (0, 0.2)$			
Bias	0.004	0.009	0.007	0.013		-0.002	-0.006	0.002	-0.006
SEE	0.125	0.270	0.114	0.266		0.088	0.190	0.080	0.188
SSE	0.128	0.294	0.120	0.290		0.091	0.209	0.085	0.207
CP	0.940	0.927	0.930	0.928		0.946	0.928	0.937	0.925
	$(\gamma_0, \xi_0) = (0.5, 0)$					$(\gamma_0, \xi_0) = (0.5, 0)$			
Bias	-0.001	-0.008	0.001	-0.003		-0.001	-0.006	-0.001	-0.006
SEE	0.114	0.237	0.100	0.232		0.081	0.168	0.071	0.164
SSE	0.121	0.262	0.114	0.278		0.087	0.190	0.076	0.186
CP	0.930	0.922	0.916	0.906		0.927	0.927	0.934	0.926
	$(\gamma_0, \xi_0) = (0.5, 0.2)$					$(\gamma_0, \xi_0) = (0.5, 0.2)$			
Bias	0.004	0.007	-0.003	-0.005		-0.001	-0.006	-0.001	-0.006
SEE	0.114	0.242	0.102	0.236		0.081	0.171	0.072	0.167
SSE	0.126	0.276	0.111	0.267		0.084	0.195	0.080	0.189
CP	0.926	0.918	0.920	0.911		0.945	0.917	0.926	0.915

References

Cheng, S. C., Wei, L. J. (2000). Inferences for a semiparametric model with panel data. Biometrika, 87, 89-97.

Table 6: Estimation results based on GEP with $\mathbf{Z}_i \sim N(0, 0.5)$, $\lambda_0 = 2$, $\mu_0(t) = 20t$, $\Lambda_0(t) = 5t$, $b_{1i}(t) = b_{2i}(t) = b_{3i}(t) = b_i$.

	$n = 100$				$n = 200$		
β_0	0	0.2	0.5		0	0.2	0.5
				$(\gamma_0, \xi_0) = (0, 0)$			
Bias	0.000	-0.003	-0.004		0.000	0.003	0.002
SEE	0.170	0.171	0.177		0.125	0.126	0.132
SSE	0.189	0.189	0.205		0.131	0.134	0.140
CP	0.920	0.924	0.900		0.942	0.931	0.927
				$(\gamma_0, \xi_0) = (0, 0.2)$			
Bias	0.030	0.024	0.039		0.040	0.041	0.040
SEE	0.170	0.172	0.180		0.125	0.127	0.132
SSE	0.184	0.185	0.193		0.135	0.135	0.144
CP	0.920	0.929	0.921		0.916	0.922	0.918
				$(\gamma_0, \xi_0) = (0.5, 0)$			
Bias	0.007	0.003	-0.013		0.005	-0.006	-0.004
SEE	0.174	0.173	0.174		0.128	0.126	0.127
SSE	0.189	0.188	0.187		0.136	0.133	0.136
CP	0.925	0.922	0.930		0.931	0.928	0.926
				$(\gamma_0, \xi_0) = (0.5, 0.2)$			
Bias	0.043	0.047	0.037		0.042	0.035	0.035
SEE	0.175	0.172	0.177		0.128	0.126	0.127
SSE	0.189	0.191	0.185		0.133	0.136	0.134
CP	0.906	0.921	0.937		0.925	0.919	0.924

Table 7: Estimation results on β based on the proposed procedures and the procedures given in Sun et al. (2012) and He et al. (2009) with $\beta_0 = \xi_0 = \gamma_0 = 0$.

	GEP	SEP	Sun et al. (2012)	He et al. (2009)
			M_1, $n = 100$	
Bias	-0.003	-0.010	-0.004	0.009
SSE	0.162	0.167	0.261	0.206
			M_1, $n = 200$	
Bias	-0.003	0.003	-0.003	0.007
SSE	0.116	0.114	0.184	0.154
			M_2, $n = 100$	
Bias	0.004	0.003	0.004	0.003
SSE	0.123	0.129	0.306	0.184
			M_2, $n = 200$	
Bias	-0.001	0.000	-0.003	0.011
SSE	0.089	0.087	0.227	0.145
			M_3, $n = 100$	
Bias	0.001	0.003	-0.010	0.000
SSE	0.074	0.077	0.221	0.071
			M_3, $n = 200$	
Bias	0.002	0.001	0.000	-0.003
SSE	0.055	0.055	0.150	0.051

Set-up M_1: $\mu_0(t) = 10t$, $\lambda_0 = 2$, $\Lambda_0(t) = 5t$, $b_{1i} = -b_{2i} = b_{3i} = b_i$, where b_i followed a uniform distribution on (-0.5, 0.5).

Set-up M_2: $\mu_0(t) = 10t$, $\lambda_0 = 0$, $\Lambda_0(t) = 5t$, $b_{1i} = -b_{2i} = b_i$, where b_i followed a uniform distribution on (-0.5, 0.5) and b_{3i} followed an extreme value distribution with distribution function $F(t) = 1 - \exp\{-\exp(t)\}$.

Set-up M_3: $\mu_0(t) = 4t$, $\lambda_0 = 0$, $\Lambda_0(t) = 5t$, $b_{1i} = 0.2b_{2i} + 0.2b_{2i}$, $b_{2i} = \log(b_{2i}^*)$ and $b_{3i} = \exp(v_i)$, where v_i and b_{2i}^* were generated, respectively, from a normal distribution with mean 0 and standard deviation 0.5 and gamma distribution with mean 4 and variance 8.

Table 8: Analysis results for the skin cancer data.

	Est.	SEE	95% CI	p-value
			Proposed	
γ_1	0.529	0.072	(0.387, 0.671)	< 0.001
γ_2	0.566	0.072	(0.426, 0.706)	< 0.001
ξ_1	1.203	0.171	(0.868, 1.538)	< 0.001
ξ_2	1.038	0.171	(0.703, 1.373)	< 0.001
β_1	-0.448	0.187	(-0.814, -0.082)	0.017
β_2	1.164	0.225	(0.723, 1.604)	< 0.001
			Sun et al. (2012)	
β_1	-0.404	0.253	(-0.899, 0.091)	0.110
β_2	1.340	0.287	(0.776, 1.902)	< 0.001
			He et al. (2009)	
β_1	-0.220	0.196	(-0.605, 0.164)	0.261
β_2	1.265	0.220	(0.834, 1.695)	< 0.001