Statistica Sinica **26** (2016), 1197-1218
doi:http://dx.doi.org/10.5705/ss.202014.0030

AN ADDITIVE-MULTIPLICATIVE MEAN MODEL FOR MARKER DATA CONTINGENT ON RECURRENT EVENT WITH AN INFORMATIVE TERMINAL EVENT

Miao Han, Xinyuan Song, Liuquan Sun and Lei Liu

Shanghai University of Finance and Economics,
The Chinese University of Hong Kong,
Chinese Academy of Sciences and Northwestern University

Abstract: In many medical studies some markers, such as the medical costs incurred during each hospitalisation, are only measured when an event, such as a hospitalisation, occurs. In addition, there may exist an informative terminal event that stops the follow-up. In this article, we propose an additive-multiplicative mean model for marker data contingent on recurrent event with an informative terminal event via latent variables. Estimation procedures are developed for parameter estimation, and asymptotic properties of the proposed estimators are derived. In addition, some numerical procedures are provided for model checking. The finite sample properties of the proposed estimators are examined through simulation studies. An application to a medical cost study of chronic heart failure patients from the University of Virginia Health System is illustrated.

Key words and phrases: Estimating equations, joint modeling, latent variables, marker data, recurrent event, terminal event.

1. Introduction

In longitudinal medical studies, patients may experience a particular event repeatedly over time, such as tumor recurrences, repeated hospital visits, and multiple infection episodes. Although interest often focuses on assessing the effects of covariates on certain features of the recurrent event process, the outcomes associated with each event are also of interest. Examples include medical cost incurred at each hospitalization, and symptoms of each infection, among others. Such outcomes are referred to as markers contingent on recurrent events. In many applications, investigators are interested in both the recurrent event process and the associated marker process (e.g., Cai, Zeng, and Pan (2010)).

A number of authors have studied marker data contingent on a recurrent event, and most of the methods are based on joint modelling approaches or marginal models (Wu and Bailey (1989); Tsiatis, Degruttola, and Wulfsohn (1995); Xu and Zeger (2001); Lin and Ying (2001); Hu, Sun, and Wei (2003);

Ratcliffe, Guo, and Ten Have (2004); Herring and Yang (2007); Liu and Ying (2007); Sun, Sun, and Liu (2007); Liang, Lu, and Ying (2009); Zhou, Zhao, and Sun (2013)). For example, Xu and Zeger (2001) proposed a latent variable model for the joint distribution of repeated outcomes and a time-to-event process. Lin and Ying (2001) suggested a marginal model for repeated outcomes. Sun, Sun, and Liu (2007) and Liang, Lu, and Ying (2009) considered some joint models for repeated outcomes and recurrent events via latent variables.

The aforementioned joint models or marginal models assume that the markers and the recurrent event times are independent conditional on covariates and/or latent variables. This implies that the markers have the same distribution at both the time of event occurrence and the time of no event occurrence. In reality, however, this assumption is not true when the markers are actually zeros at the time of no event. Thus, these methods are not suitable for the analysis of some marker data, such as the medical cost data.

Recently, Cai, Zeng, and Pan (2010) proposed a semiparametric proportional mean model for the marker at each event in the absence of a terminal event. In practice, however, there may exist a terminal event such as death that stops the follow-up, and the terminal event is often strongly correlated with the marker process and the recurrent event process. A motivating example is a medical cost study of heart failure patients treated at the University of Virginia Health System. For these data, Liu, Huang, and O'Quigley (2008) and Sun et al. (2012) used joint model approaches to demonstrate that the medical costs could be correlated with both hospital visits and death, and that ignoring the dependent terminal event would lead to biased estimates in modeling the medical costs and the hospital visits. There is a clear need for analyzing marker data contingent on recurrent event that takes an informative terminal event into account. Other related work includes He, Tong, and Sun (2009) and Han et al. (2014).

Most existing models assume that the covariates have additive or multiplicative effects on the mean function of the marker process, and the additive and multiplicative mean models postulate two rather different relationships between the covariates and the mean function. To enhance the modelling capability in many applications, it seems natural to consider models that allow some covariate effects to be multiplicative while allowing others to be additive. We propose a new joint modeling for the analysis of marker data contingent on recurrent event with an informative terminal event via two latent variables, wherein latent variables are introduced in modelling markers and recurrent events to account for within-subject variation. Specifically, conditional on the terminal event not occurring, an additive-multiplicative mean model is specified for the marker process at each event, and a proportional rates model is used for the recurrent event process. The proportional hazards model is used to model the terminal event. The

proposed joint model is comprehensive and flexible in that the distributions of the latent variables and the dependence structure between two latent variables is left unspecified. Our proposed joint model generalizes the approach of Cai, Zeng, and Pan (2010) by taking the terminal event into account. Moreover, the proposed additive-multiplicative mean model allows some covariate effects to be multiplicative and others to be additive, and thus includes both the additive and multiplicative mean models as special cases.

The remainder of the article is organized as follows. In Section 2, we describe the proposed joint model. Section 3 presents an estimating procedure for regression parameters of interest, and the asymptotic properties of the proposed estimators are established. In Section 4, we develop a technique for checking the adequacy of the proposed model. Section 5 reports some results from simulation studies conducted for evaluating the proposed methods. An application to the medical cost data for chronic heart failure patients from the clinical data repository at the University of Virginia Health System is provided in Section 6, and some concluding remarks are given in Section 7. Proofs are relegated to the Supplementary Material.

2. Model Specifications

Let $N(t)$ denote the counting process associated with recurrent events, and $m(t)$ be the marker measured at time t. Let X and W be the $p \times 1$ and $q \times 1$ vectors of covariates, and $Z = (X', W')'$. Let D be the terminal event time (e.g., death), and C be the 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. Let v_1 and v_2 be two latent variables that are independent of Z and may be associated with $m(t)$, $N(t)$, and D. It is assumed that given Z, the censoring time C is independent of $\{v_1, v_2, D, N(\cdot), m(\cdot)\}$.

Since it is of interest in many studies to make inference on the marker and recurrent event processes for subjects among survivors at a given time, and $m(t)$ only exists when $N(t)$ has a jump at time t (Ye, Kalbfleisch, and Schaubel (2007); Cai, Zeng, and Pan (2010)), we assume that $m(t)$ follows the additive-multiplicative mean model

$$E\{m(t)|N(t)-N(t-)=1, D \geq t, Z, v_1, v_2\} = \alpha_0(t; v_1) g_\beta(\beta_0' X) + g_\zeta(\zeta_0' W), \quad (2.1)$$

where $\alpha_0(t; v_1)$ is an unknown subject-specific function of t and v_1, β_0 and ζ_0 are $p \times 1$ and $q \times 1$ vectors of unknown regression parameters, and g_β and g_ζ are known link functions. Hence, the parameters β_0 and ζ_0 can be interpreted as the multiplicative and additive effects of covariates on the conditional mean function of $m(t)$ at each event given survival.

For the recurrent event process, it is assumed that $N(t)$ follows the marginal model

$$E\{dN(t)|D \geq t, Z, v_1, v_2\} = g_\gamma(\gamma_0' Z) d\mu_0(t; v_2), \tag{2.2}$$

where g_γ is a known positive link function such as the exponential function, γ_0 is a $(p+q) \times 1$ vector of unknown regression parameters, and $\mu_0(t; v_2)$ is an unknown subject-specific baseline mean function with $\mu_0(0; v_2) = 0$.

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

$$\log \Lambda_0(D) = -\eta_0' Z + \varepsilon, \tag{2.3}$$

where η_0 is a $(p+q) \times 1$ vector of unknown regression parameters, $\Lambda_0(t)$ is an unspecified baseline cumulative hazard function, and ε is a random variable with the extreme-value distribution. Here, ε is allowed to be correlated with v_1 and v_2.

Models (2.1) and (2.2) are conditional given the latent variables and thus the parameters have a subject-specific interpretation, which also implies the same dependence structure of the marker process and the recurrent process. In addition, under models (2.1) and (2.2),

$$\begin{aligned} &E\{m(t)|N(t) - N(t-) = 1, D \geq t, Z\} \\ &\quad = E\{\alpha_0(t; v_1)|N(t) - N(t-) = 1, D \geq t, Z\} g_\beta(\beta_0' X) + g_\zeta(\zeta_0' W), \end{aligned}$$

$$E\{dN(t)|D \geq t, Z\} = g_\gamma(\gamma_0' Z) E\{d\mu_0(t; v_2)|D \geq t, Z\}.$$

Therefore, the parameters at (2.1) and (2.2) can also serve as marginal effects of covariates given survival.

Models (2.1) and (2.2) are very flexible in that $\alpha_0(t; v_1)$ and $\mu_0(t; v_2)$ are nonparametric, and the correlation between v_1 and v_2 can be arbitrary. Also the distributions of v_1 and v_2 are completely unspecified. When $g_\beta(x) \equiv 1$, (2.1) reduces to the additive mean model, and when $g_\zeta(x) \equiv 0$, (2.1) reduces to the proportional mean model (Cai, Zeng, and Pan (2010)). Unlike Sun et al. (2012), we do not need the assumption that $m(t)$ is independent of $N(t)$ conditional on (v_1, v_2, Z) and $D \geq t$. For a random sample of n subjects, the observed data consist of $\{m_i(t)dN_i(t), N_i(t), T_i, \delta_i, X_i, W_i, 0 \leq t \leq T_i, i = 1, \ldots, n\}$ with $Z_i = (X_i', W_i')'$. Let $\theta_0 = (\beta_0', \zeta_0')'$. Our main interest is to estimate the parameters θ_0 and γ_0.

Remark 1. At (2.1), if $\alpha_0(t; v_1) = 1$ and X and W have some common factors, then β_0 and ζ_0 are not identifiable for such link functions as $g_\beta(x) = x$ and $g_\zeta(x) = x$. Thus, for identifiability reasons, we assume that X and W do not have any common factor throughout the paper; no other conditions are needed for the identifiability of β_0 and ζ_0. The intercept could be included into $g_\zeta(\cdot)$ as well, but it seems not to be straightforward to generalize the proposed approach to deal with this situation. Further research is needed.

3. Inference Procedures

Define $\mathcal{A}_0(t;v_1,v_2)=\int_0^t \alpha_0(u;v_1)d\mu_0(u;v_2)$, and $dH(t,s)=E\{d\mathcal{A}_0(t;v_1,v_2)|\varepsilon\geq s\}$. It follows from the independent censoring assumption and (2.1) and (2.2) that

$$\begin{aligned}
&E\big[\{m(t)-g_\zeta(\zeta_0'W)\}dN(t)|T\geq t,Z\big]\\
&\quad=E\big[\{m(t)-g_\zeta(\zeta_0'W)\}dN(t)|D\geq t,Z\big]\\
&\quad=E\Big(E[\{m(t)-g_\zeta(\zeta_0'W)\}dN(t)|D\geq t,Z,v_1,v_2]|D\geq t,Z\Big)\\
&\quad=E\big[E\{m(t)-g_\zeta(\zeta_0'W)|N(t)-N(t-)=1,D\geq t,Z,v_1,v_2\}\\
&\qquad E\{dN(t)|D\geq t,Z,v_1,v_2\}|D\geq t,Z\big]\\
&\quad=g_\beta(\beta_0'X)g_\gamma(\gamma_0'Z)E\{\alpha_0(t;v_1)d\mu_0(t;v_2)|D\geq t,Z\}\\
&\quad=g_\beta(\beta_0'X)g_\gamma(\gamma_0'Z)E\{d\mathcal{A}_0(t;v_1,v_2)|\varepsilon\geq\log\Lambda_0(t)+\eta_0'Z,Z\}\\
&\quad=g_\beta(\beta_0'X)g_\gamma(\gamma_0'Z)dH(t,\log\Lambda_0(t)+\eta_0'Z). \qquad (3.1)
\end{aligned}$$

By the assumption that (v_1,v_2,ε) is independent of (Z,C), for any integrable function $h(Z,t,s)$, we have

$$\begin{aligned}
dH(t,s)&=\frac{E\{d\mathcal{A}_0(t;v_1,v_2)I(\varepsilon\geq s)\}}{E\{I(\varepsilon\geq s)\}}\\
&=\frac{E\{d\mathcal{A}_0(t;v_1,v_2)I(\varepsilon\geq s)I(\log\Lambda_0(C)+\eta_0'Z\geq s)h(Z,t,s)\}}{E\{I(\varepsilon\geq s)I(\log\Lambda_0(C)+\eta_0'Z\geq s)h(Z,t,s)\}}\\
&=\frac{E\{d\mathcal{A}_0(t;v_1,v_2)I(\log\Lambda_0(T)+\eta_0'Z\geq s)h(Z,t,s)\}}{E\{I(\log\Lambda_0(T)+\eta_0'Z\geq s)h(Z,t,s)\}}.
\end{aligned}$$

By taking $h(Z,t,s)=g_\beta(\beta_0'X)g_\gamma(\gamma_0'Z)I(\log\Lambda_0(t)+\eta_0'Z\leq s)$, we have that

$$dH(t,s)=\frac{E\{g_\beta(\beta_0'X)g_\gamma(\gamma_0'Z)d\mathcal{A}_0(t;v_1,v_2)\Psi(T,Z,t,s)\}}{E\{g_\beta(\beta_0'X)g_\gamma(\gamma_0'Z)\Psi(T,Z,t,s)\}},$$

where $\Psi(T,Z,t,s)=I\{\log\Lambda_0(T)+\eta_0'Z\geq s\geq\log\Lambda_0(t)+\eta_0'Z\}$. Since $\Psi(T,Z,t,s)=1$ implies $T\geq t$, we obtain

$$\begin{aligned}
&E[\{m(t)-g_\zeta(\zeta_0'W)\}dN(t)\Psi(T,Z,t,s)]\\
&\quad=E\Big(E[\{m(t)-g_\zeta(\zeta_0'W)\}dN(t)\Psi(T,Z,t,s)|\Psi(T,Z,t,s),Z,v_1,v_2]\Big)\\
&\quad=E[E\{g_\beta(\beta_0'X)g_\gamma(\gamma_0'Z)d\mathcal{A}_0(t;v_1,v_2)\Psi(T,Z,t,s)|\Psi(T,Z,t,s),Z,v_1,v_2\}]\\
&\quad=E\{g_\beta(\beta_0'X)g_\gamma(\gamma_0'Z)d\mathcal{A}_0(t;v_1,v_2)\Psi(T,Z,t,s)\}.
\end{aligned}$$

Thus,

$$dH(t,s)=\frac{E[\{m(t)-g_\zeta(\zeta_0'W)\}dN(t)\Psi(T,Z,t,s)]}{E\{g_\beta(\beta_0'X)g_\gamma(\gamma_0'Z)\Psi(T,Z,t,s)\}}.$$

To estimate $dH(t,s)$, we need to estimate η_0 and $\Lambda_0(t)$ of model (2.3). Let $\hat{\eta}$ be the maximum partial likelihood estimator of η_0, the solution to

$$U_\eta(\eta) = \sum_{i=1}^{n} \int_0^\tau \{Z_i - \bar{Z}^D(t;\eta)\} dN_i^D(t) = 0,$$

where τ is a prespecified constant such that $P(T_i \geq \tau) > 0$, $N_i^D(t) = I(T_i \leq t, \delta_i = 1)$,

$$\bar{Z}^D(t;\eta) = \frac{\sum_{i=1}^n Z_i Y_i(t) \exp(\eta' Z_i)}{\sum_{i=1}^n Y_i(t) \exp(\eta' Z_i)},$$

and $Y_i(t) = I(T_i \geq t)$. Let $\hat{\Lambda}_0(t)$ be the Breslow estimator of $\Lambda_0(t)$, where

$$\hat{\Lambda}_0(t) = \sum_{i=1}^{n} \int_0^t \frac{dN_i^D(u)}{\sum_{j=1}^n Y_j(u) \exp(\hat{\eta}' Z_j)}.$$

Thus, for given γ and $\theta = (\beta', \zeta')'$, $dH(t,s)$ can be estimated by

$$\begin{aligned} &d\hat{H}(t,s;\theta,\gamma) \\ &= \frac{\sum_{i=1}^n \{m_i(t) - g_\zeta(\zeta' W_i)\} dN_i(t) I\{\log \hat{\Lambda}_0(T_i) + \hat{\eta}' Z_i \geq s \geq \log \hat{\Lambda}_0(t) + \hat{\eta}' Z_i\}}{\sum_{j=1}^n g_\beta(\beta' X_j) g_\gamma(\gamma' Z_j) I\{\log \hat{\Lambda}_0(T_j) + \hat{\eta}' Z_j \geq s \geq \log \hat{\Lambda}_0(t) + \hat{\eta}' Z_j\}}. \end{aligned}$$

Define

$$\hat{\Psi}_j(t;Z) = I\{\log \hat{\Lambda}_0(T_j) + \hat{\eta}' Z_j \geq \log \hat{\Lambda}_0(t) + \hat{\eta}' Z \geq \log \hat{\Lambda}_0(t) + \hat{\eta}' Z_j\}.$$

Based on (3.1), for a given γ, using the generalized estimating equation approach (Liang and Zeger (1986)) and replacing $dH(t,s)$ by $d\hat{H}(t,s;\theta,\gamma)$, we propose an estimating function for θ_0:

$$\begin{aligned} U_\theta(\theta;\gamma) = \sum_{i=1}^{n} \int_0^\tau W(t) \{Z_i - \bar{Z}_i(t;\beta,\gamma)\} Y_i(t) \Big[&\{m_i(t) - g_\zeta(\zeta' W_i)\} dN_i(t) \\ &- g_\beta(\beta' X_i) g_\gamma(\gamma' Z_i) \{d\bar{m}_i(t;\beta,\gamma) - d\bar{g}_i(t;\theta,\gamma)\} \Big], \end{aligned} \tag{3.2}$$

where $W(t)$ is a possibly data-dependent weight function,

$$\begin{aligned} \bar{Z}_i(t;\beta,\gamma) &= \frac{\sum_{j=1}^n Z_j g_\beta(\beta' X_j) g_\gamma(\gamma' Z_j) \hat{\Psi}_j(t;Z_i)}{\sum_{j=1}^n g_\beta(\beta' X_j) g_\gamma(\gamma' Z_j) \hat{\Psi}_j(t;Z_i)}, \\ d\bar{m}_i(t;\beta,\gamma) &= \frac{\sum_{j=1}^n m_j(t) dN_j(t) \hat{\Psi}_j(t;Z_i)}{\sum_{j=1}^n g_\beta(\beta' X_j) g_\gamma(\gamma' Z_j) \hat{\Psi}_j(t;Z_i)}, \\ d\bar{g}_i(t;\theta,\gamma) &= \frac{\sum_{j=1}^n g_\zeta(\zeta' W_j) dN_j(t) \hat{\Psi}_j(t;Z_i)}{\sum_{j=1}^n g_\beta(\beta' X_j) g_\gamma(\gamma' Z_j) \hat{\Psi}_j(t;Z_i)}. \end{aligned}$$

Since γ_0 is unknown, we need to estimate γ_0. Following the approach of Zeng and Cai (2010), we specify an estimating function for γ_0:

$$U_\gamma(\gamma) = \sum_{i=1}^{n} \int_0^\tau Q(t)\{Z_i - \bar{Z}_i^N(t;\gamma)\}Y_i(t)\Big[dN_i(t) - g_\gamma(\gamma' Z_i)d\bar{N}_i(t;\gamma)\Big], \quad (3.3)$$

where $Q(t)$ is a possibly data-dependent weight function,

$$\bar{Z}_i^N(t;\gamma) = \frac{\sum_{j=1}^n Z_j g_\gamma(\gamma' Z_j)\hat{\Psi}_j(t;Z_i)}{\sum_{j=1}^n g_\gamma(\gamma' Z_j)\hat{\Psi}_j(t;Z_i)},$$

$$d\bar{N}_i(t;\gamma) = \frac{\sum_{j=1}^n dN_j(t)\hat{\Psi}_j(t;Z_i)}{\sum_{j=1}^n g_\gamma(\gamma' Z_j)\hat{\Psi}_j(t;Z_i)}.$$

Let $\hat{\gamma}$ denote the solution to the equation $U_\gamma(\gamma) = 0$ and $\hat{\theta}$ denote the solution to $U_\theta(\theta;\hat{\gamma}) = 0$. By the Law of Large Numbers and the consistency of $\hat{\eta}$ and $\hat{\Lambda}_0(t)$, it can be shown that $\hat{\theta}$ and $\hat{\gamma}$ are consistent. The asymptotic normality of $\hat{\theta}$ and $\hat{\gamma}$ is as follows.

Theorem 1. *Under the regularity conditions* (C1)−(C6) *stated in the Supplementary Material,* $n^{1/2}(\hat{\theta}-\theta_0)$ *and* $n^{1/2}(\hat{\gamma}-\gamma_0)$ *have asymptotically a joint normal distribution with mean zero and covariance matrix* $A^{-1}\Sigma(A')^{-1}$*, where* A *and* Σ *are defined in the Supplementary Material.*

To estimate the asymptotic covariance of $\hat{\theta}$ and $\hat{\gamma}$, we need to estimate A and Σ. Let $\dot{g}_l(t) = dg_l(t)/dt$ for $l = \gamma, \beta$ and ζ. Define

$$\bar{X}_i(t;\beta,\gamma) = \frac{\sum_{j=1}^n X_j \dot{g}_\beta(\beta' X_j) g_\gamma(\gamma' Z_j)\hat{\Psi}_j(t;Z_i)}{\sum_{j=1}^n g_\beta(\beta' X_j) g_\gamma(\gamma' Z_j)\hat{\Psi}_j(t;Z_i)},$$

$$\bar{W}_i(t;\theta,\gamma) = \frac{\sum_{j=1}^n W_j \dot{g}_\zeta(\zeta' W_j) dN_j(t)\hat{\Psi}_j(t;Z_i)}{\sum_{j=1}^n g_\beta(\beta' X_j) g_\gamma(\gamma' Z_j)\hat{\Psi}_j(t;Z_i)},$$

$$\bar{Z}_i^\dagger(t;\beta,\gamma) = \frac{\sum_{j=1}^n Z_j g_\beta(\beta' X_j) \dot{g}_\gamma(\gamma' Z_j)\hat{\Psi}_j(t;Z_i)}{\sum_{j=1}^n g_\beta(\beta' X_j) g_\gamma(\gamma' Z_j)\hat{\Psi}_j(t;Z_i)},$$

$$\bar{Z}_i^*(t;\gamma) = \frac{\sum_{j=1}^n Z_j \dot{g}_\gamma(\gamma' Z_j)\hat{\Psi}_j(t;Z_i)}{\sum_{j=1}^n g_\gamma(\gamma' Z_j)\hat{\Psi}_j(t;Z_i)}.$$

It is easy to see that A can be consistently estimated by $\hat{A}$, where

$$\hat{A} = \begin{pmatrix} \hat{A}_{11}(\hat{\theta},\hat{\gamma}) & \hat{A}_{12}(\hat{\theta},\hat{\gamma}) \\ 0 & \hat{A}_{22}(\hat{\gamma}) \end{pmatrix},$$

$$\hat{A}_{11}(\theta,\gamma) = n^{-1}\sum_{i=1}^{n} \int_0^\tau W(t)\{Z_i - \bar{Z}_i(t;\beta,\gamma)\}Y_i(t)$$

$$\cdot\begin{pmatrix}\{X_i\dot{g}_\beta(\beta'X_i)g_\gamma(\gamma'Z_i)-g_\beta(\beta'X_i)g_\gamma(\gamma'Z_i)\bar{X}_i(t;\beta,\gamma)\}\{d\bar{m}_i(t;\beta,\gamma)-d\bar{g}_i(t;\theta,\gamma)\}\\ W_i\dot{g}_\zeta(\zeta'W_i)dN_i(t)-g_\beta(\beta'X_i)g_\gamma(\gamma'Z_i)\bar{W}_i(t;\theta,\gamma)\end{pmatrix}',$$

$$\hat{A}_{12}(\theta,\gamma)=n^{-1}\sum_{i=1}^{n}\int_0^\tau W(t)\{Z_i-\bar{Z}_i(t;\beta,\gamma)\}Y_i(t)$$
$$\times\Big\{Z_ig_\beta(\beta'X_i)\dot{g}_\gamma(\gamma'Z_i)-g_\beta(\beta'X_i)g_\gamma(\gamma'Z_i)\bar{Z}_i^\dagger(t;\beta,\gamma)\Big\}'$$
$$\{d\bar{m}_i(t;\beta,\gamma)-d\bar{g}_i(t;\theta,\gamma)\},$$
$$\hat{A}_{22}(\gamma)=n^{-1}\sum_{i=1}^{n}\int_0^\tau Q(t)\{Z_i-\bar{Z}_i^N(t;\gamma)\}Y_i(t)\Big\{Z_i\dot{g}_\gamma(\gamma'Z_i)-\bar{Z}_i^*(t;\gamma)g_\gamma(\gamma'Z_i)\Big\}'$$
$$d\bar{N}_i(t;\gamma).$$

However, since Σ involves some Hadamard derivatives and is complicated, it is difficult to estimate Σ directly. To overcome this difficulty, we propose a resampling approach (e.g., Lin, Fleming, and Wei (1994)). Let

$$S^{(0)}(t;\eta)=n^{-1}\sum_{i=1}^{n}Y_i(t)\exp(\eta'Z_i),$$
$$\hat{M}_i^D(t)=N_i^D(t)-\int_0^t Y_i(u)\exp(\hat{\eta}'Z_i)d\hat{\Lambda}_0(u),$$
$$\hat{\Omega}=n^{-1}\sum_{i=1}^{n}\int_0^\tau\{Z_i-\bar{Z}^D(t;\hat{\eta})\}^{\otimes2}dN_i^D(t),$$

where for a vector a, $a^{\otimes2}=aa'$. Set

$$\hat{\eta}^*=\hat{\eta}+\hat{\Omega}^{-1}n^{-1}\sum_{i=1}^{n}G_i\int_0^\tau\{Z_i-\bar{Z}^D(t;\hat{\eta})\}d\hat{M}_i^D(t),$$
$$\hat{\Lambda}_0^*(t)=\hat{\Lambda}_0(t)+n^{-1}\sum_{i=1}^{n}G_i\int_0^t\frac{d\hat{M}_i^D(u)}{S^{(0)}(u;\hat{\eta})}-\int_0^t\bar{Z}^D(u;\hat{\eta})'d\hat{\Lambda}_0(u)(\hat{\eta}^*-\hat{\eta}),$$

where $(G_1,\ldots,G_n)$ are independent standard normal variables independent of the observed data. Define

$$\Phi_1^*=\sum_{i=1}^{n}G_i\int_0^\tau W(t)\{Z_i-\bar{Z}_i(t;\hat{\beta},\hat{\gamma})\}Y_i(t)$$
$$\times\Big[\{m_i(t)-g_\zeta(\hat{\zeta}'W_i)\}dN_i(t)-g_\beta(\hat{\beta}'X_i)g_\gamma(\hat{\gamma}'Z_i)\{d\bar{m}_i(t;\hat{\beta},\hat{\gamma})-d\bar{g}_i(t;\hat{\theta},\hat{\gamma})\}\Big],$$
$$\Phi_2^*=\sum_{i=1}^{n}\int_0^\tau W(t)\{Z_i-\bar{Z}_i(t;\hat{\beta},\hat{\gamma})\}Y_i(t)g_\beta(\hat{\beta}'X_i)g_\gamma(\hat{\gamma}'Z_i)$$

$$\times\Bigg[-\frac{\sum_{j=1}^n G_j\{m_j(t)-g_\zeta(\hat\zeta' W_j)\}dN_j(t)\hat\Psi_j(t;Z_i)}{\sum_{j=1}^n g_\beta(\hat\beta' X_j)g_\gamma(\hat\gamma' Z_j)\hat\Psi_j(t;Z_i)}$$
$$+\frac{\sum_{j=1}^n\{m_j(t)-g_\zeta(\hat\zeta' W_j)\}dN_j(t)\hat\Psi_j(t;Z_i)}{[\sum_{j=1}^n g_\beta(\hat\beta' X_j)g_\gamma(\hat\gamma' Z_j)\hat\Psi_j(t;Z_i)]^2}\sum_{j=1}^n G_j g_\beta(\hat\beta' X_j)g_\gamma(\hat\gamma' Z_j)\hat\Psi_j(t;Z_i)\Bigg],$$

$$\Phi_3^*=\sum_{i=1}^n\int_0^\tau W(t)\{Z_i-\bar Z_i(t;\hat\beta,\hat\gamma)\}Y_i(t)g_\beta(\hat\beta' X_i)g_\gamma(\hat\gamma' Z_i)$$
$$\times\Big[d\bar m_i(t;\hat\beta,\hat\gamma)-d\bar g_i(t;\hat\theta,\hat\gamma)-\{d\bar m_i^*(t;\hat\beta,\hat\gamma)-d\bar g_i^*(t;\hat\theta,\hat\gamma)\}\Big],$$

$$\Phi_4^*=\sum_{i=1}^n G_i\int_0^\tau Q(t)\{Z_i-\bar Z_i^N(t;\hat\gamma)\}Y_i(t)[dN_i(t)-g_\gamma(\hat\gamma_0' Z_i)d\bar N_i(t;\hat\gamma)],$$

$$\Phi_5^*=\sum_{i=1}^n\int_0^\tau Q(t)\{Z_i-\bar Z_i^N(t;\hat\gamma)\}Y_i(t)g_\gamma(\hat\gamma' Z_i)\Bigg[-\frac{\sum_{j=1}^n G_j dN_j(t)\hat\Psi_j(t;Z_i)}{\sum_{j=1}^n g_\gamma(\hat\gamma' Z_j)\hat\Psi_j(t;Z_i)}$$
$$+\frac{\sum_{j=1}^n dN_j(t)\hat\Psi_j(t;Z_i)}{[\sum_{j=1}^n g_\gamma(\hat\gamma' Z_j)\hat\Psi_j(t;Z_i)]^2}\sum_{j=1}^n G_j g_\gamma(\hat\gamma' Z_j)\hat\Psi_j(t;Z_i)\Bigg],$$

$$\Phi_6^*=\sum_{i=1}^n\int_0^\tau Q(t)\{Z_i-Z_i^N(t;\hat\gamma)\}Y_i(t)g_\gamma(\hat\gamma' Z_i)[d\bar N_i(t;\hat\gamma)-d\bar N_i^*(t;\hat\gamma)],$$

where $\bar m_i^*(t;\beta,\gamma)$, $\bar g_i^*(t;\theta,\gamma)$ and $\bar N_i^*(t;\gamma)$ are defined the same way as $\bar m_i(t;\beta,\gamma)$, $\bar g_i(t;\theta,\gamma)$ and $\bar N_i(t;\gamma)$ except that $(\hat\eta,\hat\Lambda_0)$ is replaced with $(\hat\eta^*,\hat\Lambda_0^*)$. Let E_G denote the conditional expectation with respect to $(G_1,\ldots,G_n)$ given the observed data, and $\hat U=(\hat U_1',\hat U_2')'$, where $\hat U_1=n^{-1/2}(\Phi_1^*+\Phi_2^*+\Phi_3^*)$ and $\hat U_2=n^{-1/2}(\Phi_4^*+\Phi_5^*+\Phi_6^*)$. To estimate Σ, we produce a large number of realizations of $\hat U$ by repeatedly generating the random samples $(G_1,\ldots,G_n)$, while fixing the observation data. Then Σ can be approximated by the empirical covariance matrix of $\hat U$.

Theorem 2. *Under the conditions of Theorem* 1, $E_G[\hat U^{\otimes 2}]$ *converges in probability to* Σ.

4. Model Checking

In this section, we propose some numerical procedures for assessing the adequacy of the proposed joint model. To check the adequacy of model (2.2), we can use a graphical procedure of Zeng and Cai (2010) for recurrent event data with an informative terminal event. To check model (2.3), we can use some goodness-of-fit methods for the proportional hazards model with right-censored data (e.g., Schoenfeld (1982); Therneau, Grambsch, and Fleming (1990); Lin, Wei, and Ying

(1993)). Next, we propose some numerical procedures for assessing the adequacy of model (2.1). Let

$$d\hat{M}_i^*(t) = Y_i(t)\Big[\{m_i(t) - g_\zeta(\hat{\zeta}'W_i)\}dN_i(t) - g_\beta(\hat{\beta}'X_i)g_\gamma(\hat{\gamma}'Z_i)\{d\bar{m}_i(t;\hat{\beta},\hat{\gamma}) - d\bar{g}_i(t;\hat{\theta},\hat{\gamma})\}\Big].$$

Following Lin, Wei, and Ying (1993) and Pan and Lin (2005), we consider the cumulative sums of residual

$$\mathcal{F}(z,t) = n^{-1/2}\sum_{i=1}^{n}\int_0^t I(Z_i \le z)d\hat{M}_i^*(u),$$

where the event $I(Z_i \le z)$ means that each of the components of Z_i is no larger than the corresponding component of z. Define

$$\begin{aligned}
\Phi_7^*(z,t) &= \sum_{i=1}^{n} G_i \int_0^t I(Z_i \le z)Y_i(u)\Big[\{m_i(u) - g_\zeta(\hat{\zeta}'W_i)\}dN_i(u) \\
&\quad - g_\beta(\hat{\beta}'X_i)g_\gamma(\hat{\gamma}'Z_i)\{d\bar{m}_i(u;\hat{\beta},\hat{\gamma}) - d\bar{g}_i(u;\hat{\theta},\hat{\gamma})\}\Big], \\
\Phi_8^*(z,t) &= \sum_{i=1}^{n}\int_0^t I(Z_i \le z)Y_i(u)g_\beta(\hat{\beta}'X_i)g_\gamma(\hat{\gamma}'Z_i) \\
&\quad \times\Big[-\frac{\sum_{j=1}^n G_j\{m_j(u) - g_\zeta(\hat{\zeta}'W_j)\}dN_j(u)\hat{\Psi}_j(u;Z_i)}{\sum_{j=1}^n g_\beta(\hat{\beta}'X_j)g_\gamma(\hat{\gamma}'Z_j)\hat{\Psi}_j(u;Z_i)} \\
&\quad + \frac{\sum_{j=1}^n\{m_j(u) - g_\zeta(\hat{\zeta}'W_j)\}dN_j(u)\hat{\Psi}_j(u;Z_i)}{[\sum_{j=1}^n g_\beta(\hat{\beta}'X_j)g_\gamma(\hat{\gamma}'Z_j)\hat{\Psi}_j(u;Z_i)]^2} \\
&\quad \times\sum_{j=1}^{n} G_j g_\beta(\hat{\beta}'X_j)g_\gamma(\hat{\gamma}'Z_j)\hat{\Psi}_j(u;Z_i)\Big], \\
\Phi_9^*(z,t) &= \sum_{i=1}^{n}\int_0^t I(Z_i \le z)Y_i(u)g_\beta(\hat{\beta}'X_i)g_\gamma(\hat{\gamma}'Z_i)\Big[d\bar{m}_i(u;\hat{\beta},\hat{\gamma}) - d\bar{g}_i(u;\hat{\theta},\hat{\gamma}) \\
&\quad - \{d\bar{m}_i^*(u;\hat{\beta},\hat{\gamma}) - d\bar{g}_i^*(u;\hat{\theta},\hat{\gamma})\}\Big], \\
\hat{\Gamma}_1(z,t) &= n^{-1}\sum_{i=1}^{n}\int_0^t I(Z_i \le z)Y_i(u) \\
&\quad \times\Bigg(\{X_i\dot{g}_\beta(\hat{\beta}'X_i)g_\gamma(\hat{\gamma}'Z_i) - g_\beta(\hat{\beta}'X_i)g_\gamma(\hat{\gamma}'Z_i)\bar{X}_i(u;\hat{\beta},\hat{\gamma})\}\{d\bar{m}_i(u;\hat{\beta},\hat{\gamma}) \\
&\quad - d\bar{g}_i(u;\hat{\theta},\hat{\gamma})\}W_i\dot{g}_\zeta(\hat{\zeta}'W_i)dN_i(u) - g_\beta(\hat{\beta}'X_i)g_\gamma(\hat{\gamma}'Z_i)\bar{W}_i(u;\hat{\theta},\hat{\gamma})\Bigg),
\end{aligned}$$

$$\hat{\Gamma}_2(z,t) = n^{-1}\sum_{i=1}^{n}\int_0^t I(Z_i \le z)Y_i(u)\Big\{Z_i g_\beta(\hat{\beta}'X_i)\dot{g}_\gamma(\hat{\gamma}'Z_i) - g_\beta(\hat{\beta}'X_i)g_\gamma(\hat{\gamma}'Z_i)\bar{Z}_i^\dagger(u;\hat{\beta},\hat{\gamma})\Big\}\{d\bar{m}_i(u;\hat{\beta},\hat{\gamma}) - d\bar{g}_i(u;\hat{\theta},\hat{\gamma})\},$$

and $\hat{\Gamma}(z,t) = (\hat{\Gamma}_1(z,t)', \hat{\Gamma}_2(z,t)')'$. Let the null hypothesis H_0 denote the correct specification of model (2.1). The null distribution of $\mathcal{F}(z,t)$ is established as follows.

Theorem 3. *Under the null hypothesis H_0 and the conditions of Theorem 1, the null distribution of $\mathcal{F}(z,t)$ can be approximated by the zero-mean Gaussian process $\hat{\mathcal{F}}(z,t)$, where*

$$\hat{\mathcal{F}}(z,t) = n^{-1/2}\Big[\Phi_7^*(z,t) + \Phi_8^*(z,t) + \Phi_9^*(z,t)\Big] - \hat{\Gamma}(z,t)'\hat{A}^{-1}\hat{U}. \tag{4.1}$$

Thus we can first obtain a large number of realizations of $\hat{\mathcal{F}}(z,t)$ by repeatedly generating the standard normal random sample $(G_1,\ldots,G_n)$ while fixing the observed data, and then plot $\mathcal{F}(z,t)$ along with a few realizations of $\hat{\mathcal{F}}(z,t)$. An unusual pattern of $\mathcal{F}(z,t)$ compared to the realizations of $\hat{\mathcal{F}}(z,t)$ would indicate a lack-of-fit of model (2.1). More formally, we can apply the supremum test statistic $\sup_{z,t}|\mathcal{F}(z,t)|$, whose p-value can be obtained by comparing the observed value of $\mathcal{F}(z,t)$ with a large number of realizations from $\hat{\mathcal{F}}(z,t)$.

5. Simulation Studies

We conducted simulation studies to examine the finite sample properties of the proposed estimators. In the study, we generated two covariates with X from a Bernoulli distribution with success probability 0.5, and W from a uniform distribution on $(0,1)$. The terminal event time D was generated from $\log\Lambda_0(D) = -\eta_1 X - \eta_2 W + \varepsilon$, with $\Lambda_0(t) = t/4$, $\eta_1 = -0.5$ and $\eta_2 = 1$, where ε is the extreme-value distribution. The censoring time C was taken as $C^* \wedge \tau$, where C^* was uniform on $(2,10)$ and $\tau = 6$.

Let $v_1 = \phi_1\varepsilon/4$ and $v_2 = \mu\exp(-\phi_2\varepsilon/4)$ for $\phi_1 = 0$, 1 or -1, and $\phi_2 = 0$, 1 or -1, where μ is uniform on $(0.5, 1.5)$. Given X, W, v_2, and $T = C \wedge D$, the recurrent event process $N(t)$ was generated from a Poisson process with the intensity function

$$\lambda(t) = v_2 g_\gamma(\gamma_1 X + \gamma_2 W)I(T \ge t),$$

where $g_\gamma(t) = \exp(t)$, $\gamma_1 = -1$, and $\gamma_2 = 0.5$.

At each occurred event, the marker process $m(t)$ was generated as

$$m(t) = (0.5t + v_1)g_\beta(\beta_0 X) + g_\zeta(\zeta_0 W) + \epsilon(t),$$

Table 1. Simulation results for the estimation of γ_1, γ_2, β_0, and ζ_0 with $n = 100$.

		(ϕ_1, ϕ_2)								
		(0, 0)	(1, 0)	(-1, 0)	(0, 1)	(1, 1)	(-1, 1)	(0, -1)	(1, -1)	(-1, -1)
γ_1	Bias	-0.0072	-0.0039	-0.0098	-0.0031	-0.0064	-0.0010	-0.0073	-0.0000	-0.0013
	SE	0.1886	0.1929	0.1866	0.1878	0.1993	0.1981	0.1960	0.1939	0.2007
	SEE	0.1909	0.1917	0.1906	0.1964	0.1997	0.1961	0.1936	0.1948	0.1932
	CP	0.9540	0.9460	0.9470	0.9570	0.9430	0.9480	0.9440	0.9500	0.9330
γ_2	Bias	0.0016	0.0091	0.0013	-0.0095	-0.0020	0.0049	-0.0079	0.0084	-0.0091
	SE	0.3191	0.3183	0.3244	0.3270	0.3302	0.3226	0.3352	0.3233	0.3347
	SEE	0.3131	0.3123	0.3120	0.3286	0.3289	0.3269	0.3150	0.3133	0.3158
	CP	0.9510	0.9390	0.9370	0.9510	0.9430	0.9410	0.9270	0.9360	0.9280
β_0	Bias	0.0192	0.0273	-0.0052	0.0123	0.0104	0.0115	0.0241	0.0335	0.0169
	SE	0.2686	0.3409	0.3155	0.2860	0.3555	0.2876	0.2700	0.2890	0.3065
	SEE	0.2676	0.3180	0.3144	0.2869	0.3640	0.2765	0.2572	0.2698	0.2932
	CP	0.9510	0.9540	0.9590	0.9580	0.9540	0.9540	0.9490	0.9440	0.9560
ζ_0	Bias	-0.0064	-0.0038	-0.0216	0.0002	-0.0048	-0.0181	-0.0209	-0.0025	0.0005
	SE	0.3469	0.3822	0.3449	0.3360	0.3801	0.3524	0.3605	0.3773	0.3560
	SEE	0.3362	0.3685	0.3472	0.3316	0.3743	0.3391	0.3445	0.3717	0.3540
	CP	0.9280	0.9340	0.9360	0.9430	0.9440	0.9290	0.9280	0.9370	0.9450

Note: Bias is the sample mean of the estimate minus the true value, SE is the sampling standard error, SEE is the sample mean of the standard error estimate, and CP is the 95% empirical coverage probability.

where $g_\beta(t) = \exp(t)$, $g_\zeta(t) = t$, $\beta_0 = 0.5$, $\zeta_0 = 1$, $\epsilon(t)$ normal with mean ψ and variance 0.25 for all t, and ψ normal with mean 0 and variance 0.25. For each simulation study, we set the weight functions $W(t) = Q(t) = 1$. The results presented below are based on 1,000 replications with sample sizes $n = 100$, 200 and 400. The asymptotic variance was estimated using the resampling method with 100 realizations, which were found to be adequate. All simulations were conducted in MATLAB (the simulation code is available upon request).

Tables 1, 2, and 3 present the simulation results on the estimates of γ_1, γ_2, β_0 and ζ_0 with $n = 100$, 200, and 400, respectively. In these tables, Bias is the sample mean of the estimate minus the true value, SE is the sampling standard error of the estimate, SEE is the sample mean of the standard error estimate, and CP is the 95% empirical coverage probability based on the normal approximation. It can be seen from the tables that the proposed estimators are nearly unbiased, there is a good agreement between the estimated and the empirical standard errors, and the 95% empirical coverage probabilities are reasonable. The results are better when the sample size increases from 100 to 400.

Table 2. Simulation results for the estimation of γ_1, γ_2, β_0, and ζ_0 with $n = 200$.

		(ϕ_1, ϕ_2)								
		(0, 0)	(1, 0)	(-1, 0)	(0, 1)	(1, 1)	(-1, 1)	(0, -1)	(1, -1)	(-1, -1)
γ_1	Bias	-0.0038	-0.0047	-0.0122	0.0002	0.0024	0.0044	0.0011	-0.0014	-0.0006
	SE	0.1308	0.1303	0.1337	0.1359	0.1286	0.1323	0.1327	0.1324	0.1371
	SEE	0.1318	0.1322	0.1326	0.1361	0.1359	0.1364	0.1342	0.1337	0.1345
	CP	0.9560	0.9500	0.9420	0.9380	0.9600	0.9590	0.9480	0.9560	0.9550
γ_2	Bias	-0.0049	-0.0003	0.0006	-0.0026	0.0063	0.0030	-0.0085	0.0090	0.0093
	SE	0.2195	0.2212	0.2170	0.2261	0.2314	0.2217	0.2241	0.2252	0.2341
	SEE	0.2165	0.2184	0.2182	0.2276	0.2288	0.2286	0.2230	0.2228	0.2213
	CP	0.9430	0.9410	0.9550	0.9540	0.9440	0.9520	0.9490	0.9360	0.9280
β_0	Bias	0.0045	0.0063	-0.0019	0.0039	-0.0001	-0.0015	0.0023	0.0055	0.0012
	SE	0.1764	0.1832	0.1878	0.1866	0.2108	0.1781	0.1687	0.1637	0.1983
	SEE	0.1690	0.1796	0.1815	0.1774	0.2071	0.1785	0.1608	0.1611	0.1858
	CP	0.9460	0.9540	0.9510	0.9550	0.9490	0.9490	0.9500	0.9490	0.9510
ζ_0	Bias	0.0121	0.0020	-0.0063	-0.0128	-0.0113	-0.0066	0.0034	-0.0124	0.0073
	SE	0.2483	0.2674	0.2425	0.2335	0.2588	0.2469	0.2471	0.2680	0.2535
	SEE	0.2380	0.2611	0.2459	0.2322	0.2612	0.2399	0.2460	0.2640	0.2519
	CP	0.9350	0.9380	0.9500	0.9410	0.9460	0.9330	0.9400	0.9330	0.9400

Note: Bias is the sample mean of the estimate minus the true value, SE is the sampling standard error, SEE is the sample mean of the standard error estimate, and CP is the 95% empirical coverage probability.

For comparison, we considered the method of Cai, Zeng, and Pan (2010) (denoted by CZP) by treating the terminal event as noninformative. We used the same setup as in Table 3, but with $m(t) = (0.5t + v_1)g_\beta(\beta_0 X + \zeta_0 W) + \epsilon(t)$. The comparison results are presented in Table 4. As expected, when the terminal event is noninformative (i.e., $(\phi_1, \phi_2) = (0, 0)$), the CZP estimators are unbiased. Under such a situation, both methods provide reasonable and comparable estimates, and the variances of our method are only slightly larger than those of the CZP method. This is because the latter utilizes the assumption of an independent terminal event in estimation. However, when the terminal event is informative, the CZP estimators can have large bias. We also considered other setups and obtained similar results.

To investigate the performance of the model-checking method, we also conducted some simulations to assess the size and power of the supremum test statistic $\sup_{z,t} |F(z,t)|$, using the same setup as in Table 3 except that the marker process $m(t)$ was

$$m(t) = (0.5t + v_1)g_\beta(\beta_0 X) + g_\zeta(\zeta_0 W) + 1.5kXW + \epsilon(t),$$

Table 3. Simulation results for the estimation of γ_1, γ_2, β_0, and ζ_0 with $n = 400$.

		(ϕ_1, ϕ_2)								
		(0, 0)	(1, 0)	(-1, 0)	(0, 1)	(1, 1)	(-1, 1)	(0, -1)	(1, -1)	(-1, -1)
γ_1	Bias	0.0027	-0.0015	-0.0009	0.0001	0.0051	0.0012	0.0002	0.0017	-0.0041
	SE	0.0898	0.0941	0.0928	0.0967	0.0918	0.0922	0.0921	0.0966	0.0917
	SEE	0.0917	0.0921	0.0918	0.0959	0.0952	0.0955	0.0936	0.0939	0.0937
	CP	0.9570	0.9390	0.9430	0.9420	0.9540	0.9500	0.9510	0.9400	0.9520
γ_2	Bias	-0.0003	-0.0007	0.0068	-0.0106	0.0020	-0.0007	0.0070	-0.0023	0.0078
	SE	0.1569	0.1564	0.1502	0.1589	0.1640	0.1612	0.1596	0.1608	0.1557
	SEE	0.1535	0.1532	0.1533	0.1603	0.1601	0.1608	0.1574	0.1580	0.1579
	CP	0.9430	0.9460	0.9500	0.9570	0.9350	0.9410	0.9460	0.9440	0.9570
β_0	Bias	0.0036	0.0036	-0.0031	-0.0051	0.0054	-0.0045	0.0022	0.0023	0.0069
	SE	0.1177	0.1241	0.1214	0.1252	0.1369	0.1205	0.1085	0.1142	0.1177
	SEE	0.1142	0.1205	0.1206	0.1216	0.1383	0.1214	0.1098	0.1106	0.1204
	CP	0.9460	0.9420	0.9500	0.9370	0.9540	0.9470	0.9550	0.9450	0.9570
ζ_0	Bias	0.0053	-0.0032	-0.0028	0.0029	-0.0000	0.0002	0.0040	0.0139	-0.0072
	SE	0.1733	0.1887	0.1816	0.1655	0.1866	0.1704	0.1736	0.1910	0.1773
	SEE	0.1712	0.1875	0.1751	0.1671	0.1858	0.1712	0.1738	0.1844	0.1777
	CP	0.9420	0.9460	0.9370	0.9560	0.9410	0.9460	0.9470	0.9410	0.9440

Note: Bias is the sample mean of the estimate minus the true value, SE is the sampling standard error, SEE is the sample mean of the standard error estimate, and CP is the 95% empirical coverage probability.

with $k = 0$, 1, 2, 3 and 4. We considered the null hypothesis H_0 as $k = 0$. Table 5 presents the empirical sizes and powers of the proposed test at the significance level of 0.05. The results suggest 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 expected, the power increases as the value of k increases.

6. An Application

We illustrate our method using the medical cost data of chronic heart failure patients treated at the University of Virginia Health System (Liu, Huang, and O'Quigley (2008); Sun et al. (2012)). In the study, there were 1,475 patients aged 60-89 years who were first diagnosed with heart failure and treated in 2004. For each patient, the observed information includes the clinical visit times (in months) and the medical cost for each hospital visit. In addition, three baseline covariates were measured: gender, race, and age. The follow-up ended with each patient's last hospital admission up to July 31, 2006, or death date. For our analysis, we focus on the effects of gender, race, and age on the medical cost with an informative terminal event (death).

Table 4. Comparison results on estimation of $(\gamma_1, \gamma_2, \beta_0, \zeta_0)$ with $n = 400$.

		(ϕ_1, ϕ_2)								
		(0, 0)	(1, 0)	(-1, 0)	(0, 1)	(1, 1)	(-1, 1)	(0, -1)	(1, -1)	(-1, -1)
		Ours								
γ_1	Bias	0.0038	0.0023	0.0032	0.0049	0.0038	-0.0021	-0.0049	0.0029	-0.0016
	SE	0.0912	0.0935	0.0882	0.0947	0.0992	0.0965	0.0930	0.0875	0.0975
γ_2	Bias	0.0004	-0.0003	-0.0014	-0.0046	0.0013	0.0027	0.0073	0.0017	-0.0027
	SE	0.1471	0.1538	0.1499	0.1649	0.1576	0.1634	0.1638	0.1623	0.1586
β_0	Bias	-0.0019	-0.0058	-0.0026	0.0015	0.0017	-0.0018	-0.0000	-0.0067	-0.0033
	SE	0.0886	0.1053	0.1015	0.0946	0.1134	0.0963	0.0831	0.0952	0.1011
ζ_0	Bias	-0.0017	-0.0004	-0.0014	-0.0043	0.0000	0.0047	0.0068	-0.0031	0.0055
	SE	0.1741	0.1932	0.1928	0.1876	0.2296	0.1986	0.1671	0.1740	0.2148
		CZP								
γ_1	Bias	0.0029	0.0018	0.0041	0.0596	0.0563	0.0507	-0.0562	-0.0494	-0.0503
	SE	0.0864	0.0880	0.0838	0.0859	0.0897	0.0885	0.0869	0.0832	0.0889
γ_2	Bias	0.0003	0.0008	-0.0015	-0.1125	-0.1047	-0.1042	0.1106	0.1072	0.0971
	SE	0.1399	0.1420	0.1389	0.1499	0.1420	0.1493	0.1477	0.1475	0.1458
β_0	Bias	-0.0010	-0.0651	0.0650	0.0186	-0.0534	0.0880	-0.0153	-0.0700	0.0477
	SE	0.0780	0.0843	0.0832	0.0833	0.0942	0.0772	0.0715	0.0800	0.0799
ζ_0	Bias	-0.0042	0.1152	-0.1308	-0.0378	0.1112	-0.1716	0.0377	0.1267	-0.0851
	SE	0.1579	0.1684	0.1507	0.1549	0.1867	0.1587	0.1490	0.1510	0.1667

Note: Bias is the sample mean of the estimate minus the true value, and SE is the sampling standard error. CZP stands for the method of Cai, Zeng, and Pan (2010).

Table 5. The empirical sizes and powers of the model test with $n = 400$.

	(ϕ_1, ϕ_2)								
	(0, 0)	(1, 0)	(-1, 0)	(0, 1)	(1, 1)	(-1, 1)	(0, -1)	(1, -1)	(-1, -1)
$k = 0$	0.046	0.044	0.041	0.045	0.048	0.052	0.044	0.050	0.042
$k = 1$	0.372	0.319	0.366	0.393	0.396	0.365	0.346	0.329	0.378
$k = 2$	0.801	0.749	0.805	0.791	0.812	0.794	0.741	0.697	0.819
$k = 3$	0.949	0.926	0.956	0.958	0.945	0.947	0.919	0.894	0.948
$k = 4$	0.980	0.982	0.982	0.978	0.979	0.977	0.973	0.960	0.978

Following Sun et al. (2012), we take $m(t)$ as the log-transformed cost. For covariates, let Z_1 be a binary indicator of gender (male=1, female=0), Z_2 be a binary indicator of race (white=1, nonwhite=0), and Z_3 denote the age group, taking values 0, 1, and 2 for 60-69, 70-79, and 80-89 years, respectively. Let τ be the longest follow-up time. For illustration, we assumed that the hospital visit process and the death time can be described by model (2.2) with $g_\gamma(x) = \exp(x)$ and model (2.3), respectively. For model (2.1) with three covariates, we examined all the possible combinations of covariate effects in terms of being additive or

multiplicative. Thus, we considered eight models for $m(t)$ with $g_\beta(x) = \exp(x)$ and $g_\zeta(x) = x$.
The additive-effects-only model (denoted by AM):

$$E\{m(t)|N(t) - N(t-) = 1, D \geq t, Z, v_1, v_2\} = \alpha_0(t; v_1) + g_\zeta(\zeta_1 Z_1 + \zeta_2 Z_2 + \zeta_3 Z_3).$$

The multiplicative-effects-only model (denoted by MM):

$$E\{m(t)|N(t) - N(t-) = 1, D \geq t, Z, v_1, v_2\} = \alpha_0(t; v_1) g_\beta(\beta_1 Z_1 + \beta_2 Z_2 + \beta_3 Z_3).$$

The additive gender effect and multiplicative race and age effects model (denoted by AMM1):

$$E\{m(t)|N(t)-N(t-) = 1, D \geq t, Z, v_1, v_2\} = \alpha_0(t; v_1) g_\beta(\beta_2 Z_2+\beta_3 Z_3)+g_\zeta(\zeta_1 Z_1).$$

The additive race effect and multiplicative gender and age effects model (denoted by AMM2):

$$E\{m(t)|N(t)-N(t-) = 1, D \geq t, Z, v_1, v_2\} = \alpha_0(t; v_1) g_\beta(\beta_1 Z_1+\beta_3 Z_3)+g_\zeta(\zeta_2 Z_2).$$

The additive age effect and multiplicative gender and race effects model (denoted by AMM3):

$$E\{m(t)|N(t)-N(t-) = 1, D \geq t, Z, v_1, v_2\} = \alpha_0(t; v_1) g_\beta(\beta_1 Z_1+\beta_2 Z_2)+g_\zeta(\zeta_3 Z_3).$$

The additive race and age effects and multiplicative gender effect model (denoted by AMM4):

$$E\{m(t)|N(t)-N(t-) = 1, D \geq t, Z, v_1, v_2\} = \alpha_0(t; v_1) g_\beta(\beta_1 Z_1)+g_\zeta(\zeta_2 Z_2+\zeta_3 Z_3).$$

The additive gender and age effects and multiplicative race effect model (denoted by AMM5):

$$E\{m(t)|N(t)-N(t-) = 1, D \geq t, Z, v_1, v_2\} = \alpha_0(t; v_1) g_\beta(\beta_2 Z_2)+g_\zeta(\zeta_1 Z_1+\zeta_3 Z_3).$$

The additive gender and race effects and multiplicative age effect model (denoted by AMM6):

$$E\{m(t)|N(t)-N(t-) = 1, D \geq t, Z, v_1, v_2\} = \alpha_0(t; v_1) g_\beta(\beta_3 Z_3)+g_\zeta(\zeta_1 Z_1+\zeta_2 Z_2).$$

We took $W(t) = Q(t) = 1$, and used 500 Monte Carlo samples to estimate the asymptotic variance. The analysis results are summarized in Table 6. These results suggest that both race and age had significant effects on the medical cost and the hospital visits, but gender did not seem to have significant effect on the medical cost and the hospital visits regardless of which model is used. In

Table 6. Joint analysis of the medical cost data for heart failure patients.

	Male			White			Age		
	Est.	SE	p-value	Est.	SE	p-value	Est.	SE	p-value
η	0.2502	0.1182	0.0343	-0.2597	0.1282	0.0427	0.4707	0.0732	0.0000
γ	-0.0046	0.0319	0.8847	-0.1275	0.0374	0.0007	0.1309	0.0211	0.0000
AM	0.0698	0.0827	0.3986	-0.3618	0.1001	0.0003	-0.1212	0.0487	0.0128
MM	0.0125	0.0135	0.3543	-0.0608	0.0182	0.0008	-0.0209	0.0078	0.0073
AMM1	0.0742	0.0839	0.3767	-0.0607	0.0164	0.0002	-0.0212	0.0082	0.0093
AMM2	0.0107	0.0135	0.4268	-0.3708	0.1009	0.0002	-0.0189	0.0077	0.0146
AMM3	0.0128	0.0138	0.3550	-0.0586	0.0159	0.0002	-0.1267	0.0495	0.0105
AMM4	0.0111	0.0132	0.4005	-0.3644	0.1003	0.0003	-0.1197	0.0489	0.0145
AMM5	0.0777	0.0834	0.3519	-0.0585	0.0160	0.0003	-0.1283	0.0493	0.0092
AMM6	0.0662	0.0831	0.4255	-0.3684	0.1007	0.0003	-0.0192	0.0077	0.0126

Note: Est. is the estimate of the parameter, and SE is the standard error estimate.

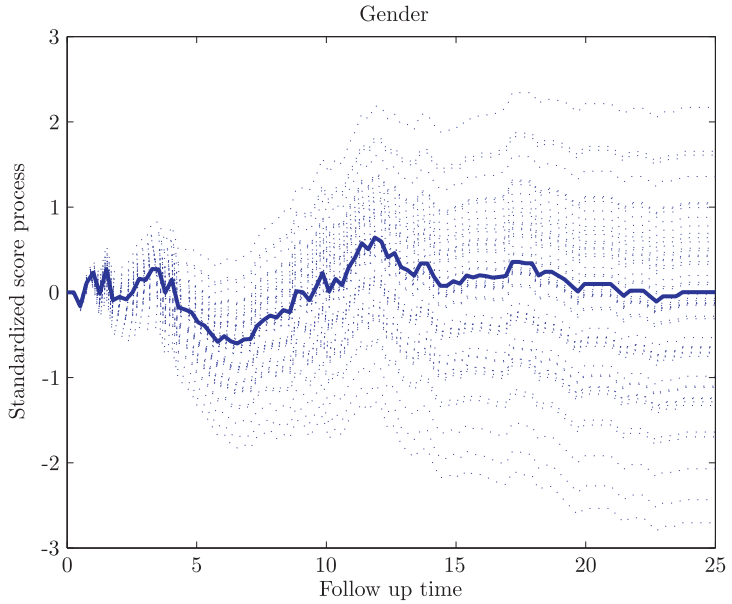

Figure 1. Plot of the standardized score process versus follow-up time for gender in the Cox model. Bold line: observed process; dotted lines: 50 simulated process.

particular, white patients were likely to be at lower risk for hospitalization and had lower medical costs compared with nonwhite patients. Older patients tended to be at higher risk for hospitalization and had less medical costs. In addition, white patients had lower mortality rate, while male and older patients tended to have higher mortality rates. These findings are consistent with those obtained

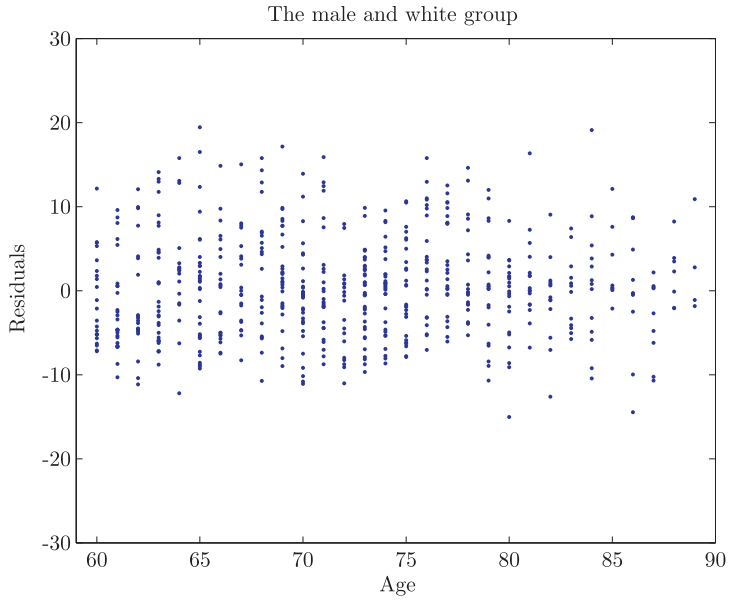

Figure 2. The residual plot of the hospital process versus age of the subjects in the male and white group.

by Liu, Huang, and O'Quigley (2008) and Sun et al. (2012).

For model checking, we first applied the standardized score process (e.g., Lin, Wei, and Ying (1993)) to check the adequacy of model (2.3). Figure 1 presents the observed score process for gender along with 50 simulated processes. The plots for race and age are similar and thus ignored. These plots indicate that model (2.3) fit the data well. To assess the goodness-of-fit of model (2.2), as discussed in Zeng and Cai (2010), we examined the residuals for each subject, $\int_0^{T_i}[dN_i(t) - g_\gamma(\hat{\gamma}'Z_i)d\bar{N}_i(t;\hat{\gamma})]$. When model (2.2) is correct, these statistics should have an approximate mean zero and be independent of Z_i. Thus, a graphical way to assess the adequacy of model (2.2) is to plot the residuals against the covariate Z_i. Figure 2 displays the residuals for each subject versus age within the male and white group. Other residual plots are similar and thus ignored. The results show that the residuals fluctuate around zero and appear to be random, indicating no evidence against model (2.2). Finally, we used the model-checking techniques introduced in Section 3 to evaluate the performances of the eight models. The p-values of the supremum test statistics for these models are presented in Table 7, based on 500 realizations of the statistic $\sup_{z,t}|\hat{\mathcal{F}}(z,t)|$. These results suggest that there is little evidence against the eight assumed models.

In order to examine which model fit the data better, we used the Akaike information criterion (AIC) and the Bayesian information criterion (BIC), where

Table 7. Model checking for the medical cost data of heart failure patients.

	AM	MM	AMM1	AMM2	AMM3	AMM4	AMM5	AMM6
$\sup_{z,t}\lvert\mathcal{F}(z,t)\rvert$	22.0984	21.2676	21.2313	21.6541	21.9006	22.1285	21.8783	21.6134
p-value	0.3960	0.4360	0.3620	0.4020	0.3820	0.4000	0.4180	0.3920
AIC	7.3323	7.2519	7.2416	7.3199	7.2770	7.3418	7.2667	7.3105
BIC	7.3431	7.2626	7.2524	7.3307	7.2878	7.3526	7.2775	7.3213

AIC $= 2(p+q)/n + \log(\text{RSS}/n)$, BIC $= (p+q)\log(n)/n + \log(\text{RSS}/n)$, and $\text{RSS} = \sum_{i=1}^{n} \hat{M}_i^*(\tau)^2$ with $\hat{M}_i^*(t)$ defined in Section 3. The results are presented in Table 7; they indicate that under AIC and BIC, model AMM1 is among the best and should be recommended. In addition, the performance of model AMM1 is close to that of model MM, without a significant improvement of fit. Thus, the simple multiplicative model MM is also recommended. In spite of the similar results produced by AMM1 and MM in the present application, these two models reveal different insights in general. AMM1 incorporates both additive and multiplicative forms of covaraites, and is therefore capable of identifying multiplicative and additive covariate effects simultaneously. Instead, MM accommodates all covariates in a multiplicative form, thereby interpreting each of the covariates in a similar manner.

7. Concluding Remarks

In this article, we have proposed a joint model for analyzing marker data contingent on recurrent event via latent variables when there exists an informative terminal event. The proposed joint model is flexible and robust in that the association among the marker process, the recurrent event process, and the terminal event is modeled nonparametrically. The additive-multiplicative mean model allows for both additive and multiplicative covariate effects. An estimating procedure was proposed to yield consistent and asymptotically normal estimators. Simulation results indicated that the proposed methods perform well, and an application to a medical cost study was illustrated.

The proposed joint model assumed that the covariates are time-independent but, following Zeng and Cai (2010), the proposed method can be extended in a straightforward manner to deal with time-dependent covariates. In particular, let $Z(t) = (X(t)', W(t)')'$, $\mathcal{Z}(t) = \{Z(s), 0 \le s \le t\}$. Then models (2.1), (2.2), and (2.3) are

$$E\{m(t)|N(t)-N(t-)=1, D\ge t, \mathcal{Z}(t), v_1, v_2\} = \alpha_0(t;v_1)g_\beta(\beta_0'X(t))+g_\zeta(\zeta_0'W(t)),$$
$$E\{dN(t)|D\ge t, \mathcal{Z}(t), v_1, v_2\} = g_\gamma(\gamma_0'Z(t))d\mu_0(t;v_2),$$
$$\log\int_0^D \exp\{\eta_0'Z(s)\}\Lambda_0(s) = \varepsilon,$$

respectively, where ε is independent of $Z(\cdot)$ with the extreme-value distribution. Furthermore, here the proportional hazards model was used for the terminal event. Other competing models, such as the additive hazards model, the accelerated failure time model, and a general transformation model can be also used.

One limitation of the proposed approach is that all covariate effects are assumed to be linear. This assumption may be too restrictive. It would be desirable to provide a more general joint model that could accommodate both linear and nonlinear covariate effects. The proposed method involves two weight functions $W(t)$ and $Q(t)$, and it turns out to be very difficult to derive optimal weights without specification of the covariance function of the marker and recurrent event processes (Sun et al. (2012)). Further research is needed to develop a simple and more efficient inference procedure.

In model (2.1), there is a problem with the choice of X and W. In practice, if the investigator is interested in studying the mean ratio of some covariates, then those covariates should be included in the multiplicative part. If the investigator is interested in studying the mean difference of some other covariates, then such covariates should be included in the additive part. Based on biological and possibly empirical grounds, the covariates anticipated to have a large impact in mean ratios should be taken as X, and those which could have a large impact in mean differences should be taken as W. If the covariates are of small dimension as in the data application, we can identify multiplicative and additive covariate effects by fitting different models, and use AIC and BIC as model selection criterion. When the biological process is not clear, it would be desirable to develop some data-driven methods for the classification of covariates.

We have assumed that X and W do not have any common factor for identifiability reasons. In the application, we considered eight models by classifying the covariates into either additive or multiplicative components and then compared the models by AIC/BIC. Our proposed estimation procedure cannot be extended in a straightforward manner to deal with the situation where X and W have some common factors. There a different estimating procedure is needed for parameter estimation. Then, as suggested by one referee, an alternative method could put each of the covariates into both components and estimate a general model, then test whether the additive/multiplicative effect is significant or test at least one of additive/multiplciative effects is significant. This results in a p-value for the effect of the covariate relevant in practice. For future research, it would be interesting to see how it compares to the current approach.

Supplementary Materials

Supplementary material available at *Statistica Sinica* online includes the regularity conditions (C1)−(C6) and the proofs of Theorems 1−3.

Acknowledgement

The authors thank the Co-Editor, Zhiliang Ying, an associate editor, and two referees for their insightful comments and suggestions that greatly improved the article. This research was partly supported by the National Natural Science Foundation of China Grants (No. 11231010, 11171330 and 11471277), Key Laboratory of RCSDS, CAS (No. 2008DP173182), BCMIIS, IRTSHUFE, GRF 14601115 from the Research Grant Council of the Hong Kong Special Administration Region, and AHRQ R01 HS 020263.

References

Cai, J., Zeng, D. and Pan, W. (2010). Semiparametric proportional means model for marker data contingent on recurrent event. *Lifetime Data Anal.* **16**, 250-270.

Han, M., Song, X., Sun, L. and Liu, L. (2014). Joint modeling of longitudinal data with informative observation times and dropouts. *Statist. Sinica* **24**, 1487-1504.

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

Herring, A. H. and Yang, J. (2007). Bayesian modeling of multiple episode occurrence and severity with a terminating event. *Biometrics* **63**, 381-388.

Hu, X., Sun, J. and Wei, L. J. (2003). Regression parameter estimation from panel counts. *Scand. J. Statist.* **30**, 25-43.

Liang, K. Y. and Zeger, S. L. (1986). Longitudinal data analysis using generalized linear models. *Biometrika* **73**, 13-22.

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

Lin, D. Y., Fleming, T. R. and Wei, L. J. (1994). Confidence bands for survival curves under the proportional hazards model. *Biometrika* **81**, 73-81.

Lin, D. Y., Wei, L. J. and Ying, Z. (1993). Checking the Cox model with cumulative sums of martingale-based residuals. *Biometrika* **80**, 557-572.

Lin, D. Y. and Ying, Z. (2001). Semiparametric and nonparametric regression analysis of longitudinal data. *J. Amer. Statist. Assoc.* **96**, 103-126.

Liu, M. and Ying, Z. (2007). Joint analysis of longitudinal data with informative right censoring. *Biometrics* **63**, 363-371.

Liu, L., Huang, X. and O'Quigley, J. (2008). 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.

Pan, Z. and Lin, D. Y. (2005). Goodness-of-fit methods for generalized linear mixed models. *Biometrics* **61**, 1000-1009.

Ratcliffe, S. J., Guo, W. and Ten Have, T. R. (2004). Joint modeling of longitudinal and survival data via a common frailty. *Biometrics* **60**, 892-899.

Schoenfeld, D. (1982). Partial residuals for the proportional hazards regression model. *Biometrika* **69**, 239-241.

Sun, J., Sun, L. and Liu, D. (2007). Regression analysis of longitudinal data in the presence of informative observation and censoring times. *J. Amer. Statist. Assoc.* **102**, 1397-1406.

Sun, L., Song, X., Zhou, J. and Liu, L. (2012). Joint analysis of longitudinal data with informative observation times and a dependent terminal event. *J. Amer. Statist. Assoc.* **107**, 688-700.

Therneau, T. M., Grambsch, P. M. and Fleming, T. R. (1990). Martingale-based residuals for survival models. *Biometrika* **77**, 147-160.

Tsiatis, A. A., Degruttola, V. and Wulfsohn, M. S. (1995). Modeling the relationship of survival to longitudinal data measured with error. Applications to survival and CD4 counts in patients with AIDS. *J. Amer. Statist. Assoc.* **90**, 27-37.

Wu, M. C. and Bailey, K. R. (1989). Estimation and comparison of changes in the presence of informative right censoring: conditional linear model. *Biometrics* **45**, 939-955.

Xu, J. and Zeger, S. L. (2001). Joint analysis of longitudinal data comprising repeated measures and times to events. *Appl. Statist.* **50**, 375-387.

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

Zeng, D. and Cai, J. (2010). Semiparametric additive rate model for recurrent events with informative terminal event. *Biometrika* **97**, 669-712.

Zhou, J., Zhao, X. and Sun, L. (2013). A new inference approach for joint models of longitudinal data with informative observation and censoring times. *Statist. Sinica* **23**, 571-593.

School of Statistics and Management, Shanghai University of Finance and Economics, Shanghai, 200433, P.R. China.

E-mail: han.miao@mail.shufe.edu.cn

Department of Statistics, The Chinese University of Hong Kong.

E-mail: xysong@sta.cuhk.edu.hk

Institute of Applied Mathematics, Academy of Mathematics and Systems Science, Chinese Academy of Sciences, Beijing, 100190, P.R. China.

E-mail: slq@amt.ac.cn

Department of Preventive Medicine, Northwestern University, Chicago, IL 60611, USA.

E-mail: lei.liu@northwestern.edu

(Received August 2014; accepted October 2015)