SCIENCE CHINA
Mathematics

Progress of Projects Supported by NSFC
• ARTICLES •

June 2015 Vol. 58 No. 6: 1163–1178
doi: 10.1007/s11425-015-5000-x

An additive-multiplicative rates model for multivariate recurrent events with event categories missing at random

YE Peng[1,*], SUN LiuQuan[1], ZHAO XingQiu[2] & XU Wei[3]

[1] *Institute of Applied Mathematics, Academy of Mathematics and Systems Science, Chinese Academy of Sciences, Beijing* 100080*, China;*
[2] *Department of Applied Mathematics, The Hong Kong Polytechnic University Shenzhen Research Institute, The Hong Kong Polytechnic University, Hong Kong, China;*
[3] *Dalla Lana School of Public Health, University of Toronto, Toronto, Canada*
Email: yepeng@amss.ac.cn, slq@amt.ac.cn, xingqiu.zhao@polyu.edu.hk

Received September 15, 2014; accepted December 24, 2014; published online March 16, 2015

Abstract Multivariate recurrent event data arises when study subjects may experience more than one type of recurrent events. In some situations, however, although event times are always observed, event categories may be partially missing. In this paper, an additive-multiplicative rates model is proposed for the analysis of multivariate recurrent event data when event categories are missing at random. A weighted estimating equations approach is developed for parameter estimation, and the resulting estimators are shown to be consistent and asymptotically normal. In addition, a model-checking technique is presented to assess the adequacy of the model. Simulation studies are conducted to evaluate the finite sample behavior of the proposed estimators, and an application to a platelet transfusion reaction study is provided.

Keywords additive-multiplicative rates model, missing data, multivariate recurrent events, semiparametric model, weighted estimating equation

MSC(2010) 62N01, 62G05

Citation: Ye P, Sun L Q, Zhao X Q, et al. An additive-multiplicative rates model for multivariate recurrent events with event categories missing at random. Sci China Math, 2015, 58: 1163–1178, doi: 10.1007/s11425-015-5000-x

1 Introduction

Recurrent event data arises frequently in many clinical and longitudinal studies where the event of interest may occur more than once over time. Examples include repeated transient ischemic attacks in cerebrovascular disease study, multiple opportunistic infections in HIV clinical trials, and recurrent pulmonary exacerbations in cystic fibrosis trials. Other examples of recurrent events that often occur in practice include hospitalizations, tumor metastases and bleeding incidents. A lot of work exists with respect to the analysis of recurrent event data [1, 8, 16, 21]. For example, Lin et al. [8] developed rigorous parameter estimation procedures for the marginal proportional rate model. Schaubel et al. [21] studied a semiparametric additive rate model for recurrent events. All the aforementioned methods are for the analysis of univariate recurrent events. In many settings, especially in medical studies, study subjects may experience several different types of related recurrent events, leading to multivariate recurrent event data. For example, infections in bone marrow transplantation can be classified into bacterial, fungal

*Corresponding author

and viral origins; transient ischemic attacks may be differentiated according to location in cardiovascular trials; and childhood asthma outcomes may be subtyped as physician office visits and hospitalizations. For the analysis of multivariate recurrent event data, some estimation procedures have been developed in the literature [2, 4, 24, 28]. For example, Cai and Schaubel [2] proposed a class of proportional marginal means and rates models for evaluating the effects of the covariates on the recurrent event processes. Chen et al. [4] presented a general additive marginal rate model for analyzing the multivariate recurrent event data. Sun et al. [24] studied a semiparametric multiplicative rates model that allows for time-dependent covariate effects. Zhu et al. [28] considered regression analysis of multivariate recurrent events in the presence of a dependent terminal event based on a joint modeling approach.

However, in practice, although the occurrence of an event is always observed, the specific category may be missing due to a variety of reasons. For example, in the study of technique failure rates for end-stage renal disease patients receiving continuous ambulatory peritoneal dialysis [19], although the dates of all technique failures are available, the causes of some technique failures are uncertain due to technical difficulties. Another example is a platelet transfusion reaction study [27] wherein hematology/oncology patients may experience different types of febrile nonhemolytic transfusion reactions. The occurrence time of each transfusion reaction is always recorded, but the specific types are partially missing. For multivariate recurrent event data with missing event types, the complete-case analysis is to discard those events with missing categories, which may lead to biased parameter estimation unless the missingness mechanism is missing completely at random (MCAR) [11], i.e., the event category missingness occurs randomly among events. When the category missingness mechanism depends only on the observed quantities, but not on the missing categories, it is termed missing at random (MAR) [11]. Compared with the MCAR assumption, MAR is a weaker assumption and often holds in practice.

The issue of missing event types in recurrent event data is analogous to the competing risk issue with missing cause of failure, which has been investigated by [5, 13, 14, 22]. For example, Gao and Tsiatis [5] considered the linear transformation competing risks model with missing cause of failure. Lu and Liang [14] studied competing risks data with missing cause of failure under the semiparametric additive hazards model. In recent years, some methods have been developed to analyze multivariate recurrent event data with missing event types under the MAR assumption. For example, Schaubel and Cai [19, 20] studied the multiple-event-type proportional means and rates model based on weighted estimating equations and multiple imputation methods, respectively. Chen and Cook [3] presented a multivariate random effects model using the likelihood approach. Lin et al. [10] proposed a nonparametric estimation of the mean function in which local likelihood methods are employed to estimate the event category probabilities.

In application, some covariate effects are additive while others are multiplicative or certain covariates have both the additive and multiplicative effects. To enhance the modelling capability in many applications, it seems natural to consider models which allow some covariate effects to be multiplicative while allowing others to be additive. Recently, a class of additive-multiplicative rate models [12, 23] have been proposed to analyze recurrent events of single type. In this paper, we consider a semiparametric additive-multiplicative rates model to study multivariate recurrent event data with event types missing at random. A weighted estimating equations approach is developed to estimate model parameters. Simulation results in Section 5 show that the proposed methods work well under the MAR assumption.

The remainder of the article is organized as follows. In Section 2, we specify the model and present an estimating equation method for parameter estimation. Section 3 gives the asymptotic properties of the resulting estimators with proofs outlined in Appendix. A model-checking technique is described in Section 4. Simulation studies are conducted to evaluate the proposed methods in Section 5. An application on the platelet transfusion reaction study is provided in Section 6 and some concluding remarks are summarized in Section 7.

2 Model and estimation procedures

Consider a study that involves n independent subjects and each subject may experience K different types of recurrent events. For each pair (i,k), let $N^*_{ik}(t)$ denote the number of events of type k that have occurred up to time t for subject i, and assume that $dN^*_{ik}(t) \in \{0,1\}$, and $dN^*_{ik}(t)dN^*_{il}(t) = 0$ for $k \neq l$. Let C_{ik} denote the right censoring times of event type k for subject i. Due to censoring, the observed recurrent event processes are $N_{ik}(t) = \int_0^t Y_{ik}(s)dN^*_{ik}(s)$, where $Y_{ik}(t) = I(C_{ik} \geqslant t)$, and $I(\cdot)$ is the indicator function. Let $W_{ik}(t)$ and $X_{ik}(t)$ be vectors of external time-dependent covariates of dimensions p and q (see [6]), which represent additive and multiplicative effects on the recurrent event processes, respectively. The proposed additive-multiplicative rates model takes the form

$$E\{dN^*_{ik}(t) \mid W_{ik}(t), X_{ik}(t)\} = g(\beta_0^{\mathrm{T}} W_{ik}(t))dt + h(\gamma_0^{\mathrm{T}} X_{ik}(t))d\mu_{0k}(t), \tag{1}$$

where $g(\cdot)$ and $h(\cdot)$ are known link functions, β_0 and γ_0 are $p \times 1$ and $q \times 1$ vectors of unknown regression coefficients, and $\mu_{0k}(t)$ is the category-specific baseline mean function. Obviously, (1) defines a very rich family of models through the link functions g and h, which includes the proportional rate model [2] and the additive rate model [4] as special cases. In practice, the choice of these appropriate link functions may be based on prior data or the desiring interpretation of the regression parameters.

We first consider the situation where event categories are fully observed. Let $Z_{ik}(t) = \{W^{\mathrm{T}}_{ik}(t), X^{\mathrm{T}}_{ik}(t)\}^{\mathrm{T}}$ and define

$$M^*_{ik}(t,\theta) = N_{ik}(t) - \int_0^t Y_{ik}(s)\{g(\beta^{\mathrm{T}} W_{ik}(t))dt + h(\gamma^{\mathrm{T}} X_{ik}(t))d\mu_{0k}(t)\},$$

where $\theta = (\beta^{\mathrm{T}}, \gamma^{\mathrm{T}})^{\mathrm{T}}$ is included in the parameter space denoted by Θ, which contains the true value $\theta_0 = (\beta_0^{\mathrm{T}}, \gamma_0^{\mathrm{T}})^{\mathrm{T}}$ as its interior point, and is assumed to be compact for the technical proof. Under (1) and independent censoring assumption, we have $E\{dM^*_{ik}(t,\theta_0) \mid Z_{ik}(t)\} = 0$. By applying the idea of generalized estimating equations [7], we specify the following estimating equations for $\mu_{0k}(t)$ and θ_0, respectively:

$$\sum_{i=1}^n \int_0^t dM^*_{ik}(s,\theta) = 0, \tag{2}$$

$$\sum_{i=1}^n \sum_{k=1}^K \int_0^\tau Z_{ik}(s,\theta)dM^*_{ik}(s,\theta) = 0, \tag{3}$$

where τ is a prespecified time point satisfying $P(Y_{ik}(\tau) = 1) > 0$ for $k = 1,\dots,K$ and $i = 1,\dots,n$. For fixed θ, based on (2), some simple calculations give that

$$\tilde{\mu}_{0k}(t,\theta) = \sum_{i=1}^n \int_0^t \frac{dN_{ik}(s) - Y_{ik}(s)g(\beta^{\mathrm{T}} W_{ik}(s))ds}{\sum_{i=1}^n Y_{ik}(s)h(\gamma^{\mathrm{T}} X_{ik}(s))}, \tag{4}$$

for $k = 1,\dots,K$. Substituting (4) into (3) leads to the following estimating equation for θ_0:

$$\sum_{i=1}^n \sum_{k=1}^K \int_0^\tau \{Z_{ik}(s) - \bar{Z}_k(s,\gamma)\}\{dN_{ik}(t) - Y_{ik}(t)g(\beta^{\mathrm{T}} W_{ik}(t))dt\} = 0, \tag{5}$$

where

$$\bar{Z}_k(t,\gamma) = \frac{\sum_{i=1}^n Y_{ik}(t)h(\gamma^{\mathrm{T}} X_{ik}(t))Z_{ik}(t)}{\sum_{i=1}^n Y_{ik}(t)h(\gamma^{\mathrm{T}} X_{ik}(t))}.$$

Let $\tilde{\theta}$ be the solution to (5) and the baseline mean function estimator is given by $\tilde{\mu}_{0k}(t,\tilde{\theta})$.

We now consider the setting where event times are always observed, but some event categories are missing at random. Let $\delta_i(t)$ denote the type of event which occurred to subject i at time t. Define $\delta_{ik}(t) = I\ (\delta_i(t) = k)$, and set $\xi_i(t) = 1$ when an event occurs at time t and $\delta_i(t)$ is known and 0 otherwise. For a random sample of n subjects, the observed data consists of $\{N_{ik}(t), C_{ik}, Z_{ik}(t), \xi_i(t), \xi_i(t)\delta_i(t); t \leqslant C_{ik}, i = 1,\dots,n, k = 1,\dots,K\}$. A complete-case analysis is to simply ignore events of missing types, which not only leads to efficiency loss due to smaller sample size but also generates biased estimators when the missingness mechanism is MAR.

Define $dN_{i\cdot}(t) = \sum_{k=1}^{K} dN_{ik}(t)$. Since $dN_{ik}(t)dN_{il}(t) = 0$ for $k \neq l$, it follows that $dN_{ik}(t) = \delta_{ik}(t)dN_{i\cdot}(t)$, and

$$\begin{aligned} dN_{ik}(t) &= \xi_i(t)dN_{ik}(t) + (1-\xi_i(t))dN_{ik}(t) \\ &= \xi_i(t)dN_{ik}(t) + (1-\xi_i(t))\delta_{ik}(t)dN_{i\cdot}(t) \\ &= \xi_i(t)dN_{ik}(t) + \delta_{ik}(t)dN_{i\cdot}^c(t), \end{aligned}$$

where $dN_{i\cdot}^c(t) = (1-\xi_i(t))dN_{i\cdot}(t)$. Under (1), we have

$$E[\xi_i(t)dN_{ik}(t) + \delta_{ik}(t)dN_{i\cdot}^c(t) - Y_{ik}(t)\{g(\beta_0^{\mathrm{T}}W_{ik}(t))dt + h(\gamma_0^{\mathrm{T}}X_{ik}(t))d\mu_{0k}(t)\}] = 0.$$

However, by which we can not establish operational estimating equations since $\delta_{ik}(t)$ is unavailable when $dN_{i\cdot}^c(t) = 1$. Let $V_i(t)$ be a covariate vector which catches the pertinent information in the event history at time t for subject i. Under the MAR assumption, we have

$$\begin{aligned} E[\delta_{ik}(t) \mid dN_{i\cdot}^c(t) = 1, V_i(t)] &= E[\delta_{ik}(t) \mid dN_{i\cdot}(t) = 1, \xi_i(t) = 0, V_i(t)] \\ &= E[\delta_{ik}(t) \mid dN_{i\cdot}(t) = 1, V_i(t)] \\ &= E[\delta_{ik}(t) \mid dN_{i\cdot}(t) = 1, \xi_i(t) = 1, V_i(t)], \end{aligned}$$

which implies that

$$E[\delta_{ik}(t) \mid dN_{i\cdot}^c(t) = 1, V_i(t)]$$

can be estimated based on those events whose categories are known. Define

$$\pi_{ik}(t) = E[\delta_{ik}(t) \mid dN_{i\cdot}(t) = 1, V_i(t)].$$

We posit a parametric model $\pi_{ik}(t;\eta_0)$ for $\pi_{ik}(t)$. As discussed in Schaubel and Cai [19, 20], we propose to fit $\pi_{ik}(t;\eta_0)$ by the following generalized logistic model:

$$\log\left\{\frac{\pi_{ik}(t;\eta_0)}{\pi_{i1}(t;\eta_0)}\right\} = \eta_0^{\mathrm{T}}V_{ik}(t), \quad k = 2, \ldots, K, \tag{6}$$

where the covariate $V_{ik}(t)$ contains the elements of $V_i(t)$ which pertain to category k, and $k = 1$ is arbitrarily selected as the reference category. Although the logistic regression is widely used, other parametric forms may be possible, such as the generalized probit model. Let $\hat{\eta}$ be the solution to the following estimating equation

$$\sum_{i=1}^{n}\sum_{k=2}^{K}\int_0^{\tau} V_{ik}(t)[\delta_{ik}(t) - \pi_{ik}(t;\eta)]\xi_i(t)dN_{i\cdot}(t) = 0. \tag{7}$$

It is easy to show that $\hat{\eta}$ is a consistent estimator of η_0. Then the event category probabilities can be estimated by

$$\pi_{ik}(t;\hat{\eta}) = \frac{\exp\{\hat{\eta}^{\mathrm{T}}V_{ik}(t)\}}{\sum_{l=1}^{K}\exp\{\hat{\eta}^{\mathrm{T}}V_{il}(t)\}}, \quad k = 1, \ldots, K,$$

where $V_{i1}(t) = 0$. Let

$$dM_{ik}(t,\theta,\eta) = \xi_i(t)dN_{ik}(t) + \pi_{ik}(t,\eta)dN_{i\cdot}^c(t) - Y_{ik}(t)\{g(\beta^{\mathrm{T}}W_{ik}(t))dt + h(\gamma^{\mathrm{T}}X_{ik}(t))d\mu_{0k}(t)\}.$$

If (1) and (6) are correctly specified, we have $E[dM_{ik}(t,\theta_0,\eta_0)] = 0$. Thus by applying the generalized estimating equation approach [7] and exploiting the consistency of $\hat{\eta}$, we specify the following estimating equations for $\mu_{0k}(t)$ and θ_0, respectively:

$$\sum_{i=1}^{n}\int_0^{t} dM_{ik}(s,\theta,\hat{\eta}) = 0, \quad 0 \leqslant t \leqslant \tau, \tag{8}$$

$$\sum_{i=1}^{n}\sum_{k=1}^{K}\int_0^{\tau} Z_{ik}(s)dM_{ik}(s,\theta,\hat{\eta})=0. \tag{9}$$

It follows from (8) that

$$\hat{\mu}_{0k}(t,\theta,\hat{\eta})=\sum_{i=1}^{n}\int_0^{t}\frac{\xi_i(s)dN_{ik}(s)+\pi_{ik}(s,\hat{\eta})dN_{i.}^{c}(s)-Y_{ik}(s)g(\beta^{\mathrm{T}}W_{ik}(s))ds}{\sum_{i=1}^{n}Y_{ik}(s)h(\gamma^{\mathrm{T}}X_{ik}(s))}. \tag{10}$$

Substituting (10) into (9), we get a weighted estimating equation for θ_0:

$$\begin{aligned}U(\theta)=\sum_{i=1}^{n}\sum_{k=1}^{K}\int_0^{\tau}&\{Z_{ik}(s)-\bar{Z}_k(s,\gamma)\}\{\xi_i(s)dN_{ik}(s)+\pi_{ik}(s,\hat{\eta})dN_{i.}^{c}(s)\\&-Y_{ik}(s)g(\beta^{T}W_{ik}(s))ds\}=0.\end{aligned} \tag{11}$$

The standard Newton-Raphson algorithm can be used to solve the above equation. In general, the initial value of θ is chosen to be 0, such as in the simulation studies below, and the Newton-Raphson algorithm always converges for the situation considered here. Let $\hat{\theta}$ be the solution to (11), and the corresponding baseline mean function estimator is given by $\hat{\mu}_{0k}(t)=\hat{\mu}_{0k}(t,\hat{\theta},\hat{\eta})$. By the law of large numbers and the consistency of $\hat{\eta}$, we show in the Appendix that under some regularity conditions, $\hat{\theta}$ is unique and consistent in a small neighborhood of θ_0.

3 Asymptotic properties

In this section, we will establish the asymptotic properties of the proposed parameter estimators in the setting where event categories are missing at random. Let $\dot{g}(x)=dg(x)/dx$ and $\dot{h}(x)=dh(x)/dx$. The following regularity conditions are assumed to hold for $i=1,\dots,n$ and $k=1,\dots,K$.

(C1) $\{N_{ik}(\cdot),Y_{ik}(\cdot),Z_{ik}(\cdot)\}_{k=1}^{K}$ are independent and identically distributed for $i=1,\dots,n$.

(C2) $P(C_{ik}\geqslant\tau)>0$, and $N_{ik}(\tau)<M<\infty$ almost surely.

(C3) $Z_{ik}(t)$ and $V_{ik}(t)$ are almost surely of bounded variation on $[0,\tau]$.

(C4) g is nonnegative and $h(\gamma^{\mathrm{T}}X_{ik}(t))$ is locally bounded away from 0 for γ in a small neighborhood of γ_0; g and h are twice continuously differentiable.

(C5) Matrices $\Omega(\eta_0)$ and $A=A(\theta_0)$ are nonsingular, where

$$\Omega(\eta)=E\bigg\{\sum_{k=1}^{K}\int_0^{\tau}V_{1k}(t)\pi_{1k}(t;\eta)\bigg\{V_{1k}(t)-\sum_{l=1}^{K}V_{1l}(t)\pi_{1l}(t;\eta)\bigg\}^{\mathrm{T}}\xi_1(t)dN_{1.}(t)\bigg\},$$

$$A(\theta)=E\bigg\{\sum_{k=1}^{K}\int_0^{\tau}Y_{1k}(t)\{Z_{1k}(t)-\bar{z}_k(t,\gamma)\}\begin{bmatrix}\dot{g}(\beta^{\mathrm{T}}W_{1k}(t))W_{1k}(t)dt\\ \dot{h}(\gamma^{\mathrm{T}}X_{1k}(t))X_{1k}(t)d\mu_{0k}(t)\end{bmatrix}^{\mathrm{T}}\bigg\},$$

and $\bar{z}_k(t,\gamma)$ is the limit of $\bar{Z}_k(t,\gamma)$.

The asymptotic properties of $\hat{\theta}$ are summarized in the following theorem with the proof outlined in Appendix.

Theorem 1. *Under the regularity conditions* (C1)–(C5), *$\hat{\theta}$ is strongly consistent to θ_0, and $n^{1/2}(\hat{\theta}-\theta_0)$ is asymptotically normal with mean zero and covariance matrix $A^{-1}\Sigma(A^{-1})^{\mathrm{T}}$, where*

$$\Sigma=E\bigg[\bigg(\sum_{k=1}^{K}\Phi_{ik}(\theta_0,\eta_0)\bigg)^{\otimes2}\bigg],$$

$$\Phi_{ik}(\theta,\eta)=\int_0^{\tau}\{Z_{ik}(t)-\bar{z}_k(t,\gamma)\}dM_{ik}(t,\theta,\eta)+\Psi_k(\gamma,\eta)\Omega(\eta)^{-1}\sum_{l=1}^{K}\Gamma_{il}(\eta),$$

$$\Gamma_{ik}(\eta)=\int_0^{\tau}V_{ik}(t)\{\delta_{ik}(t)-\pi_{ik}(t,\eta)\}\xi_i(t)dN_{i.}(t),$$

$$\hat{\Psi}_k(\gamma,\eta) = n^{-1}\sum_{i=1}^{n}\int_0^{\tau}\{Z_{ik}(t)-\bar{Z}_k(t,\gamma)\}\bigg\{V_{ik}(t)-\sum_{l=1}^{K}V_{il}(t)\pi_{il}(t,\eta)\bigg\}^{\mathrm{T}}\pi_{ik}(t,\eta)dN_{i.}^{c}(t),$$

$\Psi_k(\gamma,\eta)$ is limit of $\hat{\Psi}_k(\gamma,\eta)$ and $a^{\otimes 2}=aa^{\mathrm{T}}$ for any vector a.

The covariance matrix $A^{-1}\Sigma(A^{-1})^{\mathrm{T}}$ can be consistently estimated by $\hat{A}^{-1}\hat{\Sigma}(\hat{A}^{-1})^{\mathrm{T}}$, where $\hat{\Sigma} = n^{-1}\sum_{i=1}^{n}(\sum_{k=1}^{K}\hat{\Phi}_{ik}(\hat{\theta},\hat{\eta}))^{\otimes 2}$,

$$\hat{A} = n^{-1}\sum_{i=1}^{n}\sum_{k=1}^{K}\int_0^{\tau}Y_{ik}(t)\{Z_{ik}(t)-\bar{Z}_k(t,\hat{\gamma})\}\begin{bmatrix}\dot{g}(\hat{\beta}^{\mathrm{T}}W_{ik}(t))W_{ik}(t)dt\\ \dot{h}(\hat{\gamma}^{\mathrm{T}}X_{ik}(t))X_{ik}(t)d\hat{\mu}_{0k}(t)\end{bmatrix}^{\mathrm{T}},$$

$$\hat{\Phi}_{ik}(\theta,\eta) = \int_0^{\tau}\{Z_{ik}(t)-\bar{Z}_k(t,\gamma)\}d\hat{M}_{ik}(t,\theta,\eta)+\hat{\Psi}_k(\gamma,\eta)\hat{\Omega}(\eta)^{-1}\sum_{l=1}^{K}\Gamma_{il}(\eta),$$

$$\hat{\Omega}(\eta) = n^{-1}\sum_{i=1}^{n}\sum_{k=1}^{K}\int_0^{\tau}V_{ik}(t)\pi_{ik}(t,\eta)\bigg\{V_{ik}(t)-\sum_{l=1}^{K}V_{il}(t)\pi_{il}(t,\eta)\bigg\}^{\mathrm{T}}\xi_i(t)dN_{i.}(t),$$

and $d\hat{M}_{ik}(t,\theta,\eta)=\xi_i(t)dN_{ik}(t)+\pi_{ik}(t,\eta)dN_{i.}^{c}(t)-Y_{ik}(t)\{g(\beta^{\mathrm{T}}W_{ik}(t))dt+h(\gamma^{\mathrm{T}}X_{ik}(t))d\hat{\mu}_{0k}(t,\theta,\eta)\}$.

Define

$$S_k^0(t,\gamma)=n^{-1}\sum_{i=1}^{n}Y_{ik}(t)h(\gamma^{\mathrm{T}}X_{ik}(t)),\quad S_k^W(t,\beta)=n^{-1}\sum_{i=1}^{n}Y_{ik}(t)\dot{g}(\beta^{\mathrm{T}}W_{ik}(t))W_{ik}(t),$$

$$S_k^X(t,\gamma)=n^{-1}\sum_{i=1}^{n}Y_{ik}(t)\dot{h}(\gamma^{\mathrm{T}}X_{ik}(t))X_{ik}(t),$$

$$\bar{W}_k(t,\theta)=S_k^W(t,\beta)/S_k^0(t,\gamma),\quad \bar{X}_k(t,\gamma)=S_k^X(t,\gamma)/S_k^0(t,\gamma),$$

and

$$D_k(t,\theta)=-\begin{bmatrix}\int_0^t\bar{w}_k(s,\theta)ds\\ \int_0^t\bar{x}_k(s,\gamma)d\mu_{0k}(s)\end{bmatrix},$$

where $s_k^0(t,\gamma)$, $s_k^w(t,\beta)$ and $s_k^x(t,\gamma)$ are the limits of $S_k^0(t,\gamma)$, $S_k^W(t,\beta)$ and $S_k^X(t,\gamma)$, respectively; $\bar{w}_k(t,\theta)$ $=s_k^w(t,\beta)/s_k^0(t,\gamma)$ and $\bar{x}_k(t,\gamma)=s_k^x(t,\gamma)/s_k^0(t,\gamma)$. The limiting results for $\hat{\mu}_{0k}(t)$ are given in the next theorem.

Theorem 2. *$\hat{\mu}_{0k}(t)$ converges almost surely to $\mu_{0k}(t)$ uniformly in $t\in[0,\tau]$. Furthermore, $n^{1/2}\{\hat{\mu}_{0k}(t)$ $-\mu_{0k}(t)\}$ converges weakly to a zero-mean Gaussian process with covariance function at (s,t) equal to $\omega_k(s,t)=E[\phi_{ik}(s,\theta_0,\eta_0)\phi_{ik}(t,\theta_0,\eta_0)]$, where*

$$\begin{aligned}\phi_{ik}(t,\theta_0,\eta_0) &= D_k(t,\theta_0)^{\mathrm{T}}A^{-1}\sum_{l=1}^{K}\Phi_{il}(\theta_0,\eta_0)+B_k(t,\gamma_0,\eta_0)\Omega(\eta_0)^{-1}\sum_{l=1}^{K}\Gamma_{il}(\eta_0)\\ &\quad+\int_0^t s_k^0(s,\gamma_0)^{-1}dM_{ik}(s,\theta_0,\eta_0),\\ \hat{B}_k(t,\gamma,\eta) &= n^{-1}\sum_{i=1}^{n}\int_0^t S_k^0(s,\gamma)^{-1}\pi_{ik}(s,\eta)\bigg\{V_{ik}(s)-\sum_{l=1}^{K}V_{il}(t)\pi_{il}(s,\eta)\bigg\}^{\mathrm{T}}dN_{i.}^{c}(s),\end{aligned}$$

and $B_k(t,\gamma,\eta)$ is the limit of $\hat{B}_k(t,\gamma,\eta)$.

The covariance function $\omega_k(s,t)$ can be consistently estimated by $\hat{\omega}_k(s,t)=n^{-1}\sum_{i=1}^{n}\hat{\phi}_{ik}(s)\hat{\phi}_{ik}(t)$, where

$$\hat{\phi}_{ik}(t)=\hat{D}_k(t,\hat{\theta})^{\mathrm{T}}\hat{A}^{-1}\sum_{l=1}^{K}\hat{\Phi}_{il}(\hat{\theta},\hat{\eta})+\hat{B}_k(t,\hat{\gamma},\hat{\eta})\hat{\Omega}(\hat{\eta})^{-1}\sum_{l=1}^{K}\Gamma_{il}(\hat{\eta})+\int_0^t S_k^0(s,\hat{\gamma})^{-1}d\hat{M}_{ik}(s,\hat{\theta},\hat{\eta})$$

and

$$\hat{D}_k(t,\hat{\theta})=-\begin{bmatrix}\int_0^t\bar{W}_k(s,\hat{\theta})ds\\ \int_0^t\bar{X}_k(s,\hat{\gamma})d\hat{\mu}_{0k}(s)\end{bmatrix}.$$

4 Model checking techniques

To assess the adequacy of (1), we propose a goodness-of-fit test statistic. Let $M_{ik}(t) = M_{ik}(t, \theta_0, \eta_0)$ and $\hat{M}_{ik}(t) = \hat{M}_{ik}(t, \hat{\theta}, \hat{\eta})$. Following Lin et al. [8], we consider the following cumulative sums of residuals:

$$\mathcal{L}_k(t, z) = n^{-1/2} \sum_{i=1}^{n} \int_0^t I(Z_{ik}(u) \leqslant z) d\hat{M}_{ik}(u),$$

where the event $I(Z_{ik}(u) \leqslant z)$ means that each component of $Z_{ik}(u)$ is bounded by the corresponding component of z. For the null distribution of $\mathcal{L}_k(t, z)$, we have the following theorem with the proof given in Appendix.

Theorem 3. *Under the regularity conditions listed in Section* 3*, the null distribution of* $\mathcal{L}_k(t, z)$ *converges weakly to a zero-mean Gaussian process with covariance function at* (t_1, z_1) *and* (t_2, z_2) *equal to* $E[\sigma_{ik}(t_1, z_1)\sigma_{ik}(t_2, z_2)]$*, where*

$$\begin{aligned}
\sigma_{ik}(t, z) &= \int_0^t \left[I(Z_{ik}(u) \leqslant z) - \frac{G_k(u, z)}{s_k^0(u, \gamma_0)} \right] dM_{ik}(u) + \{E_k(t, z) - F_k(t, z)\}\Omega(\eta_0)^{-1} \\
&\quad \times \sum_{l=1}^{K} \Gamma_{il}(\eta_0) - \{H_k(t, z) + L_k(t, z)\}^{\mathrm{T}} A^{-1} \sum_{l=1}^{K} \Phi_{il}(\theta_0, \eta_0), \\
E_k(t, z) &= E\left\{ \int_0^t I(Z_{ik}(u) \leqslant z) \left[V_{ik}(u) - \sum_{l=1}^{K} \pi_{il}(u, \eta_0) V_{il}(u) \right]^{\mathrm{T}} \pi_{ik}(u, \eta_0) dN_{i.}^c(u) \right\}, \\
L_{1k}(t, z) &= E\{Y_{ik}(t) I(Z_{ik}(t) \leqslant z) \dot{g}(\beta_0^{\mathrm{T}} W_{ik}(t)) W_{ik}(t)\}, \\
L_{2k}(t, z) &= E\{Y_{ik}(t) I(Z_{ik}(t) \leqslant z) \dot{h}(\gamma_0^{\mathrm{T}} X_{ik}(t)) X_{ik}(t)\}, \\
G_k(t, z) &= E\{Y_{ik}(t) h(\gamma_0^{\mathrm{T}} X_{ik}(t)) I(Z_{ik}(t) \leqslant z)\}, \\
L_k(t, z) &= \begin{bmatrix} \int_0^t L_{1k}(u, z) du \\ \int_0^t L_{2k}(u, z) d\mu_{0k}(u) \end{bmatrix},
\end{aligned}$$

and

$$H_k(t, z) = \int_0^t G_k(u, z) dD_k(u, \theta_0), \quad F_k(t, z) = \int_0^t G_k(u, z) dB_k(u, \gamma_0, \eta_0).$$

It follows from Theorem 3 that the null distribution of $\mathcal{L}_k(t, z)$ can be approximated by the zero-mean Gaussian process $\tilde{\mathcal{L}}_k(t, z) = n^{-1/2} \sum_{i=1}^{n} \hat{\Upsilon}_{ik}(t, z)$, where

$$\begin{aligned}
\hat{\Upsilon}_{ik}(t, z) &= \int_0^t \left[I(Z_{ik}(u) \leqslant z) - \frac{\hat{G}_k(u, z)}{S_k^0(u, \hat{\gamma})} \right] d\hat{M}_{ik}(u) + \{\hat{E}_k(t, z) - \hat{F}_k(t, z)\}\hat{\Omega}(\hat{\eta})^{-1} \\
&\quad \times \sum_{l=1}^{K} \Gamma_{il}(\hat{\eta}) - \{\hat{H}_k(t, z) + \hat{L}_k(t, z)\}^{\mathrm{T}} \hat{A}^{-1} \sum_{l=1}^{K} \hat{\Phi}_{il}(\hat{\theta}, \hat{\eta}), \\
\hat{E}_k(t, z) &= n^{-1} \sum_{i=1}^{n} \int_0^t I(Z_{ik}(u) \leqslant z) \left[V_{ik}(u) - \sum_{l=1}^{K} \pi_{il}(u, \hat{\eta}) V_{il}(u) \right]^{\mathrm{T}} \pi_{ik}(u, \hat{\eta}) dN_{i.}^c(u), \\
\hat{L}_{1k}(t, z) &= n^{-1} \sum_{i=1}^{n} Y_{ik}(t) I(Z_{ik}(t) \leqslant z) \dot{g}(\hat{\beta}^{\mathrm{T}} W_{ik}(t)) W_{ik}(t), \\
\hat{L}_{2k}(t, z) &= n^{-1} \sum_{i=1}^{n} Y_{ik}(t) I(Z_{ik}(t) \leqslant z) \dot{h}(\hat{\gamma}^{\mathrm{T}} X_{ik}(t)) X_{ik}(t), \\
\hat{G}_k(t, z) &= n^{-1} \sum_{i=1}^{n} Y_{ik}(t) h(\hat{\gamma}^{\mathrm{T}} X_{ik}(t)) I(Z_{ik}(t) \leqslant z),
\end{aligned}$$

$$\hat{L}_k(t,z) = \begin{bmatrix} \int_0^t \hat{L}_{1k}(u,z)du \\ \int_0^t \hat{L}_{2k}(u,z)d\hat{\mu}_{0k}(u) \end{bmatrix},$$

and

$$\hat{H}_k(t,z) = \int_0^t \hat{G}_k(u,z)d\hat{D}_k(u,\hat{\theta},\hat{\eta}), \quad \hat{F}_k(t,z) = \int_0^t \hat{G}_k(u,z)d\hat{B}_k(u,\hat{\gamma},\hat{\eta}).$$

It is difficult to estimate the asymptotic covariance function of $\mathcal{L}_k(t,z)$ analytically because the limiting process of $\mathcal{L}_k(t,z)$ does not have an independent increment structure. To this end, we can appeal to the resampling approach [8, 9]. Let $\{G_1,\ldots,G_n\}$ be independent standard normal variables which are independent of the observed data. It can be showed that the null distribution of $\mathcal{L}_k(t,z)$ can be approximated by the conditional distribution of $\hat{\mathcal{L}}_k(t,z)$, where

$$\hat{\mathcal{L}}_k(t,z) = n^{-1/2}\sum_{i=1}^n \hat{\Upsilon}_{ik}(t,z)G_i. \tag{12}$$

Thus, we can obtain a large number of realizations of $\hat{\mathcal{L}}_k(t,z)$ by repeatedly generating the standard normal random sample $\{G_1,\ldots,G_n\}$ while fixing the observed data. To check the adequacy of (1), we could plot $\mathcal{L}_k(t,z)$ along with a few realizations of $\hat{\mathcal{L}}_k(t,z)$ to see if there exist some unusual patterns. Since $\mathcal{L}_k(t,z)$ is expected to fluctuate randomly around 0 under the assumed model, a formal lack-of-fit test can be constructed based on the supremum test statistic $\sup_{0\leqslant t\leqslant \tau,z}|\mathcal{L}_k(t,z)|$, with which the p-value can be obtained by comparing the observed value of $\sup_{0\leqslant t\leqslant \tau,z}|\mathcal{L}_k(t,z)|$ to a large number of realizations from $\sup_{0\leqslant t\leqslant \tau,z}|\hat{\mathcal{L}}_k(t,z)|$.

5 Simulation studies

In this section, simulation studies were conducted to examine the finite-sample properties of the proposed estimators. We considered the setting where there exist $K=2$ event categories. Two covariates W_i and X_i are generated from Bernoulli distribution with success probability 0.5 and uniform distribution on $(0,1)$, respectively. To induce positive correlation among the within-subject events, a frailty variable R_i was introduced, following a gamma distribution with mean $E[R_i]=1$ and variances $\sigma^2=0, 0.25, 0.5$. To avoid yielding too many recurrent events for one subject, we set $R_i^*=\min(R_i,1)$. The kth recurrent event times for subject i were generated from a Poisson process with intensity function

$$\lambda_{ik}(t) = R_i^*\{\beta_k W_i + \exp(\gamma_k X_i)\lambda_{0k}\},$$

where $(\beta_1,\beta_2,\gamma_1,\gamma_2)=(0.5,0.3,0.5,1)$, $\lambda_{01}=0.5$ and $\lambda_{02}=0.625$. The censoring times $C_{ik}=C_i$ were generated from $U(0,\tau)$ with $\tau=5$. Under the above settings, the average number of observed events per subject was approximately 2 for $k=1$ and that was approximately 3 for $k=2$. The event categories were set to missing with probability

$$P\{\xi_i(t)=0 \mid dN_{i\cdot}(t)=1,\tilde{Z}_i(t)\} = \frac{\exp(\alpha^{\mathrm{T}}\tilde{Z}_i(t))}{1+\exp(\alpha^{\mathrm{T}}\tilde{Z}_i(t))},$$

where $\tilde{Z}_i(t)=(1,t,N_{i\cdot}(t-),W_i,X_i)^{\mathrm{T}}$ with $N_{i\cdot}(t-)$ counting the total number of events before t for subject i. In this study, we set $\alpha=(\alpha_0,\alpha_t,\alpha_{NZ})^{\mathrm{T}}$, with $\alpha_0=-1$, $\alpha_t=-0.2$ and $\alpha_{NZ}=(\alpha_N,\alpha_W,\alpha_X)$ $=(0,0,0),(0.1,0.2,0.3)$ or $(0.1,0.5,1)$. The percentage of events with missing categories ranged from 21% to 48%.

For comparison, three methods were employed to estimate regression parameters: (i) the full-data (FF) analysis, which is based on the data with all event types being known; (ii) the complete-case (CC) method, which excludes the events with missing types; (iii) the proposed weighted estimating equations (WEE) method. The FF analysis was conducted before setting event types to missing. For the WEE method, the event category probability was fitted by the following logistic model: $\log\{\frac{\pi_{i2}(t,\eta_0)}{\pi_{i1}(t,\eta_0)}\}=\eta_0^{\mathrm{T}}V_{i2}(t)$, where $V_{i2}(t)=[1,t,N_{i\cdot}(t-),W_i,X_i]^{\mathrm{T}}$. The results presented below are based on 500 replications with sample sizes $n=100$ and 200.

Table 1 Comparison results on estimation of $\beta_1 = 0.5$

			Bias				MSE			
n	α_{NZ}	σ^2	FF	WEE	WEE*	CC	FF	WEE	WEE*	CC
100	(0,0,0)	0	0.004	0.004	−0.008	0.004	0.015	0.018	0.017	0.019
		0.250	0.004	0.004	0.003	0.004	0.015	0.017	0.016	0.018
		0.500	−0.007	−0.008	−0.019	−0.009	0.014	0.017	0.017	0.018
	(0.1,0.2,0.3)	0	0.004	0.002	−0.005	−0.088	0.015	0.020	0.017	0.029
		0.250	0.004	0.005	−0.004	−0.086	0.015	0.018	0.018	0.027
		0.500	−0.007	−0.008	−0.017	−0.094	0.014	0.019	0.017	0.029
	(0.1,0.5,1)	0	0.004	0.002	−0.010	−0.229	0.015	0.023	0.020	0.081
		0.250	0.004	0.007	−0.005	−0.223	0.015	0.021	0.019	0.074
		0.500	−0.007	−0.006	−0.016	−0.223	0.014	0.022	0.018	0.075
200	(0,0,0)	0	−0.003	−0.002	−0.010	−0.002	0.007	0.009	0.008	0.010
		0.250	0.003	0.002	−0.002	0.002	0.007	0.008	0.008	0.009
		0.500	−0.019	−0.016	−0.026	−0.015	0.008	0.010	0.010	0.010
	(0.1,0.2,0.3)	0	−0.003	−0.004	−0.011	−0.092	0.007	0.009	0.008	0.020
		0.250	0.003	0.002	−0.008	−0.088	0.007	0.009	0.008	0.018
		0.500	−0.019	−0.016	−0.024	−0.101	0.009	0.011	0.009	0.021
	(0.1,0.5,1)	0	−0.003	−0.005	−0.011	−0.232	0.007	0.012	0.009	0.069
		0.250	0.003	0.001	−0.011	−0.229	0.007	0.011	0.008	0.065
		0.500	−0.019	−0.017	−0.015	−0.237	0.009	0.014	0.009	0.071

All the simulation results are summarized in Tables 1–4. In these tables, Bias is the sample mean of the estimate minus the true value; ESD is the empirical standard deviation of the estimate; ASE is the average estimated standard error; MSE is the sample mean of the squared errors between the estimate and the true value; and CP is the 95% empirical coverage probability based on the normal approximation.

The estimation results of $\beta_1 = 0.5$ and $\gamma_1 = 0.5$ are summarized in Tables 1 and 2, similar results are obtained (not reported here) for β_2 and γ_2. Both Tables 1 and 2 show that the CC estimators are nearly unbiased only when $\alpha_{NZ} = (0, 0, 0)$ (i.e., the event types are missing completely at random). However, the CC estimators have large biases when α_{NZ} is not equal to zero and the bias of the CC estimator increases as the correlation between the missingness probability and covariates increases. Both the FF and WEE estimators are essentially unbiased in all settings. Furthermore, the WEE estimators are more efficient than the CC estimators, and are only slightly less efficient than the FF estimators. In addition, as pointed out by the associate editor, we try to fit the event category probability by the probit model for robustness check, and the corresponding simulation results are also presented in the WEE* columns of Tables 1 and 2. It can be seen that although the probit model approach gives unbiased estimates for β_1, it leads to biased estimates for γ_1 unless the missing mechanism is MCAR ($\alpha_{NZ} = 0$); while the proposed logistic model approach is valid under MAR and is robust.

Tables 3 and 4 give the accuracy of asymptotic approximation to the distributions of the regression parameter estimators. It can be seen that there is a good agreement between the average estimated standard errors and empirical standard deviations, and the 95% empirical coverage probabilities are reasonable. The performance of the WEE estimators becomes better when the sample size increases from 100 to 200. We also considered other setups and the results were similar to those given above.

6 An application

Now we apply the proposed method to a set of bivariate recurrent event data from a platelet transfusion reaction study on hematology/oncology patients. In this study, all patients may experience different febrile nonhemolytic transfusion reactions (FNHTRs) defined as the presence of fever, chills, rigors, hives,

Table 2 Comparison results on estimation of $\gamma_1 = 0.5$

			Bias				MSE			
n	α_{NZ}	σ^2	FF	WEE	WEE*	CC	FF	WEE	WEE*	CC
100	(0,0,0)	0	0.027	0.018	0.040	0.023	0.120	0.136	0.120	0.160
		0.250	0.017	0.003	0.016	0.007	0.133	0.152	0.133	0.164
		0.500	−0.014	−0.017	0.034	−0.015	0.131	0.152	0.135	0.168
	(0.1,0.2,0.3)	0	0.027	0.019	0.077	−0.253	0.120	0.159	0.131	0.212
		0.250	0.017	−0.001	0.086	−0.255	0.133	0.164	0.133	0.211
		0.500	−0.014	−0.034	0.051	−0.292	0.131	0.168	0.150	0.234
	(0.1,0.5,1)	0	0.027	0.019	0.078	−0.713	0.120	0.195	0.142	0.669
		0.250	0.017	−0.001	0.122	−0.716	0.133	0.191	0.158	0.666
		0.500	−0.014	−0.035	0.065	−0.758	0.131	0.206	0.136	0.731
200	(0,0,0)	0	−0.005	−0.003	0.027	−0.006	0.060	0.067	0.060	0.072
		0.250	−0.004	0.001	0.015	0.000	0.060	0.067	0.064	0.074
		0.500	0.003	0.003	0.025	0.008	0.064	0.072	0.064	0.081
	(0.1,0.2,0.3)	0	−0.005	−0.001	0.058	−0.264	0.060	0.073	0.063	0.137
		0.250	−0.004	−0.007	0.071	−0.262	0.060	0.074	0.066	0.142
		0.500	0.003	0.006	0.075	−0.251	0.064	0.078	0.073	0.135
	(0.1,0.5,1)	0	−0.005	0.001	0.094	−0.719	0.060	0.090	0.074	0.587
		0.250	−0.004	0.002	0.104	−0.712	0.060	0.086	0.073	0.583
		0.500	0.003	0.008	0.095	−0.712	0.064	0.093	0.080	0.580

Table 3 Simulation results for the accuracy of asymptotic approximation to the distributions of estimators for β_1 and β_2 using the WEE method

			β_1			β_2		
n	α_{NZ}	σ^2	ASE	ESD	CP	ASE	ESD	CP
100	(0,0,0)	0	0.129	0.133	0.940	0.147	0.150	0.948
		0.250	0.128	0.129	0.944	0.146	0.147	0.954
		0.500	0.129	0.129	0.948	0.149	0.152	0.946
	(0.1,0.2,0.3)	0	0.137	0.142	0.942	0.154	0.160	0.948
		0.250	0.135	0.135	0.944	0.153	0.158	0.940
		0.500	0.137	0.136	0.956	0.156	0.161	0.946
	(0.1,0.5,1)	0	0.148	0.152	0.938	0.165	0.169	0.948
		0.250	0.148	0.145	0.952	0.164	0.167	0.932
		0.500	0.148	0.149	0.948	0.166	0.173	0.942
200	(0,0,0)	0	0.091	0.092	0.964	0.104	0.106	0.946
		0.250	0.091	0.091	0.940	0.104	0.112	0.920
		0.500	0.091	0.096	0.930	0.106	0.103	0.944
	(0.1,0.2,0.3)	0	0.096	0.097	0.946	0.109	0.109	0.948
		0.250	0.097	0.096	0.952	0.109	0.118	0.920
		0.500	0.097	0.104	0.940	0.111	0.110	0.956
	(0.1,0.5,1)	0	0.105	0.108	0.940	0.117	0.116	0.956
		0.250	0.106	0.105	0.946	0.117	0.130	0.914
		0.500	0.105	0.116	0.914	0.118	0.123	0.946

and other recorded symptoms developed within 4–6 hours post transfusion. The data were collected at five university teaching hospitals in Toronto coded A-E over three consecutive summers from 1996–

Table 4 Simulation results for the accuracy of asymptotic approximation to the distributions of estimators for γ_1 and γ_2 using the WEE method

			γ_1			γ_2		
n	α_{NZ}	σ^2	ASE	ESD	CP	ASE	ESD	CP
100	(0,0,0)	0	0.355	0.368	0.952	0.254	0.260	0.936
		0.250	0.354	0.390	0.920	0.253	0.262	0.954
		0.500	0.367	0.390	0.942	0.266	0.272	0.940
	(0.1,0.2,0.3)	0	0.376	0.398	0.944	0.267	0.276	0.946
		0.250	0.377	0.405	0.926	0.265	0.276	0.946
		0.500	0.390	0.408	0.950	0.278	0.283	0.930
	(0.1,0.5,1)	0	0.411	0.441	0.934	0.285	0.299	0.936
		0.250	0.414	0.437	0.944	0.283	0.298	0.946
		0.500	0.426	0.453	0.946	0.295	0.309	0.942
200	(0,0,0)	0	0.246	0.258	0.944	0.179	0.175	0.948
		0.250	0.249	0.259	0.956	0.178	0.162	0.964
		0.500	0.257	0.269	0.942	0.189	0.191	0.954
	(0.1,0.2,0.3)	0	0.261	0.271	0.938	0.187	0.186	0.936
		0.250	0.264	0.272	0.948	0.186	0.170	0.970
		0.500	0.272	0.280	0.948	0.197	0.198	0.962
	(0.1,0.5,1)	0	0.285	0.300	0.948	0.200	0.200	0.940
		0.250	0.288	0.293	0.944	0.200	0.188	0.968
		0.500	0.297	0.305	0.954	0.209	0.212	0.948

1998 [15]. We considered a subset of the data which included 254 patients who were followed up during the 1997 summer and a total of 1395 transfusions were recorded. The observed number of transfusions per patient ranges from 1 (77 patients) to 50 (1 patient). Since the occurrence of FNHTRs is temporary, it is reasonable to treat a reaction as a recurrent event. There are 1201 transfusions eligible for our analysis and 314 transfusion reactions being observed. The mean number of reactions per patient is 1.3 (sd $=1.8$) and the median follow-up time is 8 days. In the following analysis, for simplicity, we will classify all reactions into two types: The reaction accompanied with fever (denoted by Type I reaction) and the reaction with no fever (denoted by Type II reaction). Based on the above classification, among the 314 observed transfusion reactions, there were 181 Type I reactions, 115 Type II reactions, and 18 reactions with missing types. Thus we get bivariate recurrent event data in the presence of missing event types.

Following Zhao et al. [27], we let $N_{i1}^*(t)$ and $N_{i2}^*(t)$ denote the numbers of Types I and II reactions which had occurred over interval $[0,t]$ for subject i, respectively. Also we define the covariate W_i as the gender of patient i (1 if female, 0 if male), and X_i as the age of patient i when entering the study (1 if age in $(0,27]$, 2 if age in $(27,42]$, 3 if age in $(42,55]$, 4 if age greater than 55). Let τ be the largest follow-up time (164 days). Our goal is to estimate the effects of gender and age on the risk of the two types of transfusion reactions. In the interest of flexibility, both covariates are assumed to be type-specific. We propose the following four models to fit the data: The additive gender effect and multiplicative age effect model (denoted by AMM1) $E\{dN_{ik}^*(t) \mid W_i, X_i\} = \beta_k W_i dt + \exp(\gamma_k X_i) d\mu_{0k}(t)$; the additive age effect and multiplicative gender effect model (denoted by AMM2) $E\{dN_{ik}^*(t) \mid W_i, X_i\} = \gamma_k X_i dt + \exp(\beta_k W_i) d\mu_{0k}(t)$; the additive rates model (denoted by AM) $E\{dN_{ik}^*(t) \mid W_i, X_i\} = (\beta_k W_i + \gamma_k X_i)dt + d\mu_{0k}(t)$; the multiplicative rates model (denoted by MM) $E\{dN_{ik}^*(t) \mid W_i, X_i\} = \exp(\beta_k W_i + \gamma_k X_i) d\mu_{0k}(t)$, for $k=1,2$. The event category probability was modeled by the following logistic regression $\pi_{ik}(t;\eta_0) = \frac{\exp\{\eta_0^{\mathrm{T}} V_{ik}\}}{\sum_{l=1}^2 \exp\{\eta_0^{\mathrm{T}} V_{il}\}}$, $k=1,2$, where $V_{i2} = (1, W_i, X_i)^{\mathrm{T}}$ and $V_{i1} = 0$.

Table 5 gives the estimates of covariate effects with the corresponding standard error estimates in parentheses using the above four models. All results show that neither FNHTR rates seems to be correlated with the gender of the patients, and older patients may have higher risk of Type II platelet transfusion

Table 5 Analysis results for the FNHTRs data

	Type I		Type II			
Model	$\hat{\beta}_1$	$\hat{\gamma}_1$	$\hat{\beta}_2$	$\hat{\gamma}_2$	AIC	BIC
AMM1	0.0025 (0.0097)	0.0548 (0.0809)	0.0003 (0.0072)	0.2354 (0.1009)	0.7837	0.8413
AMM2	0.0531 (0.2252)	0.0028 (0.0040)	−0.0265 (0.5270)	0.0075 (0.0034)	0.7835	0.8412
AM	0.0023 (0.0097)	0.0027 (0.0040)	−0.0004 (0.0072)	0.0075 (0.0034)	0.7837	0.8413
MM	0.0505 (0.1956)	0.0544 (0.0791)	0.0103 (0.2277)	0.2345 (0.0976)	0.7839	0.8415

Note. Type I denotes the reaction with fever. Type II denotes the reaction with no fever.

reactions. These results were similar to those obtained by Zhao et al. [27]. In order to examine which model fits the data best, we employ the Akaike's information criterion (AIC) and Bayesian information criterion (BIC) to evaluate the performance of the four models, where

$$\mathrm{AIC} = 2(p+q)/n + \log(\mathrm{RSS}/n), \quad \mathrm{BIC} = (p+q)\log(n)/n + \log(\mathrm{RSS}/n),$$

and RSS$= \sum_{i=1}^{n}\sum_{k=1}^{2} r_{ik}^2$ with r_{ik} being the residual of event Type k for subject i. The analysis results are also presented in Table 5. Under the measures of AIC and BIC, model AMM2 is the best among the four models, which indicates that the effect of age on the recurrence of FNHTRs is more likely to be additive while the effect of gender is more likely to be multiplicative.

Finally, we apply the model-checking technique developed in Section 4 to assess the adequacy of model AMM2 for the data. We calculated the statistics $\mathcal{L}_k(t,z)$ $(k=1,2)$, and obtained

$$\sup_{0\leqslant t\leqslant \tau, z} |\mathcal{L}_1(t,z)| = 0.7628 \quad \text{and} \quad \sup_{0\leqslant t\leqslant \tau, z} |\mathcal{L}_2(t,z)| = 0.4213$$

with p-values of 0.198 and 0.672, respectively, based on 500 realizations of the corresponding statistics $\sup_{0\leqslant t\leqslant \tau, z} |\hat{\mathcal{L}}_1(t,z)|$ and $\sup_{0\leqslant t\leqslant \tau, z} |\hat{\mathcal{L}}_2(t,z)|$. These results suggest that model AMM2 fits the data well.

7 Concluding remarks

In this paper, we proposed an additive-multiplicative rates model for multivariate recurrent event data with missing event categories under the MAR assumption. A weighted estimating equation approach was developed for parameter estimation, where weights are equal to the corresponding category-specific probabilities when event categories are missing. The resulting estimators were shown to be consistent and asymptotically normal. Simulation results indicated that the proposed methods perform well in finite samples, and a real-data example was provided.

For the additive-multiplicative rates model, a direct classification between additive and multiplicative effects can be done by hypothesis testing. In many applications, however, one may face a problem of the selection of $W_{ik}(t)$ and $X_{ik}(t)$. In general, based on some priori knowledge, the covariates anticipated to have a large impact on rate ratios should be added to the multiplicative part, and those which could have a large impact on absolute rates should be incorporated into the additive part. When there was little information on the underlying recurrent event processes, it would be desirable to develop some data-driven methods for the classification of covariates. For example, if covariates are of small dimension, we can fit all possible models based on different combinations of covariates, and choose the best fitted model based on some model selection criterion.

It would be worthwhile for us to analyze multivariate recurrent event data with missing event categories under other competing models, such as the semiparametric transformation model [26]. In addition, Lin et al. [10] developed fully nonparametric estimators of the event mean function in which the missingness mechanism is completely unspecified. We will consider similar nonparametric methods for regression analysis in the further research.

Acknowledgements This work was supported by National Natural Science Foundation of China (Grant Nos. 11231010, 11171330 and 11371299), Key Laboratory of Random Complex Structures and Data Science, Chinese Academy of Sciences (Grant No. 2008DP173182), Beijing Center for Mathematics and Information Interdisciplinary Sciences, the Research Grant Council of Hong Kong (Grant Nos. 504011 and 503513), and The Hong Kong Polytechnic University. The authors thank Dr. Bruce J. Patterson for providing them with example data. The authors thank the Associate Editor-in-Chief, Professor SHAO QiMan, the Associate Editor and the two reviewers for their constructive and insightful comments and suggestions that greatly improved the paper.

References

1 Andersen P K, Gill R D. Cox's regression model for counting processes: A large sample study. Ann Statist, 1982, 10: 1100–1120
2 Cai J, Schaubel D E. Marginal mean/rates models for multiple type recurrent event data. Lifetime Data Anal, 2004, 10: 121–138
3 Chen B E, Cook R J. The analysis of multivariate recurrent events with partially missing event types. Lifetime Data Anal, 2009, 15: 41–58
4 Chen X, Wang Q, Cai J, et al. Semiparametric additive marginal regression models for multiple type recurrent events. Lifetime Data Anal, 2012, 18: 504–527
5 Gao G, Tsiatis A A. Semiparametric estimators for the regression coefficients in the linear transformation competing risks model with missing cause of failure. Biometrika, 2005, 92: 875–891
6 Kalbfleisch J D, Prentice R L. The Statistical Analysis of Failure Time Data, 2nd ed. Hoboken: John Wiley & Sons, 2002
7 Liang K Y, Zeger S L. Longitudinal data analysis using generalized linear models. Biometrika, 1986, 73: 13–22
8 Lin D Y, Wei L J, Yang I, et al. Semiparametric regression for the mean and rate functions of recurrent events. J R Stat Soc Ser B Stat Methodol, 2000, 62: 711–730
9 Lin D Y, Wei L J, Ying Z. Checking the Cox model with cumulative sums of martingale-based residuals. Biometrika, 1993, 80: 557–572
10 Lin F, Cai J, Fine J P, et al. Nonparametric estimation of the mean function for recurrent event data with missing event category. Biometrika, 2013, 100: 727–740
11 Little R J A, Rubin D B. Statistical Analysis with Missing Data. New Hoboken: John Wiley & Sons, 2002
12 Liu Y, Wu Y, Cai J, et al. Additive-multiplicative rates model for recurrent events. Lifetime Data Anal, 2010, 16: 353–373
13 Lu K, Tsiatis A A. Multiple imputation methods for estimating regression coefficients in the competing risk model with missing cause of failure. Biometrics, 2001, 57: 1191–1197
14 Lu W, Liang Y. Analysis of competing risks data with missing cause of failure under additive hazards model. Statist Sinica, 2008, 18: 219–234
15 Patterson B J, Freedman J, Blanchette V, et al. Effect of premedication guidelines and leukoreduction on the rate of febrile nonhaemolytic platelet transfusion reactions. Transfusion Medicine, 2000, 10: 199–206
16 Pepe M S, Cai J. Some graphical displays and marginal regression analyses for recurrent failure times and time dependent covariates. J Amer Statist Assoc, 1993, 88: 811–820
17 Pollard D. Empirical Processes: Theory and Applications. Hayward: Institute of Mathematical Statistics, 1990
18 Rudin W B. Principles of Mathematical Analysis, 3rd ed. New York: McGraw-Hill Science/Engineering/Math, 1976
19 Schaubel D E, Cai J. Rate/mean regression for multiple-sequence recurrent event data with missing event category. Scand J Statist, 2006, 33: 191–207
20 Schaubel D E, Cai J. Multiple imputation methods for recurrent event data with missing event category. Canad J Statist, 2006, 34: 677–692
21 Schaubel D E, Zeng D, Cai J. A semiparametric additive rates model for recurrent event data. Lifetime Data Anal, 2006, 12: 389–406
22 Song X, Sun L, Mu X, et al. Additive hazards regression with censoring indicators missing at random. Canad J Statist, 2010, 38: 333–351
23 Sun L, Kang F. An additive-multiplicative rates model for recurrent event data with informative terminal event. Lifetime Data Anal, 2013, 19: 117–137
24 Sun L, Zhu L, Sun J. Regression analysis of multivariate recurrent event data with time-varying covariate effects. J Multivariate Anal, 2009, 100: 2214–2223
25 van der Vaart A W, Wellner J A. Weak Convergence and Empirical Processes. New York: Springer-Verlag, 1996
26 Zeng D, Lin D Y. Efficient estimation of semiparametric transformation models for counting processes. Biometrika, 2006, 93: 627–640
27 Zhao X, Liu L, Liu Y, et al. Analysis of multivariate recurrent event data with time-dependent covariates and

informative censoring. Biom J, 2012, 54: 585–599
28 Zhu L, Sun J, Tong X, et al. Regression analysis of multivariate recurrent event data with a dependent terminal event. Lifetime Data Anal, 2010, 16: 478–490

Appendix

Proof of Theorem 1. Firstly, we prove the strong consistency. Based on (11), we have $U(\theta) = \sum_{i=1}^n \sum_{k=1}^K \int_0^\tau \{Z_{ik}(s,\theta) - \bar{Z}_k(s,\gamma)\} dM_{ik}(s,\theta,\hat{\eta})$. Let $\bar{z}_k^{(1)}(t,\gamma) = \partial \bar{z}_k(t,\gamma)/\partial\gamma^{\mathrm{T}}$ and $\hat{A}_1(\theta) = -n^{-1} \partial U(\theta)/\partial\theta^{\mathrm{T}}$, where

$$\begin{aligned}\hat{A}_1(\theta) = n^{-1}\sum_{i=1}^n\sum_{k=1}^K\int_0^\tau & Y_{ik}(s)[Z_{ik}(s)-\bar{Z}_k(s,\gamma)]\\ &\times[\dot{g}(\beta^{\mathrm{T}}W_{ik}(s))W_{ik}(s)^{\mathrm{T}}ds, \dot{h}(\gamma^{\mathrm{T}}X_{ik}(s))X_{ik}(s)^{\mathrm{T}}d\mu_{0k}(s)]\\ &+ n^{-1}\sum_{i=1}^n\sum_{k=1}^K\int_0^\tau \bar{Z}_k^{(1)}(s,\gamma)dM_{ik}(s,\theta,\hat{\eta})\end{aligned} \tag{A.1}$$

with $\bar{Z}_k^{(1)}(t,\gamma) = \partial\bar{Z}_k(t,\gamma)/\partial\gamma^{\mathrm{T}}$. Since any function of bounded variation can be written as difference of two increasing functions, the processes $\bar{Z}_k(t,\gamma)$ and $\bar{Z}_k^{(1)}(t,\gamma)$ can be written as sums and products of monotone functions in t and γ. Thus they are manageable [17, p. 38]. Using the uniform strong law of large numbers [17, p. 41], it follows that $\bar{Z}_k(t,\gamma)$ and $\bar{Z}_k^{(1)}(t,\gamma)$ converge almost surely to $\bar{z}_k(t,\gamma)$ and $\bar{z}_k^{(1)}(t,\gamma)$ uniformly in $t \in [0,\tau]$ and γ. Thus, $\hat{A}_1(\theta)$ converges almost surely to a nonrandom function $A_1(\theta)$ uniformly in $\theta \in \Theta$, and it is obvious that $A_1(\theta_0) = A$, where A is given in (C5). It follows from the strong law of large numbers and the consistency of $\hat{\eta}$ that almost surely,

$$n^{-1}U(\theta_0) \to E\Big[\sum_{k=1}^K\int_0^\tau\{Z_{ik}(s)-\bar{z}_k(s,\gamma_0)\}dM_{ik}(s,\theta_0,\eta_0)\Big] = 0. \tag{A.2}$$

The uniform convergence of $\hat{A}_1(\theta)$, the continuity of $A_1(\theta)$, and the nonsingularity of A imply that there exists a small neighborhood of θ_0 inside of which the eigenvalues of $\hat{A}_1(\theta)$ are bounded away from zero for all large n. Thus, it follows from the inverse function theorem [18, p. 221] that there is a small neighborhood of θ_0, inside of which there exists a unique solution $\hat{\theta}$ to $U(\theta) = 0$ for every sufficiently large n. Since the radius of the neighborhood can be taken arbitrarily small, $\hat{\theta}$ is strongly consistent.

Now we show the asymptotic normality of $\hat{\theta}$. Write

$$n^{-1/2}U(\theta_0) = \sum_{k=1}^K\sum_{l=1}^2 U_{kl}, \tag{A.3}$$

where

$$\begin{aligned}U_{k1} &= n^{-1/2}\sum_{i=1}^n\int_0^\tau[Z_{ik}(s)-\bar{Z}_k(s,\gamma_0)]dM_{ik}(s,\theta_0,\eta_0),\\ U_{k2} &= n^{-1/2}\sum_{i=1}^n\int_0^\tau[Z_{ik}(s)-\bar{Z}_k(s,\gamma_0)][dM_{ik}(s,\theta_0,\hat{\eta})-dM_{ik}(s,\theta_0,\eta_0)].\end{aligned}$$

Using the functional central limit theorem [17, p. 53], we have

$$U_{k1} = n^{-1/2}\sum_{i=1}^n\int_0^\tau[Z_{ik}(s)-\bar{z}_k(s,\gamma_0)]dM_{ik}(s,\theta_0,\eta_0) + o_p(1). \tag{A.4}$$

In view of (7), by the Taylor expansion and some straightforward calculations, we have

$$U_{k2} = n^{-1/2}\sum_{i=1}^n\int_0^\tau[Z_{ik}(s)-\bar{Z}_k(s,\gamma_0)][\pi_{ik}(s,\hat{\eta})-\pi_{ik}(s,\eta_0)]dN_{i\cdot}^c(s),$$

$$= n^{-1/2} \sum_{i=1}^{n} \Psi_k(\gamma_0, \eta_0)\Omega(\eta_0)^{-1} \sum_{k=1}^{K} \Gamma_{ik}(\eta_0) + o_p(1), \tag{A.5}$$

where $\Psi_k(\gamma, \eta)$, $\Omega(\eta)$ and $\Gamma_{ik}(\eta)$ are defined in Theorem 1. It follows from (A.3)–(A.5) that

$$n^{-1/2}U(\theta_0) = n^{-1/2} \sum_{i=1}^{n} \sum_{k=1}^{K} \Phi_{ik}(\theta_0, \eta_0) + o_p(1), \tag{A.6}$$

where $\Phi_{ik}(\theta, \eta) = \int_0^\tau [Z_{ik}(s) - \bar{z}_k(s, \gamma)]dM_{ik}(s, \theta, \eta) + \Psi_k(\gamma, \eta)\Omega(\eta)^{-1} \sum_{l=1}^{K} \Gamma_{il}(\eta)$. Using the Taylor expansion and $\hat{A}_1(\theta_0) \to A$ almost surely, we have

$$n^{1/2}(\hat{\theta} - \theta_0) = A^{-1}n^{-1/2} \sum_{i=1}^{n} \sum_{k=1}^{K} \Phi_{ik}(\theta_0, \eta_0) + o_p(1). \tag{A.7}$$

By the multivariate central limit theorem, $n^{1/2}(\hat{\theta} - \theta_0)$ is asymptotically normal with mean zero and covariance matrix $A^{-1}\Sigma(A^{-1})^{\mathrm{T}}$, where Σ is defined in Theorem 1. □

Proof of Theorem 2. First write

$$\begin{aligned}\hat{\mu}_{0k}(t) - \mu_{0k}(t) = \{\hat{\mu}_{0k}(t) - \hat{\mu}_{0k}(t; \theta_0, \hat{\eta})\} + \{\hat{\mu}_{0k}(t; \theta_0, \hat{\eta}) - \hat{\mu}_{0k}(t; \theta_0, \eta_0)\} \\ + \{\hat{\mu}_{0k}(t; \theta_0, \eta_0) - \mu_{0k}(t)\}.\end{aligned} \tag{A.8}$$

By the Taylor expansion, we get

$$\hat{\mu}_{0k}(t) - \hat{\mu}_{0k}(t; \theta_0, \hat{\eta}) = \hat{D}_k(t, \theta^*)^{\mathrm{T}}(\hat{\theta} - \theta_0), \tag{A.9}$$

where θ^* lies between θ_0 and $\hat{\theta}$, and $\hat{D}_k(t, \theta)$ is given in Theorem 2. In a similar manner, we have

$$\hat{\mu}_{0k}(t; \theta_0, \hat{\eta}) - \hat{\mu}_{0k}(t; \theta_0, \eta_0) = \hat{B}_k(t, \gamma_0, \eta^*)(\hat{\eta} - \eta_0), \tag{A.10}$$

where η^* lies between $\hat{\eta}$ and η_0, and $\hat{B}_k(t, \gamma, \eta)$ is defined in Theorem 2. Under the conditions (C1)–(C4), $\hat{D}_k(t, \theta^*)$ and $\hat{B}_k(t, \gamma_0, \eta^*)$ are of bounded variation asymptotically uniformly in $t \in [0, \tau]$. Using the strong consistency of $\hat{\theta}$ and $\hat{\eta}$, we get that both $\hat{\mu}_{0k}(t) - \hat{\mu}_{0k}(t; \theta_0, \hat{\eta})$ and $\hat{\mu}_{0k}(t; \theta_0, \hat{\eta}) - \hat{\mu}_{0k}(t; \theta_0, \eta_0)$ converge almost surely to 0 uniformly in $t \in [0, \tau]$.

For the third term on the right-hand side of (A.8), after some algebraic manipulations, we obtain

$$\hat{\mu}_{0k}(t; \theta_0, \eta_0) - \mu_{0k}(t) = n^{-1} \sum_{i=1}^{n} \int_0^t \frac{dM_{ik}(s, \theta_0, \eta_0)}{S_k^0(s, \gamma_0)}. \tag{A.11}$$

By the uniform strong law of large numbers and [8, Lemma 1], it can be seen that $\hat{\mu}_{0k}(t; \theta_0, \eta_0) - \mu_{0k}(t)$ converges almost surely to 0 uniformly in $t \in [0, \tau]$. Thus, it follows from (A.8)–(A.11) that $\hat{\mu}_{0k}(t)$ converges almost surely to $\mu_{0k}(t)$ uniformly in $t \in [0, \tau]$.

Now we prove the weak convergence of $\hat{\mu}_{0k}(t)$. Based on (A.7) and (A.9), we get

$$n^{1/2}\{\hat{\mu}_{0k}(t) - \hat{\mu}_{0k}(t; \theta_0, \hat{\eta})\} = D_k(t, \theta_0)^{\mathrm{T}} A^{-1} n^{-1/2} \sum_{i=1}^{n} \sum_{k=1}^{K} \Phi_{ik}(\theta_0, \eta_0) + o_p(1), \tag{A.12}$$

where $D_k(t, \theta)$ is the limit of $\hat{D}_k(t, \theta)$. It follows from (A.10) and (7) that

$$n^{1/2}\{\hat{\mu}_{0k}(t; \theta_0, \hat{\eta}) - \hat{\mu}_{0k}(t; \theta_0, \eta_0)\} = B_k(t, \gamma_0, \eta_0)\Omega(\eta_0)^{-1} n^{-1/2} \sum_{i=1}^{n} \sum_{l=1}^{K} \Gamma_{il}(\eta_0) + o_p(1), \tag{A.13}$$

uniformly in $t \in [0, \tau]$, where $B_k(t, \gamma, \eta)$ is the limit of $\hat{B}_k(t, \gamma, \eta)$. In addition, by (A.11), we have that uniformly in $t \in [0, \tau]$,

$$n^{1/2}\{\hat{\mu}_{0k}(t; \theta_0, \eta_0) - \mu_{0k}(t)\} = n^{-1/2} \sum_{i=1}^{n} \int_0^t \frac{dM_{ik}(s, \theta_0, \eta_0)}{s_k^0(s, \gamma_0)} + o_p(1), \tag{A.14}$$

where $s_k^0(s,\gamma_0)$ is the limit of $S_k^0(s,\gamma_0)$. Therefore, it follows from (A.8) and (A.12)–(A.14) that

$$n^{1/2}\{\hat{\mu}_{0k}(t)-\mu_{0k}(t)\}=n^{-1/2}\sum_{i=1}^{n}\phi_{ik}(t,\theta_0,\eta_0)+o_p(1), \tag{A.15}$$

uniformly in $t\in[0,\tau]$, where $\phi_{ik}(t;\theta,\eta)$ is as defined in Theorem 2. Because $\phi_{ik}(t;\theta_0,\eta_0)$ are independent and identically distributed zero-mean random variables for each t, the multivariate central limit theorem implies that $n^{1/2}\{\hat{\mu}_{0k}(t)-\mu_{0k}(t)\}$ converges in finite-dimensional distributions to a zero-mean Gaussian process. Note that $D_k(t,\theta_0)$ and $B_k(t,\gamma_0,\eta_0)$ are deterministic functions, and the third term of $\phi_{ik}(t;\beta_0,\gamma_0)$ can be written as sums of monotone functions of t. Hence, $n^{1/2}\{\hat{\mu}_{0k}(t)-\mu_{0k}(t)\}$ is tight [25, p. 215], and converges weakly to a zero-mean Gaussian process with covariance function $\omega_k(s,t)=E\{\phi_{ik}(s,\theta_0,\eta_0)\phi_{ik}(t,\theta_0,\eta_0)\}$ for $s,t\in[0,\tau]$. This completes the proof. □

Proof of Theorem 3. Note that

$$\mathcal{L}_k(t,z)=n^{-1/2}\sum_{i=1}^{n}\int_0^t I(Z_{ik}(s)\leqslant z)dM_{ik}(s)+R_{k1}(t,z)+R_{k2}(t,z)+R_{k3}(t,z), \tag{A.16}$$

where

$$\begin{aligned}
R_{k1}(t,z)&=n^{-1/2}\sum_{i=1}^{n}\int_0^t I(Z_{ik}(s)\leqslant z)[\pi_{ik}(s,\hat{\eta})-\pi_{ik}(s,\eta_0)]dN_{i.}^c(s),\\
R_{k2}(t,z)&=-n^{-1/2}\sum_{i=1}^{n}\bigg\{\int_0^t I(Z_{ik}(s)\leqslant z)Y_{ik}(s)[g(\hat{\beta}^{\mathrm{T}}W_{ik}(s))ds+h(\hat{\gamma}^{\mathrm{T}}X_{ik}(s))d\mu_{0k}(s)\\
&\quad-g(\beta_0^{\mathrm{T}}W_{ik}(s))ds-h(\gamma_0^{\mathrm{T}}X_{ik}(s))d\mu_{0k}(s)]\bigg\},
\end{aligned}$$

and $R_{k3}(t,z)=-n^{-1/2}\sum_{i=1}^n\int_0^t I(Z_{ik}(s)\leqslant z)Y_{ik}(s)h(\hat{\gamma}^{\mathrm{T}}X_{ik}(s))d[\hat{\mu}_{0k}(s)-\mu_{0k}(s)]$. Similarly to (A.5), we obtain

$$R_{k1}(t,z)=n^{-1/2}\sum_{i=1}^{n}E_k(t,z)\Omega(\eta_0)^{-1}\sum_{l=1}^{K}\Gamma_{il}(\eta_0)+o_p(1), \tag{A.17}$$

where $E_k(t,z)$ is given in Theorem 3. By the Taylor expansion, uniform strong law of large numbers and (A.7), we get

$$R_{k2}(t,z)=-n^{-1/2}\sum_{i=1}^{n}L_k(t,z)^{\mathrm{T}}A^{-1}\sum_{l=1}^{K}\Phi_{il}(\theta_0,\eta_0)+o_p(1), \tag{A.18}$$

where $L_k(t,z)$ is presented in Theorem 3. In addition, by the uniform strong law of large numbers and (A.15), we have

$$\begin{aligned}
R_{k3}(t,z)=-n^{-1/2}\sum_{i=1}^{n}\bigg\{&H_k(t,z)^{\mathrm{T}}A^{-1}\sum_{l=1}^{K}\Phi_{il}(\theta_0,\eta_0)+F_k(t,z)\Omega(\eta_0)^{-1}\sum_{l=1}^{K}\Gamma_{il}(\eta_0)\\
&+\int_0^t\frac{G_k(s,z)}{s_k^0(s,\gamma_0)}dM_{ik}(s)\bigg\}+o_p(1),
\end{aligned} \tag{A.19}$$

where $G_k(t,z)$, $H_k(t,z)$ and $F_k(t,z)$ are also defined in Theorem 3. Then it follows from (A.16)–(A.19) that $\mathcal{L}_k(t,z)=n^{-1/2}\sum_{i=1}^n\sigma_{ik}(t,z)+o_p(1)$, where $\sigma_{ik}(t,z)$ is defined in Theorem 3. The multivariate central limit theorem implies that $\mathcal{L}_k(t,z)$ converges in finite-dimensional distributions to a zero-mean Gaussian process. Using the similar argument as the tightness of $n^{1/2}\{\hat{\mu}_{0k}(t)-\mu_{0k}(t)\}$, $\mathcal{L}_k(t,z)$ is tight. Therefore, $\mathcal{L}_k(t,z)$ converges weakly to a zero-mean Gaussian process with covariance function at (t_1,z_1) and (t_2,z_2) equal to $E\{\sigma_{ik}(t_1,z_1)\sigma_{ik}(t_2,z_2)\}$. By the arguments of Lin et al. [8], the limiting Gaussian process can be approximated by the zero-mean Gaussian process $\hat{\mathcal{L}}_k(t,z)$ given in (12). □