Lifetime Data Anal
DOI 10.1007/s10985-016-9375-y

Joint analysis of panel count data with an informative observation process and a dependent terminal event

Jie Zhou[1] · **Haixiang Zhang**[2] · **Liuquan Sun**[3] · **Jianguo Sun**[4]

Received: 11 March 2015 / Accepted: 19 July 2016

Abstract Panel count data occur in many clinical and observational studies, and in many situations, the observation process may be informative and also there may exist a terminal event such as death which stops the follow-up. In this article, we propose a new joint model for the analysis of panel count data in the presence of both an informative observation process and a dependent terminal event via two latent variables. For the inference on the proposed models, a class of estimating equations is developed and the resulting estimators are shown to be consistent and asymptotically normal. In addition, a lack-of-fit test is provided for assessing the adequacy of the models. Simulation studies suggest that the proposed approach works well for practical situations. A real example from a bladder cancer clinical trial is used to illustrate the proposed methods.

Keywords Estimating equation · Informative observation process · Joint modeling · Panel count data · Terminal event

1 Introduction

Panel count data usually occur in longitudinal follow-up studies that concern occurrence rates of certain recurrent events. This kind of data usually arise from event history studies that concern some recurrent events and in which subjects are moni-

✉ Haixiang Zhang
zhx_math@163.com

1 School of Mathematical Sciences, Capital Normal University, Beijing 100048, China

2 Center for Applied Mathematics, Tianjin University, Tianjin 300072, China

3 Institute of Applied Mathematics, Chinese Academy of Sciences, Beijing 100190, China

4 Department of Statistics, University of Missouri, Columbia, MO 65211, USA

Published online: 23 July 2016

tored or observed only at discrete time points instead of continuously. The fields in which one often sees such data include demographical and epidemiological studies, medical researches, reliability experiments, tumorgenicity experiments and sociological studies (Kalbfleisch and Lawless 1985; Thall and Lachin 1988; Sun 2006).

Many authors have investigated the analysis of panel count data. For example, Sun and Kalbfleisch (1995) considered the estimation of the mean function of the underlying point process that yields panel count data. Balakrishnan and Zhao (2009, 2010, 2011), Park et al. (2007), Sun and Fang (2003) and Zhao and Sun (2011) presented some nonparametric test procedures for the comparison of the mean functions of counting processes based on panel count data. Hu et al. (2003) and Sun and Wei (2000) developed some estimating equation-based methods for regression analysis of panel count data. Wellner and Zhang (2007) and Zhang (2002) also discussed regression analysis of panel count data and gave some likelihood-based approaches. Furthermore, Huang et al. (2006) and Sun et al. (2007) considered regression analysis of panel count data with dependent observation times. Li et al. (2010) proposed a class of semiparametric transformation models for panel count data with a dependent observation process. Zhang et al. (2013a) presented a robust joint model for multivariate panel count data via latent variables. Tong et al. (2009) and Zhang et al. (2013b) considered the variable selection issues on panel count data. A relatively complete references on panel count data can be found in Sun and Zhao (2013).

Most of the existing methods for panel count data assume that there is no terminal event and the observation process is independent of the underlying recurrent event process unconditionally or conditional on the covariates. In many situations, however, the follow-up of subjects could be stopped by a terminal event, such as death, which precludes further recurrent events. For example, in a tumourigenicity study, tumours would not develop after death. Furthermore, it is often the case that the terminal event is strongly correlated with the recurrent events of interest as well as the observation process. In the presence of a terminal event, there exists considerable work on the analysis of recurrent event data and longitudinal data analysis, and two approaches are commonly adopted. One is the marginal model approach (Cook and Lawless 1997; Ghosh and Lin 2002; Zhao et al. 2011), and the other is the frailty model approach (Huang and Wang 2004; Liu et al. 2004; Ye et al. 2007; Zeng and Cai 2010; Sun et al. 2012).

Although panel count data can be regarded as a special case of longitudinal data, the methods developed for longitudinal data are usually less efficient than the methods specifically developed for panel count data since the former does not take into account the special data structures of panel count data. To deal with these problems, Sun and Zhao (2013) and Zhao et al. (2013) described a marginal modeling approach that leaves the correlation between the recurrent event and the terminal event unspecified. They made use of the inverse probability weighting technique to take into account the fact that the subjects who are terminated cannot experience further occurrences of the events of interest. Sun et al. (2007) presented some shared frailty models to analyze panel count data when the response process tends to be associated with the observation process and the follow-up process. Their method requires the assumptions that one random effect is normally distributed and the observation process is a nonhomogeneous Poisson process. In practice, it is impossible to know which distribution the random effects follow, and misspecifying the distribution of random effect often leads to erroneous inference. Recently, Sun et al. (2012) considered an additive model for the analysis of longitudinal data. In this article, we will extend Sun et al. (2012)'s method to panel count data with a proportional means model. We propose a joint modeling approach to model the panel count data in the presence of both informative observation times and a dependent terminal event via two latent variables. This joint model is flexible and robust in that the distributions of the two latent variables and the dependence structures are left unspecified.

The rest of this paper is organized as follows. In Sect. 2, we introduce some notation and describe the proposed models that will be used throughout the paper. Specifically, we will describe the joint model connected through two latent variables for the recurrent event process, the observation times, and the terminal event. In Sect. 3, an estimating equation approach is developed for estimation of the regression parameters of interest. Also we establish the asymptotic properties of the proposed estimates. In Sect. 4, we develop a technique for checking the adequacy of the proposed models. Section 5 reports some results from the simulation studies conducted for evaluating the proposed methods. In Sect. 6, we apply the proposed methods to a bladder cancer study and some concluding remarks are provided in Sect. 7. Details of the proof are given in the Appendix.

2 Notation and models

Consider a recurrent event study and let $N(t)$ denote the number of the occurrences of the recurrent event of interest up to time t, $0 \leq t \leq \tau$, where τ is the longest follow-up time. For each subject, suppose that a $d \times 1$ vector of covariates X is observed. Let D be the time of the terminal event such as death, and C be the censoring time. Define $T = C \wedge D$ and $\delta = I(D \leq C)$, where $a \wedge b = \min(a, b)$ and $I(\cdot)$ is the indicator function. Let u and v be two latent variables which are independent of X. For any time t, suppose that given (u, v, X) and $D \geq t$, the mean function of $N(t)$ has the form,

$$E\{N(t)|X, D \geq t, u, v\} = \mu_N(t; u)\exp(X'\beta_0), \tag{1}$$

where $\mu_N(t; u)$ is an unknown baseline mean function and β_0 is a vector of unknown regression parameters.

Let $H(t)$ denote the observation process that yields the observation time points for $N(t)$ and assume that $H(t)$ is independent of $N(t)$ conditional on (u, v, X) and $D \geq t$. Also assume that $H(t)$ follows the rate model

$$E\{dH(t)|X, D \geq t, u, v\} = \exp(X'\gamma_0)d\mu_H(t; v), \tag{2}$$

where γ_0 is a vector of unknown regression parameters, and $\mu_H(t; v)$ is an unknown baseline mean function with $\mu_H(0; v) = 0$. Note that in the two models above, the recurrent event and observation processes are related with each other and to the terminal event through the latent variables u and v. The condition $D \geq t$ is used because it is often of interest in many studies to make inference for subjects who are currently alive (Ye et al. 2007; Zeng and Cai 2010; Zhao et al. 2011). For the terminal event, we assume that it follows the semiparametric Cox model

$$\log \Lambda_0(D) = -X'\eta_0 + \epsilon, \tag{3}$$

where η_0 is a vector of unknown regression parameters, $\Lambda_0(t)$ is an unspecified baseline cumulative hazard function, and ϵ is a random error with extreme-value distribution.

Note that the three models above are constructed through random effects u, v and random error ϵ, respectively. We use the joint distribution of the three variables to describe the association among the three models. In the following, the joint distribution of u, v and ϵ will be left unspecified. Hence the joint models (1), (2) and (3) are extensive in the sense that $\mu_N(t; u)$ and $\mu_H(t; v)$ are both nonparametric and depend on the latent random variables in an arbitrary way. By letting $u = v$, $\mu_N(t; u) = \mu_N(t)u^{\alpha}$, and $\mu_H(t; v) = v\mu_H(t)$ with α being an unknown parameter, the joint models above reduce to the models proposed in Sun et al. (2007) assuming that their restrictions are satisfied. In what follows, we will assume that given X, the censoring time C is independent of $\{u, v, D, N(\cdot), H(\cdot)\}$. For a sample of n independent subjects, the observed data consist of $\{N_i(t)dH_i(t), T_i, \delta_i, X_i, H_i(t), 0 \le t \le T_i, i = 1, \ldots, n\}$.

3 Inference procedure

Now we discuss the estimation of the parameters β_0 and γ_0 as well as others. For this, note that D may be censored and the latent variables u and v are unobservable and thus it is impossible to make inference for the parameters of interest directly. To overcome this problem, we will first consider the observed mean function given the observed endpoint T, in which the resulting nonparametric component depends on the latent variables. Then we will derive an expression of the nonparametric component which can be estimated using the observed data for given β and γ. The latent variables will disappear from the resulting estimator. To be more specific, let $\mathcal{A}_0(t; u, v) = \int_0^t \mu_N(z; u)d\mu_H(z; v)$ and define $dR(t, s) = E\{d\mathcal{A}_0(t; u, v)|\epsilon \ge s\}$. Following the assumption that (u, v, ϵ) is independent of (X, C), we obtain that

$$E\{N(t)dH(t)|X, T \ge t\} = \exp\left\{X'(\beta_0 + \gamma_0)\right\} dR(t, \log \Lambda_0(t) + X'\eta_0), \tag{4}$$

and

$$dR(t, s) = \frac{E\left[N(t)dH(t)I\left\{\log \Lambda_0(T) + X'\eta_0 \ge s \ge \log \Lambda_0(t) + X'\eta_0\right\}\right]}{E\left[\exp\{X'(\beta_0 + \gamma_0)\}I\{\log \Lambda_0(T) + X'\eta_0 \ge s \ge \log \Lambda_0(t) + X'\eta_0\}\right]}. \tag{5}$$

Here $\log \Lambda_0(T) + X'\eta_0 \ge s \ge \log \Lambda_0(t) + X'\eta_0$ implies $T \ge t$. The derivation of (4) and (5) is given in the Appendix, where the key point is that $\mathcal{A}_0(t; u, v)$ is independent of X.

To obtain an estimate of $dR(t, s)$, we need to obtain the estimates of η_0 and $\Lambda_0(t)$ from model (3). For this, according to Fleming and Harrington (1991), we can use the maximum partial likelihood estimator $\hat{\eta}$ and the Breslow estimator $\hat{\Lambda}_0(t)$. Hence for given β and γ, we have

$$d\hat{R}(t, s; \beta, \gamma) = \frac{\sum_{i=1}^{n} N_i(t) dH_i(t) I \left\{ \log \hat{\Lambda}_0(T_i) + X_i' \hat{\eta} \geq s \geq \log \hat{\Lambda}_0(t) + X_i' \hat{\eta} \right\}}{\sum_{j=1}^{n} \exp \left\{ X_j'(\beta + \gamma) \right\} I \left\{ \log \hat{\Lambda}_0(T_j) + X_j' \hat{\eta} \geq s \geq \log \hat{\Lambda}_0(t) + X_j' \hat{\eta} \right\}}.$$

For given γ, to estimate β_0, motivated by (4) and the generalized estimating equation approach (Liang and Zeger 1986), we propose the following estimating function for β_0,

$$U(\beta; \gamma) = \sum_{i=1}^{n} \int_0^{\tau} W(t) \left\{ X_i - \bar{X}_i(t; \beta, \gamma) \right\} \Delta_i(t) \Bigg[N_i(t) dH_i(t) - \exp \left\{ X_i'(\beta + \gamma) \right\} \times \frac{\sum_{j=1}^{n} N_j(t) dH_j(t) \hat{\Phi}_j(t, X_i)}{\sum_{j=1}^{n} \exp \left\{ X_j'(\beta + \gamma) \right\} \hat{\Phi}_j(t, X_i)} \Bigg], \tag{6}$$

where $\Delta_i(t) = I(T_i \geq t)$ and $W(t)$ is a possibly data-dependent weight function,

$$\hat{\Phi}_j(t, X_i) = I \left\{ \log \hat{\Lambda}_0(T_j) + X_j' \hat{\eta} \geq \log \hat{\Lambda}_0(t) + X_i' \hat{\eta} \geq \log \hat{\Lambda}_0(t) + X_j' \hat{\eta} \right\},$$

$$\bar{X}_i(t; \beta, \gamma) = \frac{\sum_{j=1}^{n} X_j \exp \left\{ X_j'(\beta + \gamma) \right\} \hat{\Phi}_j(t, X_i)}{\sum_{j=1}^{n} \exp \left\{ X_j'(\beta + \gamma) \right\} \hat{\Phi}_j(t, X_i)}.$$

Of course in reality, γ_0 is unknown. We can employ the estimation equation-based method which was also studied in Zeng and Cai (2010) to estimate γ_0,

$$\tilde{U}(\gamma) = \sum_{i=1}^{n} \int_0^{\tau} Q(t) \left\{ X_i - \bar{X}_i^*(t; \gamma) \right\} \Delta_i(t) \left\{ dH_i(t) - \exp(X_i' \gamma) d\bar{H}_i(t; \gamma) \right\} = 0, \tag{7}$$

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

$$\bar{X}_i^*(t; \gamma) = \frac{\sum_{j=1}^{n} X_j \exp(X_j' \gamma) \hat{\Phi}_j(t, X_i)}{\sum_{j=1}^{n} \exp(X_j' \gamma) \hat{\Phi}_j(t, X_i)}, \quad d\bar{H}_i(t; \gamma) = \frac{\sum_{j=1}^{n} dH_j(t) \hat{\Phi}_j(t, X_i)}{\sum_{j=1}^{n} \exp(X_j' \gamma) \hat{\Phi}_j(t, X_i)}.$$

Denote $\hat{\gamma}$ as the solution to the estimating Eq. (7), and $\hat{\beta}$ as the solution to $U(\beta; \hat{\gamma}) = 0$. By the law of large numbers and the consistency of $\hat{\eta}$ and $\hat{\Lambda}_0(t)$, we can show that

$\hat{\beta}$ and $\hat{\gamma}$ are consistent. The following theorem gives the asymptotic normality of $\hat{\beta}$ and $\hat{\gamma}$ with the proof sketched in the Appendix.

Theorem 1 *Assume that the regularity conditions C.1–C.3 stated in the Appendix hold. Then $n^{1/2}(\hat{\beta} - \beta_0)$ and $n^{1/2}(\hat{\gamma} - \gamma_0)$ have an asymptotic multivariate normal distribution with mean zero and covariance matrix $A^{-1}\Sigma(A^{-1})'$, where A and Σ are defined in the Appendix.*

To apply the results above, it is apparent that we need to estimate the asymptotic covariance matrix or A and Σ. For A, a simple consistent estimate is given by

$$\hat{A} = -n^{-1}\begin{pmatrix} \partial U(\hat{\beta};\hat{\gamma})/\partial\beta & \partial U(\hat{\beta};\hat{\gamma})/\partial\gamma \\ 0 & \partial\tilde{U}(\hat{\gamma})/\partial\gamma \end{pmatrix}.$$

However, it would be very difficult to estimate Σ as it is complicated and involves the Hadamard derivatives of $d\bar{M}_0(t, X; \eta, \Lambda)$ and $d\bar{H}_0(t, X; \eta, \Lambda)$, defined in the Appendix, with respect to Λ. Corresponding to this, we propose to use the following Monte Carlo method. Specifically, based on the proof of Theorem 1, we know that the variation of $U(\beta_0; \gamma_0)$ comes from three sources, $dM_i(t) - \exp\{X_i'(\beta_0 + \gamma_0)\}d\bar{M}(t, X_i; \hat{\eta}, \hat{\Lambda}_0)$, the empirical summation in the numerator and denominator of $\bar{M}(t, X_i; \hat{\eta}, \hat{\Lambda}_0)$, and the plug-in estimator $(\hat{\eta}, \hat{\Lambda}_0)$, where $dM_i(t)$ and $d\bar{M}(t, X; \eta, \Lambda)$ are defined in the Appendix. To describe the resampling approach to capture all these variations, let $\mathcal{Z}_1, \ldots, \mathcal{Z}_n$ denote the independent and identically distributed sample of size n from the standard normal distribution. Then the first two sources of variations can be estimated by

$$\Omega_1^* = \sum_{i=1}^{n} \mathcal{Z}_i \int_0^{\tau} W(t)\left\{X_i - \bar{X}_i(t;\hat{\beta},\hat{\gamma})\right\}\Delta_i(t)\left[N_i(t)\,dH_i(t) - \exp\left\{X_i'\left(\hat{\beta}+\hat{\gamma}\right)\right\}\right.$$
$$\left.\times\frac{\sum_{j=1}^{n} N_j(t)dH_j(t)\hat{\Phi}_j(t, X_i)}{\sum_{j=1}^{n}\exp\left\{X_j'\left(\hat{\beta}+\hat{\gamma}\right)\right\}\hat{\Phi}_j(t, X_i)}\right],$$

and

$$\Omega_2^* = \sum_{i=1}^{n} \int_0^{\tau} W(t)\left\{X_i - \bar{X}_i(t;\hat{\beta},\hat{\gamma})\right\}\Delta_i(t)\exp\left\{X_i'\left(\hat{\beta}+\hat{\gamma}\right)\right\}$$
$$\times\left[-\frac{\sum_{j=1}^{n}\mathcal{Z}_j N_j(t)dH_j(t)\hat{\Phi}_j(t, X_i)}{\sum_{j=1}^{n}\exp\left\{X_j'\left(\hat{\beta}+\hat{\gamma}\right)\right\}\hat{\Phi}_j(t, X_i)}\right.$$

$$+\frac{\sum_{j=1}^{n} N_j(t)dH_j(t)\hat{\Phi}_j(t,X_i)}{\left[\sum_{j=1}^{n}\exp\left\{X_j'(\hat{\beta}+\hat{\gamma})\right\}\hat{\Phi}_j(t,X_i)\right]^2}\sum_{j=1}^{n}\mathcal{Z}_j\exp\left\{X_j'\left(\hat{\beta}+\hat{\gamma}\right)\right\}\hat{\Phi}_j(t,X_i)\Bigg].$$

For estimation of the third part, define

$$\hat{\eta}^* = \hat{\eta} + \hat{\Omega}^{-1}n^{-1}\sum_{i=1}^{n}\mathcal{Z}_i\int_0^{\tau}\left\{X_i - \bar{X}^D(t,\hat{\eta})\right\}d\hat{M}_i^D(t),$$

and

$$\hat{\Lambda}_0^*(t) = \hat{\Lambda}_0(t) + n^{-1}\sum_{i=1}^{n}\mathcal{Z}_i\int_0^t\frac{d\hat{M}_i^D(z)}{S^{(0)}(z,\hat{\eta})} - \int_0^t\bar{X}^D(z,\hat{\eta})'d\hat{\Lambda}_0(z)\left(\hat{\eta}^*-\hat{\eta}\right),$$

where $H_i^D(t) = I(T_i \le t, \delta_i = 1)$, $\hat{M}_i^D(t) = H_i^D(t) - \int_0^t \Delta_i(z)\exp(X_i'\hat{\eta})d\hat{\Lambda}_0(z)$, $S^{(0)}(t,\eta) = n^{-1}\sum_{i=1}^{n}\Delta_i(t)\exp(X_i'\eta)$, $S^{(1)}(t,\eta) = n^{-1}\sum_{i=1}^{n}\Delta_i(t)X_i\exp(X_i'\eta)$, $\bar{X}^D(t,\hat{\eta}) = S^{(1)}(t,\hat{\eta})/S^{(0)}(t,\hat{\eta})$, $S^{(2)}(t,\eta) = n^{-1}\sum_{i=1}^{n}\Delta_i(t)X_i^{\otimes 2}\exp(X_i'\eta)$,

$$\hat{\Omega} = n^{-1}\sum_{i=1}^{n}\int_0^{\tau}\left[\frac{S^{(2)}(t,\hat{\eta})}{S^{(0)}(t,\hat{\eta})} - \left\{\frac{S^{(1)}(t,\hat{\eta})}{S^{(0)}(t,\hat{\eta})}\right\}^{\otimes 2}\right]dH_i^D(t),$$

and for a vector a, $a^{\otimes 2} = aa'$. Then the variation due to $(\hat{\eta}, \hat{\Lambda}_0)$ can be estimated by

$$\begin{aligned}\Omega_3^* = \sum_{i=1}^{n}\int_0^{\tau} W(t)\left\{X_i - \bar{X}_i\left(t;\hat{\beta},\hat{\gamma}\right)\right\}\Delta_i(t)\exp\left\{X_i'\left(\hat{\beta}+\hat{\gamma}\right)\right\}\\ \times\left[\frac{\sum_{j=1}^{n}N_j(t)dH_j(t)\hat{\Phi}_j(t,X_i)}{\sum_{j=1}^{n}\exp\left\{X_j'\left(\hat{\beta}+\hat{\gamma}\right)\right\}\hat{\Phi}_j(t,X_i)} - \frac{\sum_{j=1}^{n}N_j(t)dH_j(t)\hat{\Phi}_j^*(t,X_i)}{\sum_{j=1}^{n}\exp\left\{X_j'\left(\hat{\beta}+\hat{\gamma}\right)\right\}\hat{\Phi}_j^*(t,X_i)}\right],\end{aligned}$$

where $\hat{\Phi}_j^*(t,X_i)$ is defined the same way as $\hat{\Phi}_j(t,X_i)$ except that $(\hat{\eta},\hat{\Lambda}_0)$ is replaced with $(\hat{\eta}^*,\hat{\Lambda}_0^*)$. Similarly, we can estimate the variation of $\tilde{U}(\gamma_0)$ by the following three quantities

$$\Omega_4^* = \sum_{i=1}^{n}\mathcal{Z}_i\int_0^{\tau}Q(t)\left\{X_i - \bar{X}_i^*(t;\hat{\gamma})\right\}\Delta_i(t)\times\left\{dH_i(t) - \exp(X_i'\hat{\gamma})d\bar{H}_i(t;\hat{\gamma})\right\},$$

$$\Omega_5^* = \sum_{i=1}^{n} \int_0^{\tau} Q(t) \left\{ X_i - \bar{X}_i^*(t;\hat{\gamma}) \right\} \Delta_i(t) \exp(X_i'\hat{\gamma}) \left[-\frac{\sum_{j=1}^{n} \mathcal{Z}_j dH_j(t) \hat{\Phi}_j(t, X_i)}{\sum_{j=1}^{n} \exp(X_j'\hat{\gamma}) \hat{\Phi}_j(t, X_i)} \right.$$

$$\left. + \frac{\sum_{j=1}^{n} dH_j(t) \hat{\Phi}_j(t, X_i)}{\left[\sum_{j=1}^{n} \exp(X_j'\hat{\gamma}) \hat{\Phi}_j(t, X_i) \right]^2} \sum_{j=1}^{n} \mathcal{Z}_j \exp(X_j'\hat{\gamma}) \hat{\Phi}_j(t, X_i) \right],$$

and

$$\Omega_6^* = \sum_{i=1}^{n} \int_0^{\tau} Q(t) \left\{ X_i - \bar{X}_i^*(t;\hat{\gamma}) \right\} \Delta_i(t) \exp(X_i'\hat{\gamma}) \left[d\bar{H}_i(t;\hat{\gamma}) - d\bar{H}_i^*(t;\hat{\gamma}) \right],$$

where $\bar{H}_i^*(t;\hat{\gamma})$ is defined the same way as $\bar{H}_i(t;\hat{\gamma})$ except that $(\hat{\eta}, \hat{\Lambda}_0)$ is replaced with $(\hat{\eta}^*, \hat{\Lambda}_0^*)$. Define $\hat{\Upsilon} = (\hat{\Upsilon}_1', \hat{\Upsilon}_2')'$, where $\hat{\Upsilon}_1 = n^{-1/2}(\Omega_1^* + \Omega_2^* + \Omega_3^*)$ and $\hat{\Upsilon}_2 = n^{-1/2}(\Omega_4^* + \Omega_5^* + \Omega_6^*)$. Given the observed data, we can estimate Σ by the empirical covariance matrix of $\hat{\Upsilon}$ with repeatedly generating the random samples $(\mathcal{Z}_1, \ldots, \mathcal{Z}_n)$. The following theorem justifies the above Monte Carlo method with the proof given in the Appendix.

Theorem 2 *Let $E_{\mathcal{Z}}$ denotes the conditional expectation with respect to $(\mathcal{Z}_1, \ldots, \mathcal{Z}_n)$ given the observed data. Then $E_{\mathcal{Z}}(\hat{\Upsilon}^{\otimes 2}) \xrightarrow{P} \Sigma$, where $\xrightarrow{P}$ denotes convergence in probability.*

4 Model diagnostics

In this section, we will propose some graphical and numerical procedures for checking the adequacy of the proposed models with focus on model (1) as some procedures can be found in the literature for assessing models (2) and (3), respectively. Following Lin et al. (2000), we propose the following cumulative sums of residual,

$$\mathcal{F}(t, x) = n^{-1/2} \sum_{i=1}^{n} \int_0^{t} I(X_i \le x) d\hat{M}_i^*(z), \tag{8}$$

where $I(X_i \le x)$ means that each of the components of X_i is no larger than the corresponding component of x, and

$$d\hat{M}_i^*(t) = \Delta_i(t) \left[N_i(t) dH_i(t) - \exp\left\{ X_i' \left(\hat{\beta} + \hat{\gamma} \right) \right\} \times \frac{\sum_{j=1}^{n} N_j(t) dH_j(t) \hat{\Phi}_j(t, X_i)}{\sum_{j=1}^{n} \exp\left\{ X_j' \left(\hat{\beta} + \hat{\gamma} \right) \right\} \hat{\Phi}_j(t, X_i)} \right].$$

Here the null hypothesis H_0 is defined as the correct specification of model (1). Similarly to $U(\beta_0; \gamma_0)$, the variation of $\mathcal{F}(t, x)$ can be characterized by the following three terms:

$$\Omega_7^*(t,x) = \sum_{i=1}^n \mathcal{Z}_i \int_0^t I(X_i \le x)\Delta_i(z)\Bigg[N_i(z)dH_i(z) - \exp\left\{X_i'\left(\hat\beta+\hat\gamma\right)\right\} \times \frac{\sum_{j=1}^n N_j(z)dH_j(z)\hat\Phi_j(z,X_i)}{\sum_{j=1}^n \exp\left\{X_j'\left(\hat\beta+\hat\gamma\right)\right\}\hat\Phi_j(z,X_i)}\Bigg],$$

$$\Omega_8^*(t,x) = \sum_{i=1}^n \int_0^t I(X_i \le x)\Delta_i(z)\exp\left\{X_i'\left(\hat\beta+\hat\gamma\right)\right\} \times \Bigg[-\frac{\sum_{j=1}^n \mathcal{Z}_j N_j(z)dH_j(z)\hat\Phi_j(z,X_i)}{\sum_{j=1}^n \exp\left\{X_j'\left(\hat\beta+\hat\gamma\right)\right\}\hat\Phi_j(z,X_i)} + \frac{\sum_{j=1}^n N_j(z)dH_j(z)\hat\Phi_j(z,X_i)}{\left[\sum_{j=1}^n \exp\left\{X_j'\left(\hat\beta+\hat\gamma\right)\right\}\hat\Phi_j(z,X_i)\right]^2}\sum_{j=1}^n \mathcal{Z}_j \exp\left\{X_j'\left(\hat\beta+\hat\gamma\right)\right\}\hat\Phi_j(z,X_i)\Bigg],$$

and

$$\Omega_9^*(t,x) = \sum_{i=1}^n \int_0^t I(X_i \le x)\Delta_i(z)\exp\left\{X_i'\left(\hat\beta+\hat\gamma\right)\right\} \times \Bigg[\frac{\sum_{j=1}^n N_j(z)dH_j(z)\hat\Phi_j(z,X_i)}{\sum_{j=1}^n \exp\left\{X_j'\left(\hat\beta+\hat\gamma\right)\right\}\hat\Phi_j(z,X_i)} - \frac{\sum_{j=1}^n N_j(z)dH_j(z)\hat\Phi_j^*(z,X_i)}{\sum_{j=1}^n \exp\left\{X_j'\left(\hat\beta+\hat\gamma\right)\right\}\hat\Phi_j^*(z,X_i)}\Bigg].$$

Define

$$\hat\Gamma_1(t,x) = n^{-1}\sum_{i=1}^n \int_0^t I(X_i \le x)\,\triangle_i\,(z)\left\{X_i - \bar X_i\left(z;\hat\beta,\hat\gamma\right)\right\}\exp\left\{X_i'\left(\hat\beta+\hat\gamma\right)\right\} \times \frac{\sum_{j=1}^n N_j(z)dH_j(z)\hat\Phi_j(z,X_i)}{\sum_{j=1}^n \exp\left\{X_j'\left(\hat\beta+\hat\gamma\right)\right\}\hat\Phi_j(z,X_i)},$$

and $\hat{\Gamma}(t,x) = (\hat{\Gamma}_1(t,x)', \hat{\Gamma}_1(t,x)')'$. Then the null distribution of $\mathcal{F}(t,x)$ can be obtained from the following theorem with the proof presented in the Appendix.

Theorem 3 *Suppose that the conditions in Theorem 1 hold. Then under H_0, the null distribution of $\mathcal{F}(t,x)$ can be approximated by the following zero-mean Gaussian process,*

$$\hat{\mathcal{F}}(t,x) = n^{-1/2}\left\{\Omega_7^*(t,x) + \Omega_8^*(t,x) + \Omega_9^*(t,x) - \hat{\Gamma}(t,x)'\hat{A}^{-1}\hat{\Upsilon}\right\}. \tag{9}$$

Based on Theorem 3, it is easy to see that we can obtain a large number of realizations of $\mathcal{F}(t,x)$ by repeatedly generating the standard normal random sample $(\mathcal{Z}_1,\ldots,\mathcal{Z}_n)$ while fixing the observed data. To assess the adequacy of model (1), a simple and graphical method is to plot these realizations of $\hat{\mathcal{F}}(t,x)$ along with the observed $\mathcal{F}(t,x)$ and examine any unusual pattern of $\mathcal{F}(t,x)$ with respect to $\hat{\mathcal{F}}(t,x)$. In particular, if the model is correct, the observed $\mathcal{F}(t,x)$ is expected to be surrounded by the realizations and some illustrations can be found in Lin et al. (2000) among others. More formally, we can apply the supremum test statistic $\sup_{t,x}|\mathcal{F}(t,x)|$ to conduct the lack-of-fit test and obtain the p-value by comparing the observed value of $\sup_{t,x}|\mathcal{F}(t,x)|$ to a large number of realizations from $\sup_{t,x}|\hat{\mathcal{F}}(t,x)|$.

5 A simulation study

In this section we will present some results from an extensive simulation study. In the study, the covariate X_i was generated from a Bernoulli distribution with success probability 0.5. The terminal event time was generated through $\log(D_i/4) = -\eta_0 X_i + \epsilon_i$, where ϵ_i was generated from the extreme-value distribution. The censoring time C_i was taken as $\min(C_i^*, \tau)$ with C_i^* following the uniform distribution over (2,10) and $\tau = 6$, which yielded 23 % censoring for the terminal event. Let $u_i = \exp(\phi_1\epsilon_i/5)$ and $v_i = \rho_i \exp(-\phi_2\epsilon_i/5)$, where $\phi_1 = -1$, 0 or 1, ϕ_2=−1, 0 or 1, and ρ_i followed the uniform distribution over (0.5, 1.5). Given X_i, u_i, v_i and $T_i = \min(C_i, D_i)$, we generated the observation process from the nonhomogeneous Poisson process with intensity function

$$\lambda_i(t) = v_i \exp(\gamma_0 X_i) I(T_i \geq t).$$

The average number of observations per subject is about 3.

Also in the study, the recurrent event process $N_i(t)$ was assumed to be the Poisson process with the intensity function

$$\lambda_i^*(t|X_i, u_i, v_i, \omega_i) = \omega_i u_i \exp(\beta_0 X_i),$$

where ω_i was an independent gamma random variable with mean 1 and variance σ^2. Specifically, let $(t_{i,1},\ldots,t_{i,K_i})$ be the observation times for the ith subject, then $N_i(t_{i,j})$, $j = 1,\ldots,K_i$ were generated piecewisely by generating $N_i(t_{i,j}) - N_i(t_{i,j-1})$ from a Poisson distribution with the mean functions $\omega_i(t_{i,j} - t_{i,j-1})u_i \exp(\beta_0 X_i)$,

Table 1 Simulation results for the estimation of $\beta_0 = -0.5$ and $\gamma_0 = 0.5$

n	(ϕ_1, ϕ_2)	β_0				γ_0			
		BIAS	SSE	ESE	CP	BIAS	SSE	ESE	CP
100	$(-1, -1)$	-0.0102	0.3433	0.3265	0.928	0.0480	0.1708	0.1714	0.943
	$(-1, 0)$	-0.0145	0.3419	0.3235	0.929	0.0537	0.1668	0.1718	0.944
	$(-1, 1)$	-0.0182	0.3202	0.3250	0.947	0.0478	0.1723	0.1755	0.950
	$(0, -1)$	-0.0080	0.3544	0.3305	0.921	0.0475	0.1741	0.1729	0.947
	$(0, 0)$	0.0122	0.3437	0.3263	0.937	0.0383	0.1724	0.1702	0.947
	$(0, 1)$	0.0146	0.3426	0.3227	0.942	0.0380	0.1727	0.1741	0.947
	$(1, -1)$	0.0085	0.3589	0.3314	0.928	0.0554	0.1697	0.1714	0.945
	$(1, 0)$	0.0103	0.3448	0.3293	0.938	0.0430	0.1662	0.1694	0.952
	$(1, 1)$	0.0112	0.3470	0.3297	0.928	0.0320	0.1694	0.1732	0.955
200	$(-1, -1)$	-0.0099	0.2311	0.2310	0.946	0.0290	0.1146	0.1166	0.950
	$(-1, 0)$	0.0136	0.2378	0.2299	0.950	0.0212	0.1147	0.1134	0.938
	$(-1, 1)$	-0.0083	0.2346	0.2295	0.936	0.0185	0.1175	0.1181	0.946
	$(0, -1)$	-0.0126	0.2414	0.2317	0.936	0.0234	0.1177	0.1161	0.948
	$(0, 0)$	0.0120	0.2389	0.2307	0.941	0.0193	0.1046	0.1142	0.964
	$(0, 1)$	-0.0025	0.2402	0.2294	0.937	0.0206	0.1199	0.1177	0.946
	$(1, -1)$	0.0271	0.2466	0.2357	0.936	0.0320	0.1139	0.1163	0.939
	$(1, 0)$	0.0079	0.2411	0.2348	0.938	0.0259	0.1095	0.1142	0.959
	$(1, 1)$	0.0198	0.2457	0.2349	0.938	0.0265	0.1202	0.1181	0.948

where $t_{i,0}$ was set to be 0. It is easy to verify that $N_i(t) = N_i(t \wedge D)$ and $N_i(t)$ satisfies (1). Note that ϕ_1 and ϕ_2 reflected the dependence among the recurrent event process, the observation times and the terminal event. For example, $\phi_1 = 0$ and $\phi_2 = 0$ implied that the recurrent event process, the observation times and the terminal event were independent, while $\phi_1 \neq 0$ and $\phi_2 \neq 0$ reflected that the three processes were related with each other. In the simulation study, we set $\eta_0 = 0.5$, $\beta_0 = 0.5$ and -0.5, $\gamma_0 = 0.5$, $W(t) = Q(t) = 1$. We found that 100 resamplings work well for the variance estimation. All the results reported below were computed based on 1000 replications with sample sizes $n = 100$ and 200.

Tables 1 and 2 report the simulation results on the estimates for β_0 and γ_0. The tables include the bias (BIAS) given by the difference of sample means of estimate and the true value, the sampling standard errors (SSE), the sampling means of the estimated standard errors (SEE), and the 95 % empirical coverage probabilities (CP). It can be seen from the tables that the estimators seem to be unbiased and the proposed variance estimation seems to work well. Also the coverage probabilities are reasonable and consistent with the normal levels.

For comparison, we also considered two other methods. One is given by Sun and Wei (2000), denoted by SW below, which assumed that given covariates, $N(\cdot)$, $H(\cdot)$ and D are independent, and the other is based on the generalized estimating equation method similar to (6) but ignoring the informative observation process and the terminal event, denoted by Naive below. Here we generated the data in the same way

Table 2 Simulation results for the estimation of $\beta_0 = 0.5$ and $\gamma_0 = 0.5$

n	(ϕ_1, ϕ_2)	β_0				γ_0			
		BIAS	SSE	ESE	CP	BIAS	SSE	ESE	CP
100	(−1, −1)	0.0720	0.3372	0.3381	0.947	0.0419	0.1682	0.1690	0.944
	(−1, 0)	0.0755	0.3484	0.3366	0.954	0.0336	0.1602	0.1684	0.951
	(−1, 1)	0.0803	0.3346	0.3398	0.955	0.0435	0.1736	0.1736	0.940
	(0, −1)	0.0729	0.3426	0.3335	0.944	0.0497	0.1640	0.1699	0.949
	(0, 0)	0.0825	0.3481	0.3357	0.945	0.0447	0.1581	0.1687	0.947
	(0, 1)	0.0714	0.3423	0.3400	0.962	0.0504	0.1737	0.1756	0.945
	(1, −1)	0.0871	0.3557	0.3502	0.952	0.0480	0.1721	0.1700	0.937
	(1, 0)	0.1074	0.3465	0.3473	0.947	0.0469	0.1637	0.1685	0.948
	(1, 1)	0.0803	0.3530	0.3406	0.942	0.0446	0.1708	0.1735	0.952
200	(−1, −1)	0.0487	0.2248	0.2263	0.943	0.0227	0.1176	0.1158	0.937
	(−1, 0)	0.0426	0.2164	0.2230	0.953	0.0206	0.1151	0.1141	0.945
	(−1, 1)	0.0373	0.2178	0.2232	0.951	0.0213	0.1163	0.1177	0.950
	(0, −1)	0.0342	0.2220	0.2243	0.947	0.0312	0.1141	0.1154	0.943
	(0, 0)	0.0289	0.2296	0.2250	0.928	0.0188	0.1097	0.1133	0.958
	(0, 1)	0.0311	0.2311	0.2217	0.937	0.0118	0.1138	0.1183	0.965
	(1, −1)	0.0390	0.2376	0.2325	0.936	0.0232	0.1121	0.1151	0.947
	(1, 0)	0.0427	0.2354	0.2283	0.944	0.0114	0.1090	0.1131	0.959
	(1, 1)	0.0424	0.2266	0.2285	0.951	0.0193	0.1189	0.1172	0.935

as before except taking the parameter $\eta = 0$ and 0.5. Note that with $\eta = 0$ and $(\phi_1, \phi_2) = (0, 0)$, the model of Sun and Wei (2000) is satisfied. Table 3 gives the comparison results for estimation of β_0. As expected, the SW's method and the Naive method yielded consistent estimators when $\eta = 0$ with smaller biases. However, when $\eta = 0.5$, the SW's method and the Naive method resulted in biased estimates. In addition, the SW's method always gave the largest variance, which may be caused by the use of inefficient estimating equations. The variance of the proposed estimator is always comparable and seems to be more efficient and reliable than the other methods.

Furthermore, we conducted some simulation studies to investigate the performance of the model checking procedure described in Sect. 4. In the studies, the covariate X_i was taken as 0, 1, 2, 3 and 4 with equal probabilities, and we replaced $\eta_0 X_i$, $\gamma_0 X_i$ and $\beta_0 X_i$ with $\eta_0 X_i + 0.01 K_1 X_i^2$, $\gamma_0 X_i + 0.02 K_2 X_i^2$ and $\beta_0 X_i + 0.05 K_3 X_i^2$, respectively. The other setups were the same as before with $\eta_0 = \gamma_0 = 0.2$, $\beta_0 = -0.2$. We considered the null hypothesis H_0 as $(K_1, K_2, K_3) = (0, 0, 0)$. Table 4 reports the empirical sizes and powers of the proposed test at the significance level of 0.05 with $(\phi_1, \phi_2) = (-1, 1)$ and $n = 200$. Simulation results show that the proposed test is useful, but a little conservative. Specifically, when model (1) is correct, it will tend to accept the null hypothesis even though models (2) and (3) are misspecified. When model (1) is misspecified, the test has a reasonable power to detect deviations from the null hypothesis, especially as the value of K_3 increases. In addition, the simulation

Table 3 Comparison results on estimation of β_0 with $n = 200$

η_0	(ϕ_1, ϕ_2)	Proposed		SW		Naive	
		BIAS	SSE	BIAS	SSE	BIAS	SSE
0	(−1, −1)	−0.0029	0.2106	0.0050	0.2916	0.0019	0.2051
	(−1, 0)	−0.0090	0.2143	0.0005	0.2920	−0.0009	0.2062
	(−1, 1)	−0.0068	0.2065	−0.0052	0.2844	−0.0003	0.1984
	(0, −1)	−0.0006	0.2140	0.0176	0.2866	0.0065	0.2117
	(0, 0)	−0.0031	0.2083	0.0097	0.2875	0.0058	0.2008
	(0, 1)	−0.0151	0.2113	−0.0102	0.2960	−0.0073	0.2021
	(1, −1)	−0.0098	0.2247	0.0045	0.3110	−0.0004	0.2180
	(1, 0)	−0.0164	0.2123	−0.0033	0.2973	−0.0091	0.2069
	(1, 1)	−0.0088	0.2086	−0.0087	0.2844	−0.0012	0.1996
0.5	(−1, −1)	−0.0067	0.2333	−0.5461	0.3075	−0.0568	0.2214
	(−1, 0)	0.0012	0.2267	−0.5453	0.2944	−0.0591	0.2161
	(−1, 1)	−0.0046	0.2354	−0.5466	0.3023	−0.0818	0.2121
	(0, −1)	0.0068	0.2368	−0.4889	0.3296	0.0165	0.2285
	(0, 0)	0.0040	0.2302	−0.4761	0.3155	−0.0011	0.2287
	(0, 1)	−0.0080	0.2221	−0.4831	0.3050	−0.0194	0.2106
	(1, −1)	−0.0148	0.2452	−0.4818	0.3443	0.0452	0.2416
	(1, 0)	0.0054	0.2319	−0.4459	0.3278	0.0538	0.2315
	(1, 1)	0.0042	0.2406	−0.4405	0.3221	0.0460	0.2310

results on the estimates for β_0 are also reported in Table 4. The results indicate that the biases are mainly affected by K_3, and the biases are relatively small when $K_3 = 0$. However, additional simulation studies show that the biases could be large when the values of K_1 and K_2 increase (not reported here).

6 An application

In this section, we applied the proposed methods to the bladder cancer data discussed by Sun and Wei (2000) among others. The data arose from the study conducted by the Veterans Administration Cooperative Urological Research Group and in the study, the patients were randomly assigned to three treatment groups at the beginning of the study. The tumors were removed at the patients' clinical visits, and may occur recurrently. For each patient, the number of initial tumors before entering the study and the size of the largest initial tumor were measured as baseline covariates. In addition, the observed informations include the clinical visit times (in month) and the number of bladder tumors that occurred between clinical visits. As in Sun and Wei (2000), we will focus on 85 bladder cancer patients in the placebo (47) and the thiotepa (38) groups. Among them, 22 patients were terminated by death with 12 in the thiotepa group and 10 in the placebo group. In the analysis below, we will focus on the effects of the thiotepa treatment and the number of initial tumors on the tumor occurrence rate by taking into account both informative observation times and a dependent terminal event.

Table 4 Simulation results on the model test and estimation of β_0

$K3$	$K2$	$K1$	Size/power	BIAS	SSE	ESE	CP
0	0	0	0.030	−0.0029	0.0866	0.0815	0.934
	0	1	0.033	0.0010	0.0929	0.0852	0.931
	0	3	0.036	0.0018	0.0978	0.0937	0.933
	1	0	0.017	−0.0108	0.0844	0.0825	0.943
	1	1	0.024	−0.0098	0.0922	0.0859	0.930
	1	3	0.039	−0.0044	0.0948	0.0924	0.945
	3	0	0.051	−0.0277	0.0926	0.0900	0.911
	3	1	0.036	−0.0222	0.0902	0.0909	0.929
	3	3	0.025	−0.0312	0.0962	0.0970	0.931
1	0	0	0.036	0.2097	0.0843	0.0816	0.282
	0	1	0.034	0.2109	0.0874	0.0836	0.273
	0	3	0.031	0.2038	0.0968	0.0920	0.397
	1	0	0.048	0.2134	0.0831	0.0843	0.273
	1	1	0.049	0.2122	0.0931	0.0865	0.316
	1	3	0.045	0.2061	0.1006	0.0937	0.405
	3	0	0.152	0.2251	0.0951	0.0918	0.307
	3	1	0.142	0.2204	0.1019	0.0930	0.330
	3	3	0.079	0.2133	0.1005	0.0989	0.415
3	0	0	0.303	0.7241	0.1101	0.1037	0
	0	1	0.264	0.7110	0.1126	0.1032	0
	0	3	0.200	0.6995	0.1197	0.1106	0
	1	0	0.352	0.7521	0.1174	0.1065	0
	1	1	0.349	0.7483	0.1189	0.1091	0
	1	3	0.289	0.7253	0.1213	0.1122	0
	3	0	0.514	0.8256	0.1346	0.1196	0
	3	1	0.505	0.8215	0.1432	0.1228	0
	3	3	0.445	0.7966	0.1449	0.1267	0

For the analysis, define $N_i(t)$ to be the cumulative number of observed tumors at time $t, i = 1, \ldots, 85$. Let $X_{i1} = 1$ if patient i was in the thiotepa group and 0 otherwise, and X_{i2} to be the logarithm of the number of the initial tumors plus 1. Let τ be the longest observation time being 53 months. The application of the proposed methods in Sect. 3 with $W(t) = Q(t) = 1$ yielded $\hat{\beta}_1 = -1.5594$ and $\hat{\beta}_2 = 1.2991$ with the estimated standard errors 0.3817 and 0.3456, respectively. These results imply that both the thiotepa treatment and the initial number of tumors have significant effects on the tumor occurrence process. Specifically, the thiotepa treatment significantly reduces the bladder tumor occurrence rate, and the patients with higher number of initial tumors tend to have a higher tumor occurrence rate. These results are consistent with Sun and Wei (2000).

To assess the adequacy of the proposed models for the bladder cancer data, we applied the model checking techniques presented in Sect. 4 and obtained

$\sup_{x,t} |\mathcal{F}(x,t)| = 19.3408$. This gives a p value of 0.842 based on 1000 realizations of the corresponding statistic $\sup_{x,t} |\hat{\mathcal{F}}(x,t)|$ and indicates that model (1) seems to fit the bladder cancer data well.

7 Concluding remarks

This paper discussed the analysis of panel count data in the presence of both informative observation times and a terminal event. Some joint models were proposed to describe the recurrent event process, the observation times and the terminal event together via two latent variables. An estimating equation-based inference procedure was proposed for the estimation of parameters. Also a goodness-of-fit procedure was presented for assessing the appropriateness of the proposed models, and the simulation results indicated that the estimation procedure works well in practical situations. In addition, an illustrative example was also provided.

The proposed joint modeling approach offers a good choice for modeling panel count data in the presence of both informative observation times and a terminal event. Note that for the situation considered here, the modeling of the observation process and the terminal event could not be of interest, and the misspecification of these two models may lead to biased estimation for the recurrent event process. Simulation studies show that when the model for the recurrent event process is correct, small misspecifications of the models for the observation process and the terminal event lead to small biases for the estimation of the recurrent event process. However, when the misspecifications of these two models are severe, the biases could be large.

There remain several topics to study in the future. First, note that we only considered the time-independent covariates. In practice, they may be varying with time and thus it is desirable to extend the proposed procedure to the situation with time-dependent covariates. However, this is clearly not straightforward. Second, it would be useful to develop some variable selection procedures for the proposed joint model. For this, one possible way is to use the non-concave penalized estimating function approach (Tong et al. 2009) based on (7) and it is apparent that a lot of research efforts are needed for it.

Acknowledgments The authors thank the Editor-in Chief, Professor Mei-Ling Ting Lee, an Editor and two reviewers for their insightful comments and suggestions that greatly improved the article. The first author's research was partly supported by the National Natural Science Foundation of China (Grant No. 11301355). The second author's research was partly supported by the National Natural Science Foundation of China (Grant Nos. 11301212 and 11401146). The third author's research was partly supported by the National Natural Science Foundation of China (Grant Nos. 11231010 and 11171330), Key Laboratory of RCSDS, CAS (No. 2008DP173182) and BCMIIS.

Compliance with ethical standards

Conflict of interest The authors declare that they have no conflict of interest.

Research involving human participants and/or animals This article does not contain any studies with human participants or animals performed by any of the authors.

Informed consent Informed consent was obtained from all individual participants included in the study.

Appendix: Derivations of Eqs. (4) and (5) and the proofs of Theorems 1, 2 and 3

To obtain the asymptotic distributions of $\hat{\beta}$ and $\hat{\gamma}$, we need the following regularity conditions:

C.1. $\{N_i(\cdot), H_i(\cdot), T_i, \delta_i, X_i\}$, $i = 1, \ldots, n$ are independent and identically distributed.

C.2. $H(\tau)$ and X are bounded almost surely, $N(t)$ is of bounded variation and $P(T \geq \tau) > 0$.

C.3. A is nonsingular, where

$$
\begin{aligned}
A &= \begin{pmatrix} A_{11} & A_{12} \\ 0 & A_{22} \end{pmatrix}, \\
A_{11} &= E\left[\int_0^{\tau} W(t)\{X_i - \bar{x}(t, X_i)\}^{\otimes 2}\Delta_i(t)\right. \\
&\qquad \left. \times \exp\left\{X_i'(\beta_0 + \gamma_0)\right\} dR(t, \log \Lambda_0(t) + X_i'\eta_0)\right], \\
A_{22} &= E\left[\int_0^{\tau} Q(t)\left\{X_i - \bar{x}_i^*(t, X_i)\right\}^{\otimes 2}\Delta_i(t)\exp(X_i'\gamma_0)d\bar{H}_i(t; \gamma)\right], \\
A_{12} &= A_{11},
\end{aligned}
$$

and $\bar{x}(t, X_i)$ and $\bar{x}^*(t, X_i)$ are the limits of $\bar{X}_i(t; \beta_0, \gamma_0)$ and $\bar{X}_i^*(t; \gamma_0)$ conditional on X_i, respectively.

Derivations of Eqs. (4) and (5) Given (u, v, X) and $D \geq t$, we suppose that the mean function of the recurrent process and the rate of the observation process are independent of D. Note that $\mathcal{A}_0(t; u, v) = \int_0^t \mu_N(z; u)d\mu_H(z; v)$. Then by the independent censoring assumption and models (1) and (2), we obtain

$$
\begin{aligned}
&E\{N(t)dH(t)|X, T \geq t\} \\
&\quad = E\{N(t)dH(t)|X, D \geq t\} \\
&\quad = E[E\{N(t)dH(t)|X, D \geq t, u, v\}|X, D \geq t] \\
&\quad = \exp\left\{X'(\beta_0 + \gamma_0)\right\} E\{d\mathcal{A}_0(t; u, v)|X, D \geq t\} \\
&\quad = \exp\left\{X'(\beta_0 + \gamma_0)\right\} E\left\{d\mathcal{A}_0(t; u, v)|X, \epsilon \geq \log \Lambda_0(t) + X'\eta_0\right\},
\end{aligned}
$$

where the last equality is from (3). Define $dR(t, s) = E\{d\mathcal{A}_0(t; u, v)|\epsilon \geq s\}$. It can be checked that

$$
E\{N(t)dH(t)|X, T \geq t\} = \exp\left\{X'(\beta_0 + \gamma_0)\right\} dR(t, \log \Lambda_0(t) + X'\eta_0). \quad (10)
$$

For any integrable function $g(X, t, s)$, by the assumption that (u, v, ϵ) is independent of (X, C), we have

$$
dR(t, s) = E\{d\mathcal{A}_0(t; u, v)|\epsilon \geq s\}
$$

$$
\begin{aligned}
&= \frac{E\left\{d\mathcal{A}_0(t;u,v)I(\epsilon \geq s)\right\}}{E\left\{I(\epsilon \geq s)\right\}} \\
&= \frac{E\left[d\mathcal{A}_0(t;u,v)I(\epsilon \geq s)I\left\{\log \Lambda_0(C) + X'\eta_0 \geq s\right\} g(X,t,s)\right]}{E\left[I(\epsilon \geq s)I\left\{\log \Lambda_0(C) + X'\eta_0 \geq s\right\} g(X,t,s)\right]} \\
&= \frac{E\left[d\mathcal{A}_0(t;u,v)I\left\{\log \Lambda_0(T) + X'\eta_0 \geq s\right\} g(X,t,s)\right]}{E\left[I\left\{\log \Lambda_0(T) + X'\eta_0 \geq s\right\} g(X,t,s)\right]}. \qquad (11)
\end{aligned}
$$

In particular, we choose $g(X,t,s) = \exp\{X'(\beta_0+\gamma_0)\}I\left\{\log \Lambda_0(t) + X'\eta_0 \leq s\right\}$, the denominator of (11) becomes $E[\exp\left\{X'(\beta_0+\gamma_0)\right\}\Phi(T,X,t,s)]$, where $\Phi(T,X,t,s) = I\left\{\log \Lambda_0(T) + X'\eta_0 \geq s \geq \log \Lambda_0(t) + X'\eta_0\right\}$. Note that $\Phi(T,X,t,s) = 1$ implies $T \geq t$. similarly to (10), we get

$$
\begin{aligned}
&E\left\{N(t)dH(t)\Phi(T,X,t,s)\right\} = E\left[E\left\{N(t)dH(t)\Phi(T,X,t,s)|X,\Phi(T,X,t,s),u,v\right\}\right] \\
&\quad = E(E\left[\exp\left\{X'(\beta_0+\gamma_0)\right\} d\mathcal{A}_0(t;u,v)\Phi(T,X,t,s)|X,\Phi(T,X,t,s),u,v\right]) \\
&\quad = E\left[\exp\left\{X'(\beta_0+\gamma_0)\right\} d\mathcal{A}_0(t;u,v)\Phi(T,X,t,s)\right],
\end{aligned}
$$

which is the numerator of (11). Thus

$$
dR(t,s) = \frac{E\left[N(t)dH(t)I\left\{\log \Lambda_0(T) + X'\eta_0 \geq s \geq \log \Lambda_0(t) + X'\eta_0\right\}\right]}{E\left[\exp\left\{X'(\beta_0+\gamma_0)\right\} I\left\{\log \Lambda_0(T) + X'\eta_0 \geq s \geq \log \Lambda_0(t) + X'\eta_0\right\}\right]}.
$$

This completes the derivation. □

Proof of Theorem 1 Define

$$
\begin{aligned}
\Phi_i(t,X;\eta,\Lambda) &= I\left\{\log \Lambda(T_i) + X_i'\eta \geq \log \Lambda(t) + X'\eta \geq \log \Lambda(t) + X_i'\eta\right\}, \\
dM_i(t) &= N_i(t)dH_i(t), \\
d\bar{M}(t,X;\eta,\Lambda) &= \frac{\sum_{j=1}^n \Phi_j(t,X;\eta,\Lambda)dM_j(t)}{\sum_{j=1}^n \exp\left\{X_j'(\beta_0+\gamma_0)\right\}\Phi_j(t,X;\eta,\Lambda)}, \\
d\bar{M}_0(t,X;\eta,\Lambda) &= \frac{E\left\{\Phi_j(t,X;\eta,\Lambda)dM_j(t)|X\right\}}{E\left[\exp\left\{X_j'(\beta_0+\gamma_0)\right\}\Phi_j(t,X;\eta,\Lambda)|X\right]},
\end{aligned}
$$

and

$$
\bar{x}(t,X;\eta,\Lambda) = \frac{E\left[X_j\exp\left\{X_j'(\beta_0+\gamma_0)\right\}\Phi_j(t,X;\eta,\Lambda)|X\right]}{E\left[\exp\left\{X_j'(\beta_0+\gamma_0)\right\}\Phi_j(t,X;\eta,\Lambda)|X\right]}.
$$

Denote $\bar{x}(t, X) = \bar{x}(t, X; \eta_0, \Lambda_0)$ and $\Phi_i(t, X) = \Phi_i(t, X; \eta_0, \Lambda_0)$. Then we have

$$\begin{aligned}
& n^{-1/2} \sum_{i=1}^{n} \int_0^{\tau} W(t) \{X_i - \bar{x}(t, X_i)\} \Delta_i(t) \exp\left\{X_i'(\beta_0 + \gamma_0)\right\} \\
& \quad \times \left\{d\bar{M}(t, X_i; \hat{\eta}, \hat{\Lambda}_0) - d\bar{M}_0(t, X_i; \hat{\eta}, \hat{\Lambda}_0)\right\} \\
& = n^{-1/2} \sum_{i=1}^{n} \int_0^{\tau} W(t) \\
& \quad \times \left[\int \{x - \bar{x}(t, x)\} I(c \geq t) \frac{\exp\left\{x'(\beta_0 + \gamma_0)\right\} \Phi_i(t, x)}{E\left[\exp\left\{X_i'(\beta_0 + \gamma_0)\right\} \Phi_i(t, x)\right]} dF(x, c)\right] dM_i(t) \\
& \quad - n^{-1/2} \sum_{i=1}^{n} \int \left[\int_0^{\tau} W(t) \{x - \bar{x}(t, x)\} I(c \geq t) \exp\left\{x'(\beta_0 + \gamma_0)\right\} \Phi_i(t, x)\right. \\
& \quad \left. \times \frac{E\left\{\Phi_i(t, x) dM_i(t)\right\}}{\left(E\left[\exp\left\{X_i'(\beta_0 + \gamma_0)\right\} \Phi_i(t, x)\right]\right)^2}\right] \exp\left\{X_i'(\beta_0 + \gamma_0)\right\} dF(x, c) + o_p(1),
\end{aligned} \tag{12}$$

where $F(x, c)$ is the joint probability measure of (X_i, T_i). According to Fleming and Harrington (1991, Page 299), we obtain

$$\hat{\eta} - \eta_0 = \Omega^{-1} n^{-1} \sum_{i=1}^{n} \int_0^{\tau} \left\{X_i - \bar{x}^D(t)\right\} dM_i^D(t) + o_p(n^{-1/2}),$$

and

$$\hat{\Lambda}_0(t) - \Lambda_0(t) = n^{-1} \sum_{i=1}^{n} \int_0^{t} \frac{dM_i^D(z)}{s^{(0)}(z; \eta_0)} - \int_0^{t} \bar{x}^D(z)' d\Lambda_0(z)(\hat{\eta} - \eta_0) + o_p(n^{-1/2}),$$

where $M_i^D(t) = H_i^D(t) - \int_0^t \Delta_i(z) \exp(X_i'\eta_0) d\Lambda_0(z)$, and Ω, $s^{(0)}(t; \eta_0)$ and $\bar{x}^D(t)$ are the limits of $\hat{\Omega}$, $S^{(0)}(t; \eta_0)$ and $\bar{X}^D(t; \eta_0)$, respectively. Denote $dR_\eta(t, X)$ and $dR_\Lambda(t, X)$ as the derivative and the Hadamard derivative of $d\bar{M}_0(t, X; \eta_0, \Lambda_0)$ with respect to η and Λ, respectively. Then by the functional delta method, we get

$$\begin{aligned}
& n^{-1/2} \sum_{i=1}^{n} \int_0^{\tau} W(t) \{X_i - \bar{x}(t, X_i)\} \Delta_i(t) \exp\left\{X_i'(\beta_0 + \gamma_0)\right\} \\
& \quad \times \left\{d\bar{M}_0(t, X_i; \hat{\eta}, \hat{\Lambda}_0) - d\bar{M}_0(t, X_i; \eta_0, \Lambda_0)\right\} \\
& = n^{-1/2} \sum_{i=1}^{n} \int_0^{\tau} \left[P_1 \Omega^{-1} \left\{X_i - \bar{x}^D(t)\right\} + \frac{B_1(t)}{s^{(0)}(t; \eta_0)}\right] dM_i^D(t) + o_p(1),
\end{aligned} \tag{13}$$

where

$$P_1 = E\left[\int_0^\tau W(t)\{X_i - \bar{x}(t, X_i)\}\Delta_i(t)\exp\left\{X_i'(\beta_0+\gamma_0)\right\}\left\{dR_\eta(t, X_i) - \left(\int_0^t \bar{x}^D(z)' d\Lambda_0(z)\right) dR_\Lambda(t, X_i)\right\}\right],$$

and

$$B_1(t) = E\left[\int_t^\tau W(z)\{X_i - \bar{x}(z, X_i)\}\Delta_i(z)\exp\left\{X_i'(\beta_0+\gamma_0)\right\} dR_\Lambda(z, X_i)\right].$$

It follows from (12) and (13) that

$$\begin{aligned}
& n^{-1/2}U(\beta_0;\gamma_0) \\
&= n^{-1/2}\sum_{i=1}^n \int_0^\tau W(t)\left\{X_i - \bar{X}_i(t;\beta_0,\gamma_0)\right\}\Delta_i(t)\Big[dM_i(t) \\
&\quad - \exp\left\{X_i'(\beta_0+\gamma_0)\right\} d\bar{M}_0(t, X_i;\eta_0,\Lambda_0) \\
&\quad - \exp\left\{X_i'(\beta_0+\gamma_0)\right\}\left\{d\bar{M}(t, X_i;\hat{\eta},\hat{\Lambda}_0) - d\bar{M}_0(t, X_i;\hat{\eta},\hat{\Lambda}_0)\right\} \\
&\quad - \exp\left\{X_i'(\beta_0+\gamma_0)\right\}\left\{d\bar{M}_0(t, X_i;\hat{\eta},\hat{\Lambda}_0) - d\bar{M}_0(t, X_i;\eta_0,\Lambda_0)\right\}\Big] \\
&= n^{-1/2}\sum_{i=1}^n \xi_i + o_p(1),
\end{aligned} \tag{14}$$

where

$$\begin{aligned}
\xi_i &= \int_0^\tau W(t)\{X_i - \bar{x}(t, X_i)\}\Delta_i(t)[dM_i(t) - \exp\left\{X_i'(\beta_0+\gamma_0)\right\} d\bar{M}_0(t, X_i;\eta_0,\Lambda_0)] \\
&- \int_0^\tau W(t)\left[\int \{x - \bar{x}(t, x)\} I(c \geq t)\frac{\exp\left\{x'(\beta_0+\gamma_0)\right\}\Phi_i(t, x)}{E\left[\exp\left\{X_i'(\beta_0+\gamma_0)\right\}\Phi_i(t, x)\right]} dF(x, c)\right] dM_i(t) \\
&+ \int\left[\int_0^\tau W(t)\{x - \bar{x}(t, x)\} I(c \geq t)\exp\left\{x'(\beta_0+\gamma_0)\right\}\Phi_i(t, x)\right. \\
&\times \left.\frac{E\left\{\Phi_i(t, x) dM_i(t)\right\}}{(E\left[\exp\left\{X_i'(\beta_0+\gamma_0)\right\}\Phi_i(t, x)\right])^2}\right]\exp\left\{X_i'(\beta_0+\gamma_0)\right\} dF(x, c) \\
&- \int_0^\tau \left[P_1\Omega^{-1}\left\{X_i - \bar{x}^D(t)\right\} + \frac{B_1(t)}{s^{(0)}(t;\eta_0)}\right] dM_i^D(t).
\end{aligned}$$

Moreover, we define the following notations:

$$d\bar{H}(t, X;\eta,\Lambda) = \frac{\sum_{j=1}^n dH_j(t)\Phi_j(t, X;\eta,\Lambda)}{\sum_{j=1}^n \exp(X_j'\gamma_0)\Phi_j(t, X;\eta,\Lambda)},$$

$$d\bar{H}_0(t, X; \eta, \Lambda) = \frac{E\left[dH_j(t)\Phi_j(t, X; \eta, \Lambda)|X\right]}{E\left[\exp(X_j'\gamma_0)\Phi_j(t, X; \eta, \Lambda)|X\right]},$$

$$\bar{x}^*(t, X; \eta, \Lambda) = \frac{E\left[X_j \exp(X_j'\gamma_0)\Phi_j(t, X; \eta, \Lambda)|X\right]}{E\left[\exp(X_j'\gamma_0)\Phi_j(t, X; \eta, \Lambda)|X\right]}.$$

Denote $\bar{x}^*(t, X) = \bar{x}^*(t, X; \eta_0, \Lambda_0)$. In a similar manner, we can get

$$\begin{aligned}
& n^{-1/2}\sum_{i=1}^{n}\int_0^\tau Q(t)\left\{X_i - \bar{x}^*(t, X_i)\right\}\Delta_i(t)\exp(X_i'\gamma_0) \\
& \quad \times \left\{d\bar{H}(t, X; \hat{\eta}, \hat{\Lambda}_0) - d\bar{H}_0(t, X; \hat{\eta}, \hat{\Lambda}_0)\right\} \\
& = n^{-1/2}\sum_{i=1}^{n}\int_0^\tau Q(t)\left[\int \left\{x - \bar{x}^*(t, x)\right\} I(c \geq t)\frac{\exp(x'\gamma_0)\Phi_i(t, x)}{E\left[\exp(X_i'\gamma_0)\Phi_i(t, x)\right]}dF(x, c)\right]dH_i(t) \\
& \quad - n^{-1/2}\sum_{i=1}^{n}\int \left[\int_0^\tau Q(t)\left\{x - \bar{x}^*(t, x)\right\} I(c \geq t)\exp(x'\gamma_0)\Phi_i(t, x)\right. \\
& \quad \left. \times \frac{E\left\{\Phi_i(t, x)dH_i(t)\right\}}{(E\left[\exp(X_i'\gamma_0)\Phi_i(t, x)\right])^2}\right]\exp(X_i'\gamma_0)dF(x, c) + o_p(1), \qquad (15)
\end{aligned}$$

and

$$\begin{aligned}
& n^{-1/2}\sum_{i=1}^{n}\int_0^\tau Q(t)\left\{X_i - \bar{x}^*(t, X_i)\right\}\Delta_i(t)\exp(X_i'\gamma_0) \\
& \quad \times \left\{d\bar{H}_0(t, X_i; \hat{\eta}, \hat{\Lambda}_0) - d\bar{H}_0(t, X_i; \eta_0, \Lambda_0)\right\} \\
& = n^{-1/2}\sum_{i=1}^{n}\int_0^\tau \left[P_2\Omega^{-1}\left\{X_i - \bar{x}^D(t)\right\} + \frac{B_2(t)}{s^{(0)}(t; \eta_0)}\right]dM_i^D(t) + o_p(1), \qquad (16)
\end{aligned}$$

where

$$\begin{aligned}
P_2 = E\Bigg[\int_0^\tau & Q(t)\left\{X_i - \bar{x}^*(t, X_i)\right\}\Delta_i(t)\exp(X_i'\gamma_0)\Big\{dR_\eta^*(t, X_i) \\
& - \left(\int_0^t \bar{x}^D(z)'d\Lambda_0(z)\right)dR_\Lambda^*(t, X_i)\Big\}\Bigg],
\end{aligned}$$

$$B_2(t) = E\left[\int_t^\tau Q(z)\left\{X_i - \bar{x}^*(z, X_i)\right\}\Delta_i(z)\exp(X_i'\gamma_0)dR_\Lambda^*(z, X_i)\right],$$

and $dR_\eta^*(t, X)$ and $dR_\Lambda^*(t, X)$ as the derivative and the Hadamard derivative of $d\bar{H}_0(t, X; \eta_0, \Lambda_0)$ with respectively to η and Λ, respectively. It follows from (15) and (16) that

$$n^{-1/2}\tilde{U}(\gamma_0) = n^{-1/2}\sum_{i=1}^{n}\zeta_i + o_p(1), \tag{17}$$

where

$$\begin{aligned}
\zeta_i = & \int_0^\tau Q(t)\left\{X_i - \bar{x}^*(t, X_i)\right\}\Delta_i(t)\exp(X_i'\gamma_0) \\
& \times\left[dH_i(t) - \exp(X_i'\gamma_0)d\bar{H}_0(t, X_i; \eta_0, \Lambda_0)\right] \\
& - \int_0^\tau Q(t)\left[\int\{x - \bar{x}^*(t, x)\}I(c \geq t)\frac{\exp(x'\gamma_0)\Phi_i(t, x)}{E\left[\exp(X_i'\gamma_0)\Phi_i(t, x)\right]}dF(x, c)\right]dH_i(t) \\
& + \int\left[\int_0^\tau Q(t)\{x - \bar{x}^*(t, x)\}I(c \geq t)\exp(x'\gamma_0)\Phi_i(t, x)\right. \\
& \left.\times\frac{E\left\{\Phi_i(t, x)dH_i(t)\right\}}{(E\left[\exp(X_i'\gamma_0)\Phi_i(t, x)\right])^2}\right]\exp(X_i'\gamma_0)dF(x, c) \\
& - \int_0^\tau\left[P_2\Omega^{-1}\{X_i - \bar{x}^D(t)\} + \frac{B_2(t)}{s^{(0)}(t; \eta_0)}\right]dM_i^D(t).
\end{aligned}$$

Notice that $-n^{-1}\partial U(\beta_0; \gamma_0)/\partial\beta$, $-n^{-1}\partial U(\beta_0; \gamma_0)/\partial\gamma$ and $-n^{-1}\partial\tilde{U}(\gamma_0)/\partial\gamma$ convergence in probability to A_{11}, A_{12} and A_{22}, respectively. Then using (14), (17) and the Taylor expansion, we obtain

$$\begin{aligned}
n^{1/2}\begin{pmatrix}\hat{\beta} - \beta_0 \\ \hat{\gamma} - \gamma_0\end{pmatrix} &= A^{-1}n^{-1/2}\begin{pmatrix}U(\beta_0; \gamma_0) \\ \tilde{U}(\gamma_0)\end{pmatrix} + o_p(1) \\
&= A^{-1}n^{-1/2}\sum_{i=1}^{n}\begin{pmatrix}\xi_i \\ \zeta_i\end{pmatrix} + o_p(1).
\end{aligned}$$

Thus,

$$n^{1/2}\begin{pmatrix}\hat{\beta} - \beta_0 \\ \hat{\gamma} - \gamma_0\end{pmatrix} \xrightarrow{\mathcal{L}} N(0, A^{-1}\Sigma(A^{-1})'),$$

where $\xrightarrow{\mathcal{L}}$ denotes convergence in distribution, $\Sigma = E\{(\xi_i', \zeta_i')'^{\otimes 2}\}$. This completes the proof. □

Proof of Theorem 2 Since $\hat{\beta}$, $\hat{\gamma}$, $\hat{\eta}$ and $\hat{\Lambda}_0(t)$ are consistent, we can obtain that conditional on the observed data,

$$\begin{aligned}
n^{-1/2}\Omega_1^* = & \sum_{i=1}^{n}\mathcal{Z}_i\int_0^\tau W(t)\{X_i - \bar{x}(t, X_i)\}\Delta_i(t)[dM_i(t) \\
& - \exp\left\{X_i'(\beta_0 + \gamma_0)\right\}d\bar{M}_0(t, X_i; \eta_0, \Lambda_0)] + o_p(1).
\end{aligned} \tag{18}$$

Similarly,

$$
\begin{aligned}
n^{-1/2}\Omega_2^* = & -n^{-1/2}\sum_{i=1}^{n}\mathcal{Z}_i\int_0^\tau W(t)\left[\int\{x-\bar{x}(t,x)\}I(c\geq t)\right.\\
&\left.\times\frac{\exp\left\{x'(\beta_0+\gamma_0)\right\}\Phi_i(t,x)}{E\left[\exp\left\{X_i'(\beta_0+\gamma_0)\right\}\Phi_i(t,x)\right]}dF(x,c)\right]dM_i(t)\\
&+n^{-1/2}\sum_{i=1}^{n}\mathcal{Z}_i\int\left[\int_0^\tau W(t)\{x-\bar{x}(t,x)\}I(c\geq t)\right.\\
&\left.\times\exp\left\{x'(\beta_0+\gamma_0)\right\}\Phi_i(t,x)\frac{E\left\{\Phi_i(t,x)dM_i(t)\right\}}{\left(E\left[\exp\left\{X_i'(\beta_0+\gamma_0)\right\}\Phi_i(t,x)\right]\right)^2}\right]\\
&\times\exp\left\{X_i'(\beta_0+\gamma_0)\right\}dF(x,c)+o_p(1).
\end{aligned}
\tag{19}
$$

Define

$$
d\hat{M}(t,X;\eta,\Lambda)=\frac{\sum_{j=1}^{n}\Phi_j(t,X;\eta,\Lambda)dM_j(t)}{\sum_{j=1}^{n}\exp\left\{X_j'\left(\hat{\beta}+\hat{\gamma}\right)\right\}\Phi_j(t,X;\eta,\Lambda)}.
$$

Thus,

$$
\begin{aligned}
\Omega_3^* = & \sum_{i=1}^{n}\int_0^\tau W(t)\left\{X_i-\bar{X}_i\left(t;\hat{\beta},\hat{\gamma}\right)\right\}\Delta_i(t)\exp\left\{X_i'\left(\hat{\beta}+\hat{\gamma}\right)\right\}\\
&\times\left\{d\hat{M}\left(t,X;\hat{\eta},\hat{\Lambda}_0\right)-d\hat{M}\left(t,X;\hat{\eta}^*,\hat{\Lambda}_0^*\right)\right\}.
\end{aligned}
$$

Moreover, we notice that

$$
\begin{aligned}
& d\hat{M}\left(t,X;\hat{\eta},\hat{\Lambda}_0\right)-d\hat{M}\left(t,X;\hat{\eta}^*,\hat{\Lambda}_0^*\right)\\
&=\left\{d\hat{M}\left(t,X;\hat{\eta},\hat{\Lambda}_0\right)-d\bar{M}_0\left(t,X;\hat{\eta},\hat{\Lambda}_0\right)\right\}\\
&\quad-\left\{d\hat{M}\left(t,X;\hat{\eta}^*,\hat{\Lambda}_0^*\right)-d\bar{M}_0\left(t,X;\hat{\eta}^*,\hat{\Lambda}_0^*\right)\right\}\\
&\quad+\left\{d\bar{M}_0\left(t,X;\hat{\eta},\hat{\Lambda}_0\right)-d\bar{M}_0\left(t,X;\hat{\eta}^*,\hat{\Lambda}_0^*\right)\right\}.
\end{aligned}
$$

Using a similar method to the proof of (12), we get

$$
n^{-1/2}\sum_{i=1}^{n}\int_0^\tau W(t)\left\{X_i-\bar{X}_i\left(t;\hat{\beta},\hat{\gamma}\right)\right\}\Delta_i(t)
$$

$$\exp\left\{X_i'\left(\hat\beta+\hat\gamma\right)\right\}\Big[\left\{d\hat M\left(t,X;\hat\eta,\hat\Lambda_0\right)-d\bar M_0\left(t,X;\hat\eta,\hat\Lambda_0\right)\right\}$$
$$-\left\{d\hat M\left(t,X;\hat\eta^*,\hat\Lambda_0^*\right)-d\bar M_0\left(t,X;\hat\eta^*,\hat\Lambda_0^*\right)\right\}\Big]=o_p(1)$$

Similarly to (13), we obtain that conditional on the observed data,

$$n^{-1/2}\sum_{i=1}^n\int_0^\tau W(t)\left\{X_i-\bar X_i\left(t;\hat\beta,\hat\gamma\right)\right\}\Delta_i(t)\exp\left\{X_i'\left(\hat\beta+\hat\gamma\right)\right\}$$
$$\times\left\{d\bar M_0\left(t,X;\hat\eta^*,\hat\Lambda_0^*\right)-d\bar M_0\left(t,X;\hat\eta,\hat\Lambda_0\right)\right\}$$
$$=n^{-1/2}\sum_{i=1}^n\mathcal{Z}_i\int_0^\tau\left[P_1\Omega^{-1}\{X_i-\bar x^D(t)\}+\frac{B_1(t)}{s^{(0)}(t;\eta_0)}\right]dM_i^D(t)$$
$$+o_p(1). \tag{20}$$

It follows from (18), (19) and (20) that

$$\hat\Upsilon_1=n^{-1/2}(\Omega_1^*+\Omega_2^*+\Omega_3^*)=n^{-1/2}\sum_{i=1}^n\mathcal{Z}_i\xi_i+o_p(1).$$

In a similar way, we obtain

$$\hat\Upsilon_2=n^{-1/2}(\Omega_4^*+\Omega_5^*+\Omega_6^*)=n^{-1/2}\sum_{i=1}^n\mathcal{Z}_i\zeta_i+o_p(1).$$

Thus, by the Theorem 3.6.13 of van der Vaart and Wellner (1996), $E_{\mathcal{Z}}(\hat\Upsilon^{\otimes 2})\xrightarrow{P}\Sigma$. This completes the proof of this theorem. □

Proof of Theorem 3 We notice that

$$\mathcal{F}(t,x)=n^{-1/2}\sum_{i=1}^n\int_0^t I(X_i\le x)\Delta_i(z)\Big[N_i(z)dH_i(z)-\exp\left\{X_i'(\beta_0+\hat\gamma)\right\}$$
$$\times\frac{\sum_{j=1}^n N_j(z)dH_j(z)\hat\Phi_j(z,X_i)}{\sum_{j=1}^n\exp\left\{X_j'(\beta_0+\hat\gamma)\right\}\hat\Phi_j(z,X_i)}\Big]-n^{1/2}\hat\Gamma_1(t,x)'(\hat\beta-\beta_0). \tag{21}$$

Following similar arguments as in the proof of Theorem 1, we can show that the first term on the right-hand side of (21) is

$$n^{-1/2}\sum_{i=1}^n\Psi_i(t,x)-n^{1/2}\Gamma_1(t,x)'(\hat\gamma-\gamma_0)+o_p(1), \tag{22}$$

where

$$\Psi_i(t,x) = \int_0^t I(X_i \le x)\Delta_i(z)\left[dM_i(z) - \exp\left\{X_i'(\beta_0+\gamma_0)\right\} d\bar{M}_0(t, X_i; \eta_0, \Lambda_0)\right]$$
$$- \int_0^t \left[\int I(s \le x) I(c \ge z) \frac{\exp\left\{s'(\beta_0+\gamma_0)\right\} \Phi_i(z,s)}{E\left[\exp\left\{X_i'(\beta_0+\gamma_0)\right\}\Phi_i(z,s)\right]} dF(s,c)\right] dM_i(z)$$
$$+ \int \left[\int_0^t I(s \le x) I(c \ge z) \exp\left\{s'(\beta_0+\gamma_0)\right\} \Phi_i(z,s)\right.$$
$$\left.\times \frac{E\{\Phi_i(z,s) dM_i(z)\}}{\left(E\left[\exp\left\{X_i'(\beta_0+\gamma_0)\right\}\Phi_i(z,s)\right]\right)^2}\right] \exp\left\{X_i'(\beta_0+\gamma_0)\right\} dF(s,c)$$
$$- P_1^*(t,x)\Omega^{-1} \int_0^\tau \left[\{X_i - \bar{x}^D(z)\} + \frac{B_1^*(t,z,x)}{s^{(0)}(z;\eta_0)}\right] dM_i^D(z),$$

$$P_1^*(t,x) = E\Big[\int_0^t I(X_i \le x)\Delta_i(s)\exp\left\{X_i'(\beta_0+\gamma_0)\right\}\Big\{dR_\eta(s, X_i)$$
$$- \left(\int_0^s \bar{x}^D(z)' d\Lambda_0(z)\right) dR_\Lambda(s, X_i)\Big\}\Big],$$
$$B_1^*(t,z,x) = E\left[\int_z^t I(X_i \le x)\Delta_i(s)\exp\left\{X_i'(\beta_0+\gamma_0)\right\} dR_\Lambda(s, X_i)\right],$$

and $\Gamma_1(t,x)$ is the limit of $\hat{\Gamma}_1(t,x)$. Furthermore, it can be shown that the second term on the right-hand side of (21) is equivalent to

$$-n^{1/2}\Gamma_1(t,x)(\hat{\beta} - \beta_0) + o_p(1). \tag{23}$$

Thus, using (21), (22), (23) and Theorem 1, we obtain

$$\mathcal{F}(t,x) = n^{-1/2}\sum_{i=1}^n \left\{\Psi_i(t,x) - \Gamma(t,x)' A^{-1}(\xi_i', \zeta_i')'\right\} + o_p(1), \tag{24}$$

where $\Gamma(t,x) = (\Gamma_1(t,x)', \Gamma_1(t,x)')'$. By the multivariate central limit theorem, $\mathcal{F}(t,x)$ converges in finite-dimensional distribution to a zero-mean Gaussian process. It is easy to see that $\mathcal{F}(t,x)$ is tight. Then using a similar method as the proof of Theorem 2, we obtain that $\mathcal{F}(t,x)$ converges weakly to a zero-mean Gaussian process and the null distribution can be approximated by (9). This completes the proof. □

References

Balakrishnan N, Zhao X (2009) New multi-sample nonparametric tests for panel count data. Ann Stat 37:1112–1149

Balakrishnan N, Zhao X (2010) A nonparametric test for the equality of counting processes with panel count data. Comput Stat Data Anal 54:135–142

Balakrishnan N, Zhao X (2011) A class of multi-sample nonparametric tests for panel count data. Ann Inst Stat Math 63:135–156

Cook RJ, Lawless JF (1997) Marginal analysis of recurrent events and a terminating event. Stat Med 16:911–924

Fleming T, Harrington D (1991) Counting processes and survival analysis. Wiley, New York

Ghosh D, Lin DY (2002) Marginal regression models for recurrent and terminal events. Stat Sin 12:663–688

Huang CY, Wang MC (2004) Joint modeling and estimation of recurrent event processes and failure time. J Am Stat Assoc 99:1153–1165

Huang C, Wang M, Zhang Y (2006) Analysing panel count data with informative observation times. Biometrika 93:763–775

Hu X, Sun J, Wei LJ (2003) Regression parameter estimation from panel counts. Scand J Stat 30:25–43

Kalbfleisch JD, Lawless JF (1985) The analysis of panel data under a Markov assumption. J Am Stat Assoc 80:863–871

Li N, Sun L, Sun J (2010) Semiparametric transformation models for panel count data with dependent observation process. Stat Biosci 2:191–210

Liang K, Zeger S (1986) Longitudinal data analysis using generalized linear models. Biometrika 73:13–22

Lin DY, Wei LJ, Yang I, Ying Z (2000) Semiparametric regression for the mean and rate function of recurrent events. J R Stat Soc B 62:711–730

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

Park D, Sun J, Zhao X (2007) A class of two-sample nonparametric tests for panel count data. Commun Stat Theory Methods 36:1611–1625

Sun J (2006) The statistical analysis of interval-censored failure time data. Springer, New York

Sun J, Fang H (2003) A nonparametric test for panel count data. Biometrika 90:199–208

Sun J, Kalbfleisch JD (1995) Estimation of the mean function of point processes based on panel count data. Stat Sin 5:279–290

Sun J, Tong X, He X (2007) Regression analysis of panel count data with dependent observation times. Biometrics 63:1053–1059

Sun J, Wei LJ (2000) Regression analysis of panel count data with covariate-dependent observation and censoring times. J R Stat Soc B 62:293–302

Sun L, Song X, Zhou J, Liu L (2012) Joint analysis of longitudinal data with informative observation times and a dependent terminal event. J Am Stat Assoc 107:688–700

Sun J, Zhao X (2013) The statistical analysis of panel count data. Springer, New York

Thall PF, Lachin JM (1988) Analysis of recurrent events: nonparametric methods for random-interval count data. J Am Stat Assoc 83:339–347

Tong X, He X, Sun L, Sun J (2009) Variable selection for panel count data via non-concave penalized estimating function. Scand J Stat 36:620–635

van der Vaart AW, Wellner JA (1996) Weak convergence and empirical processes. Springer, New York

Wellner JA, Zhang Y (2007) Two likelihood-based semiparametric estimation methods for panel count data with covariates. Ann Stat 35:2106–2142

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

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

Zhang H, Sun J, Wang D (2013a) Variable selection and estimation for multivariate panel count data via the seamless-L0 penalty. Can J Stat 41:368–385

Zhang H, Zhao H, Sun J, Wang D, Kim K (2013b) Regression analysis of multivariate panel count data with an informative observation process. J Multivar Anal 119:71–80

Zhang Y (2002) A semiparametric pseudolikelihood estimation method for panel count data. Biometrika 89:39–48

Zhao H, Li Y, Sun J (2013) Analyzing panel count data with a dependent observation process and a terminal event. Can J Stat 41:174–191

Zhao X, Sun J (2011) Nonparametric comparison for panel count data with unequal observation processes. Biometrics 67:770–779

Zhao X, Zhou J, Sun L (2011) Semiparametric transformation models with time-varying coefficients for recurrent and terminal events. Biometrics 67:404–414