Published in final edited form as:

Lifetime Data Anal. 2010 April ; 16(2): 250–270. doi:10.1007/s10985-009-9146-0.

Semiparametric proportional means model for marker data contingent on recurrent event

Jianwen Cai,

Department of Biostatistics, University of North Carolina at Chapel Hill, Campus Box 7420,
Chapel Hill, NC 27599-7420, USA

Donglin Zeng, and

Department of Biostatistics, University of North Carolina at Chapel Hill, Campus Box 7420,
Chapel Hill, NC 27599-7420, USA

Wenqin Pan

Department of Biostatistics and Bioinformatics, Duke University, Durham, NC 27705, USA

Jianwen Cai: cai@bios.unc.edu; Donglin Zeng: dzeng@bios.unc.edu; Wenqin Pan: wendy.pan@duke.edu

Abstract

In many biomedical studies with recurrent events, some markers can only be measured when events happen. For example, medical cost attributed to hospitalization can only incur when patients are hospitalized. Such marker data are contingent on recurrent events. In this paper, we present a proportional means model for modelling the markers using the observed covariates contingent on the recurrent event. We also model the recurrent event via a marginal rate model. Estimating equations are constructed to derive the point estimators for the parameters in the proposed models. The estimators are shown to be asymptotically normal. Simulation studies are conducted to examine the finite-sample properties of the proposed estimators and the proposed method is applied to a data set from the Vitamin A Community Trial.

Keywords

Recurrent event; Recurrent marker; Joint models; Rate function; Estimating equation

1 Introduction

In many medical studies, patients may experience the same type of event repeatedly over time. Typical examples of recurrent events include cancer recurrence, multiple infection episodes, and repeated use of drugs, etc. Although interest usually focuses on the process of recurrent events, investigators are also interested in outcomes associated with each event, for example, symptoms at the time of each infection and medical cost incurred at each hospitalization due to heart attack, etc. We refer to such outcomes as the markers contingent on recurrent events. In many studies, both the recurrent event itself and the associated marker are of interest.

A motivating example is from the Vitamin A Community Trial (VACT) conducted in Brazil (Barreto et al. 1994). The study was a randomized community trial to examine the effect of

the supplementation of vitamin A on diarrhea morbidity and diarrhea severity in children living in areas where its intake was inadequate. Some previous studies showed that vitamin A supplement could reduce children's mortality by 23–34% in populations where vitamin A deficiency was endemic. However, the effect on the morbidity was little known before the study. The VACT study consisted of 1240 children, aged 6–48 months at baseline, who were assigned to receive either vitamin A or placebo every 4 months for 1 year in a small city in Northeast Brazil. During the study, a number of children had one or more diarrheal episodes, where an episode of diarrhea was defined as a sequence of days with diarrhea and the episode was considered ended if there were 3 or more days without diarrhea. At each diarrheal episode, the study also collected information on the severity of a diarrheal episode based on the duration of an episode and a severity scale which was based on the number of liquid or semi-liquid motions reported in a 24-h period. The severity scale was graded as 1 (mild), 2 (moderate) and 3 (severe). The actual assessment of the severity of one diarrhea event is the severity score defined as the severity scale multiplied by the length of the episode. Clearly, this score is a marker process contingent on recurrent diarrhea. For this study, investigators were particularly interested in answering the following two questions: (a) Did taking vitamin A reduce the number of diarrheal episodes? (b) Given that a diarrheal episode occurs, did taking vitamin A lessen the diarrheal severity?

When markers are repeatedly measured and each measurement is associated with the occurrence of the recurrent event, joint modelling is a natural way to analyze such data. Research on joint modelling of event data and repeated outcomes has been fruitful in the last decade, for example, modelling repeated outcomes and a survival endpoint is discussed in Wu and Bailey (1989), Tsiatis et al. (1995), and Xu and Zeger (2001) among others; while, modelling repeated outcomes and recurrent events is given in Ratcliffe et al. (2004) and in Herring and Yang (2007) with an additional terminal event. In these models, subject-specific random effects are introduced to account for latent dependence between the markers and the recurrent events and the interpretation of the parameters is dependent on the subject. In many studies, however, the public health interest focuses on the population-level effects of risk factors, especially for the studies such as the VACT trial. In such studies, models with population average interpretation are preferred and joint marginal models may be more suitable. Joint marginal models have been considered by Lin and Ying (2001) and Hu et al. (2003), who propose a proportional rates model for the recurrent events and assumed a marginal means model for the repeated outcomes.

All the aforementioned joint models, either with random effects or marginal models, assume that the repeated markers and the event time are independent conditional on covariates and/or subject-specific random effects. Thus, the repeated markers should have the same distribution at the time of event occurrence as at the time of no event. Obviously, this assumption is not valid when the repeated markers are actually zeros at the time when no event occurs. Therefore, none of these approaches is applicable to the VACT trial or medical cost studies. Furthermore, the parameters introduced in the model for repeated markers in the joint modelling approach can not directly answer the practical question like (b).

The repeated marker process contingent on recurrent event can also be viewed as a marked point process, which has been widely used in modelling spatial data (Diggle 2003). The study of such process usually assumes some Poisson structure (e.g., Martinussen and Scheike 2000, 2001) and it may lead to biased result if the Poisson structure does not hold. In this paper, we model the rate of recurrent events and the mean of the marker contingent on recurrent events. Particularly, a proportional rates model (Lin et al. 2000) is used to model recurrent event while a proportional means model is used for recurrent markers contingent on the occurrence of recurrent events. Our models do not need the Poisson

assumption and the model parameters have direct interpretation to answer questions like (a) and (b).

In Sect. 2, we provide details of the proposed models. Moreover, we propose an inference procedure for estimating the effects of covariates. In Sects. 3 and 4, we give the asymptotic properties of the proposed estimators along with the estimated cumulative marker measure. We then propose some test statistics for goodness-of-fit in Sect. 5. Simulation studies are reported in Sect. 6. The proposed method is then applied to data from the VACT study in Sect. 7. The paper concludes with some remarks in Sect. 8.

2 Models and inference

Suppose $X(t)$ contains all the external covariates. Let $N(t)$ denote the counting process associated with recurrent events. A marginal model for the rate function of the recurrent event is assumed to be

$$E[dN(t)|X] = \lambda(t)g_r(X(t)^T\beta)dt, \quad (1)$$

where g_r is a known link function and is usually assumed to be exponential. Let $m(t)$ denote the marker measured at time t and assumed to be positive. Since $m(t)$ only exists when $N(t) - N(t-) = 1$, we propose a conditional mean model for $m(t)$ conditional on $N(t) - N(t-) = 1$:

$$E[m(t)|N(t) - N(t-) = 1, X] = \mu(t)g_m(X(t)^T\gamma) \quad (2)$$

where g_m is another known link function. The link function g_m is exponential if the effect of the covariates on the mean marker accumulation is multiplicative; however, other links such as the inverted logistic link or linear link can also be useful in some specific contexts.

Models (1) and (2) specify a very general class of models. It includes a general class of frailty models as special cases. To see this, suppose $\xi(t)$ denote the subject specific frailty process independent of X . Assume that the recurrent events follow the frailty model

$$E[dN(t)|X, \xi] = \lambda(t)g_r(X(t)^T\beta)\psi_1(\xi(t), t)dt$$

where $\psi_1(\cdot, \cdot)$ is a bivariate function. Since the marker is associated with the recurrent event, it may also be natural to assume ξ affects the mean of $m(t)$. Then we can assume

$$E[m(t)|N(t) - N(t-) = 1, X, \xi] = \mu(t)g_m(X(t)^T\gamma)\psi_2(\xi(t), t),$$

where $\psi_2(\cdot, \cdot)$ is another bivariate function. It is easy to show

$$E[dN(t)|X] = \lambda(t)E[\psi_1(\xi(t), t)]g_r(X(t)^T\beta)dt.$$

Moreover,

$$\begin{aligned}
& E[m(t)|N(t) - N(t-)=1, X] \\
&= \mu(t)g_m(X(t)^T \gamma) E[\psi_2(\xi(t), t)|N(t) - N(t-)=1, X] \\
&= \mu(t)g_m(X(t)^T \gamma) \frac{E[\psi_2(\xi(t), t)dN(t)|X]}{E(dN(t)|X)}
\end{aligned}$$

and

$$\begin{aligned}
E[\psi_2(\xi(t), t)dN(t)|X] &= E[\psi_2(\xi(t), t)\lambda(t)g_r(X(t)^T \beta)\psi_1(\xi(t), t)dt|X] \\
&= E[\psi_1(\xi(t), t)\psi_2(\xi(t), t)\lambda(t)g_r(X(t)^T \beta)dt].
\end{aligned}$$

Hence, we obtain

$$E[m(t)|N(t) - N(t-)=1, X] = \mu(t) \frac{E[\psi_1(\xi(t), t)\psi_2(\xi(t), t)]}{E[\psi_1(\xi(t), t)]} g_m(X(t)^T \gamma).$$

Therefore, the assumed frailty models imply models (1) and (2) but with different baseline functions, i.e., the assumed frailty models are special cases of models (1) and (2). In this paper, we work with the general case as specified in models (1) and (2).

In the presence of censoring, the observations from n i.i.d subjects are

$$(N_i(t), m_i(t)\Delta N_i(t), X_i(t), t \leq C_i), i=1, \dots, n$$

where $\Delta N_i(t) = N_i(t) - N_i(t-)$ and C_i is the censoring time and is assumed to be non-informative of $N_i(t)$ and $m_i(t)$ given covariates. To make inference on β and γ , from (1) and following Lin et al. (2000), we first estimate β using the following estimating equation

$$\sum_{i=1}^n \int R_i(t) \left[X_i(t) \frac{g'_r(X_i(t)^T \beta)}{g_r(X_i(t)^T \beta)} - \frac{\sum_{j=1}^n R_j(t) g'_r(X_j(t)^T \beta) X_j(t)}{\sum_{j=1}^n R_j(t) g_r(X_j(t)^T \beta)} \right] dN_i(t) = 0, \quad (3)$$

where $R_i(t) = I(C_i > t)$ is the at-risk process for subject i and $g'_r(x)$ denotes the derivative of $g_r(x)$. We denote $\hat{\beta}$ as the solution to (3).

To estimate γ , from Eqs. 1 and 2, if we define the cumulative marker at time t as

$$M(t) = \int_0^t m(s) dN(s), \quad (4)$$

where $m(t)$ is the marker measured at time t whenever an event occurs at time t , then we obtain

$$\begin{aligned}
E[dM(t)|X] &= E[m(t)dN(t)|X] = \mu(t)g_m(X(t)^T \gamma) E[dN(t)|X] \\
&= \mu(t)\lambda(t)g_m(X(t)^T \gamma)g_r(X(t)^T \beta)dt.
\end{aligned}$$

The above expression is another type of rate model similar to (1) if we redefine $\mu(t)\lambda(t) = v(t)$. Using the same idea of constructing estimating equation for the recurrent event, we treat $V(t) = \int_0^t v(s)ds$ as a step function with jumps at the observed events. We then propose the following estimating equations for estimating $V(t)$ and γ :

$$\begin{aligned} \sum_{i=1}^n \int R_i(t) I(t \leq s) \left[dM_i(t) - g_m(X_i(t)^T \gamma) g_r(X_i(t)^T \widehat{\beta}) dV(t) \right] &= 0, \quad s \in (0, \infty), \\ \sum_{i=1}^n \int R_i(t) I(t \leq s) X_i(t) \frac{g'_m(X_i(t)^T \gamma)}{g_m(X_i(t)^T \gamma)} \\ &\times \left[dM_i(t) - g_m(X_i(t)^T \gamma) g_r(X_i(t)^T \widehat{\beta}) dV(t) \right] = 0. \end{aligned}$$

From the first equation, we obtain

$$\widehat{V}(t) = \int_0^t \frac{\sum_{i=1}^n R_i(s) dM_i(s)}{\sum_{i=1}^n R_i(s) g_m(X_i(s)^T \gamma) g_r(X_i(s)^T \widehat{\beta})}.$$

After plugging this expression into the second equation for γ , we have that the estimate for γ , denoted by $\hat{\gamma}$, solves

$$\begin{aligned} \sum_{i=1}^n \int R_i(t) \left[X_i(t) \frac{g'_m(X_i(t)^T \gamma)}{g_m(X_i(t)^T \gamma)} - \frac{\sum_{j=1}^n R_j(t) g_r(X_j(t)^T \widehat{\beta}) g'_m(X_j(t)^T \gamma) X_j(t)}{\sum_{j=1}^n R_j(t) g_r(X_j(t)^T \widehat{\beta}) g_m(X_j(t)^T \gamma)} \right] \\ \times dM_i(t) = 0. \end{aligned} \quad (5)$$

3 Asymptotic properties

We need the following assumptions.

- (C1) Let τ be the study duration. We assume $\Pr(C \geq \tau | X) > 0$.
- (C2) Covariate $X(t)$ is a bounded function with finite total variation in $[0, \tau]$.
- (C3) Both $g_r(\cdot)$ and $g_m(\cdot)$ are twice-continuously differentiable and strictly positive.
- (C4) Let

$$v_k(t) = R(t) g_r^{(k)}(X(t)^T \beta_0), \quad \omega_{kl}(t) = R(t) g_r^{(k)}(X(t)^T \beta_0) g_m^{(l)}(X(t)^T \gamma_0), \quad k, l = 0, 1, 2,$$

with the superscripts denoting the derivatives. Assume that both matrix

$$E \left\{ \int R(t) \left[X(t)^{\otimes 2} \left(\frac{g'_r(x)}{g_r(x)} \right)' \right] \Big|_{x=X(t)^T \beta_0} - \frac{E[v_2(t) X(t)^{\otimes 2}]}{E[v_0(t)]} + \left(\frac{E[v_1(t) X(t)]}{E[v_0(t)]} \right)^{\otimes 2} \right\} dN(t)$$

and matrix

$$E \left\{ \int R(t) \left[X(t)^{\otimes 2} \left(\frac{g'_m(x)}{g_m(x)} \right)' \Big|_{x=X(t)^T \gamma_0} - \frac{E\{\omega_{02}(t)X(t)^{\otimes 2}\}}{E\{\omega_{00}(t)\}} + \left(\frac{E\{\omega_{01}(t)X(t)\}}{E\{\omega_{00}(t)\}} \right)^{\otimes 2} \right] dM(t) \right\}$$

are non-singular.

Condition (C1) assumes that there is a positive chance of observing the entire event process till time τ . Condition (C3) includes a number of choices of link functions. When $g_r(x) = g_m(x) = \exp(x)$, condition (C4) is similar to the standard condition in the proportional hazards model. Under these conditions, the following theorem holds.

Theorem 1

Under conditions (C1)–(C4), $\sqrt{n}(\widehat{\beta} - \beta_0, \widehat{\gamma} - \gamma_0)$ converges in distribution to a normal distribution with mean zero.

Let $\mathcal{P}$ denote the underlying probability measure and $\mathcal{P}[Q] = E[Q]$ for any random variable Q . Correspondingly, let $\mathcal{P}_n$ denote the empirical measure based on n observations. The proof of Theorem 1 is mainly based on the linearization of Eq. 3 and use the weak convergence of $\sqrt{n}(\mathcal{P}_n - \mathcal{P})$. The details are given in the appendix. Particularly, the proof of Theorem 1 yields that the asymptotic covariance for $\sqrt{n}(\widehat{\beta} - \beta_0, \widehat{\gamma} - \gamma_0)$ is equal to

$$\mathcal{P} \left[\left(\begin{array}{c} \Sigma_{11}^{-1} S_\beta \\ \Sigma_{22}^{-1} [S_\gamma - \Sigma_{21} \Sigma_{11}^{-1} S_\beta] \end{array} \right)^{\otimes 2} \right],$$

where

$$\begin{aligned} \Sigma_{11} &= \mathcal{P} \left\{ \int R(t) \left[X(t)^{\otimes 2} \left(\frac{g'_r(x)}{g_r(x)} \right)' \Big|_{x=X(t)^T \beta_0} - \frac{\mathcal{P}\{v_2(t)X(t)^{\otimes 2}\}}{\mathcal{P}\{v_0(t)\}} + \left(\frac{\mathcal{P}\{v_1(t)X(t)\}}{\mathcal{P}\{v_0(t)\}} \right)^{\otimes 2} \right] dN(t) \right\}, \\ \Sigma_{22} &= \mathcal{P} \left\{ \int R(t) \left[X(t)^{\otimes 2} \left(\frac{g'_m(x)}{g_m(x)} \right)' \Big|_{x=X(t)^T \gamma_0} - \frac{\mathcal{P}\{\omega_{02}(t)X(t)^{\otimes 2}\}}{\mathcal{P}\{\omega_{00}(t)\}} + \left(\frac{\mathcal{P}\{\omega_{01}(t)X(t)\}}{\mathcal{P}\{\omega_{00}(t)\}} \right)^{\otimes 2} \right] dM(t) \right\}, \\ \Sigma_{21} &= -\mathcal{P} \left\{ \int R(t) \left[\frac{\mathcal{P}\{\omega_{11}(t)X(t)^{\otimes 2}\}}{\omega_{00}(t)} - \frac{\mathcal{P}\{\omega_{01}(t)X(t)\}\mathcal{P}\{\omega_{10}(t)X(t)^T\}}{(\mathcal{P}\{\omega_{00}(t)\})^2} \right] dM(t) \right\}, \end{aligned}$$

and

$$\begin{aligned}
S_\beta = & -\int R(t) \left[X(t) \frac{g_r'(X(t)^T \beta_0)}{g_r(X(t)^T \beta_0)} - \frac{\mathcal{P}\{v_1(t)X(t)\}}{\mathcal{P}\{v_0(t)\}} \right] dN(t) \\
& + \int R(t) g_r'(X(t)^T \beta_0) X(t) \frac{\mathcal{P}\{R(t)dN(t)\}}{\mathcal{P}\{R(t)g_r(X(t)^T \beta_0)\}} \\
& - \int R(t) g_r(X(t)^T \beta_0) \frac{\mathcal{P}\{R(t)dN(t)\} \mathcal{P}\{R(t)g_r(X(t)^T \beta_0)X(t)\}}{\{\mathcal{P}\{R(t)g_r(X(t)^T \beta_0)\}\}^2}, \\
S_\gamma = & -\int R(t) \left[X(t) \frac{g_m'(X(t)^T \gamma_0)}{g_m(X(t)^T \gamma_0)} - \frac{\mathcal{P}\{\omega_{01}(t)X(t)\}}{\mathcal{P}\{\omega_{00}(t)\}} \right] dM(t) \\
& + \int \omega_{01}(t) X(t) \frac{\mathcal{P}\{R(t)dM(t)\}}{\mathcal{P}\{\omega_{00}(t)\}} - \int \omega_{00}(t) \frac{\mathcal{P}\{R(t)dM(t)\} \mathcal{P}\{\omega_{01}(t)X(t)\}}{\{\mathcal{P}\{\omega_{00}(t)\}\}^2}.
\end{aligned}$$

Therefore, a consistent estimator for the asymptotic covariance can be given by

$$\mathcal{P}_n \left[\left(\begin{array}{c} \widehat{\Sigma}_{11}^{-1} \widehat{S}_\beta \\ \widehat{\Sigma}_{22}^{-1} \{ \widehat{S}_\gamma - \widehat{\Sigma}_{21} \widehat{\Sigma}_{11}^{-1} \widehat{S}_\beta \} \end{array} \right)^{\otimes 2} \right],$$

where $\widehat{\Sigma}_{kl}$, $k, l = 1, 2$, $\widehat{S}_\beta$ and $\widehat{S}_\gamma$ are defined as in the above expressions but with $\mathcal{P}$ replaced by $\mathcal{P}_n$ and the true parameters replaced by their estimators.

4 Estimation of cumulative recurrent marker

In many applications, such as in studies to assess recurrent severity score or recurrent medical cost, one may also be interested in estimating the total marker by a fixed time s for $s \leq \tau$, i.e., $E[M(s)]$.

From the proposed estimation method in Sect. 2, we can estimate this quantity by

$$\begin{aligned}
\widehat{M}(s) & \equiv \frac{1}{n} \sum_{i=1}^n \int_0^s g_r(X_i(t)^T \widehat{\beta}) g_m(X_i(t)^T \widehat{\gamma}) d\widehat{V}(t) \\
& = \frac{1}{n} \int_0^s \frac{\sum_{i=1}^n g_r(X_i(t)^T \widehat{\beta}) g_m(X_i(t)^T \widehat{\gamma}) \sum_{i=1}^n R_i(t) dM_i(t)}{\sum_{i=1}^n R_i(t) g_m(X_i(t)^T \widehat{\gamma}) g_r(X_i(t)^T \widehat{\beta})}.
\end{aligned}$$

The following theorem describes the large sample property of $\widehat{M}(s)$.

Theorem 2

Under conditions (C1)–(C4), $\sqrt{n}(\widehat{M}(s) - E[M(s)])$ converges in distribution to a mean-zero Gaussian process, denoted as $\mathcal{G}$, in $l^\infty[0, \tau]$.

The proof is given in the appendix. Moreover, from the proof, we obtain

$$\sqrt{n}(\widehat{M}(s) - E[M(s)]) = n^{-1/2} \sum_{i=1}^n \mathcal{Q}(R_i, M_i, X_i, s; \mathcal{P}, \beta_0, \gamma_0) + o_p(1),$$

where $\mathcal{Q}(R_i, M_i, X_i, s; \mathcal{P}, \beta_0, \gamma_0)$ is the influence function of $\widehat{M}(s)$ as given in the proof. Thus, we can construct the confidence band for $E[M(s)]$ for $s \in [0, \tau]$ using the Monte-Carlo method. Specifically, we generate $Q = (Q_1, \dots, Q_n)$ independently from the standard normal distribution. We then calculate

$$\widehat{M}_Q(s) = \widehat{M}(s) + n^{-1} \sum_{i=1}^n Q_i \mathcal{J}(R_i, M_i, X_i, s; \mathcal{P}_n, \widehat{\beta}, \widehat{\gamma}).$$

We repeat this process B times for a large number B . Then the collection of $\widehat{M}_Q(s)$ from B repetitions simulates the limiting process of $\widehat{M}(s)$. Particularly, we have the following result.

Theorem 3

Under conditions (C1)–(C4), conditional on the observed data, $\sqrt{n}(\widehat{M}_Q(s) - \widehat{M}(s))$ converges in distribution to $\mathcal{G}$.

As the result, we can use the simulated process $\widehat{M}_Q(s)$ to construct the confidence band for $E[M(s)]$. For example, a pointwise $(1-\alpha)$ -level confidence band for $E[M(s)]$ can be obtained by

$$[\widehat{M}(s) + z_{\alpha/2}(s), \widehat{M}(s) + z_{1-\alpha/2}(s)], \quad s \in [0, \tau],$$

where $z_{\alpha/2}(s)$ and $z_{1-\alpha/2}(s)$ are the $\alpha/2$ and $(1-\alpha/2)$ th quantiles of $\{\widehat{M}_Q(s) - \widehat{M}(s)\}$ respectively.

5 Test for goodness-of-fit

One important issue in models (1) and (2) is the choice of link functions g_r and g_m . Therefore, in this section, we propose some goodness-of-fit test statistics for the model. The construction of our test is based on the fact that when the correct link functions are used, both processes,

$$dN(t) - R(t)g_r(X(t)^T \beta_0) d\Lambda_0(t)$$

and

$$dM(t) - R(t)g_r(X(t)^T \beta_0)g_m(X(t)^T \gamma_0) dV_0(t),$$

are mean-zero processes given covariates. Hence, we consider the following two statistics

$$\mathcal{T}_1(x) = n^{-1} \sum_{i=1}^n \int R_i(t) I(X_i(t)^T \widehat{\beta} \leq x) \left[dN_i(t) - g_r(X_i(t)^T \widehat{\beta}) d\widehat{\Lambda}(t) \right]$$

and

$$\mathcal{T}_2(y) = n^{-1} \sum_{i=1}^n \int R_i(t) I(X_i(t)^T \widehat{\gamma} \leq y) \left[dM_i(t) - g_r(X_i(t)^T \widehat{\beta}) g_m(X_i(t)^T \widehat{\gamma}) d\widehat{V}(t) \right],$$

where $\widehat{\beta}$ and $\widehat{\gamma}$ are the estimates given in Sect. 2, and

$$\widehat{\Lambda}(t) = \int_0^t \frac{\sum_{i=1}^n R_i(s) dN_i(s)}{\sum_{i=1}^n R_i(s) g_r(X_i(s)^T \widehat{\beta})},$$

and

$$\widehat{V}(t) = \int_0^t \frac{\sum_{i=1}^n R_i(s) dM_i(s)}{\sum_{i=1}^n R_i(s) g_r(X_i(s)^T \widehat{\beta}) g_m(X_i(s)^T \widehat{\gamma})}.$$

Intuitively, these two processes accumulate the marginal residuals based on the linear predictors so the departure from zeros may imply that the link functions between the observed recurrent events or the markers and the linear predictors are chosen incorrectly.

The following theorem describes the asymptotic behavior of $\mathcal{T}_1(x)$ and $\mathcal{T}_2(y)$.

Theorem 4

Under models (1) and (2) and conditions (C1)–(C4), $\sqrt{n}(\mathcal{T}_1, \mathcal{T}_2)^T$ weakly converges to a bivariate Gaussian process in $l^\infty(R^2)$ with mean zero.

The proof of the theorem is provided in the appendix. From the Proof of Theorem 4, we obtain

$$\sqrt{n}(\mathcal{T}_1(x), \mathcal{T}_2(y))^T = n^{-1/2} \sum_{i=1}^n (S_1(R_i, M_i, X_i, x), S_2(R_i, M_i, X_i, y))^T + o_p(1)$$

and both $S_1(R, M, X, x)$ and $S_2(R, M, X, y)$ can be estimated consistently by $\widehat{S}_1(R, M, X, x)$ and $\widehat{S}_2(R, M, X, y)$. Therefore, we can use the same idea in the previous section to simulate the limiting Gaussian process: We generate $Q = (Q_1, \dots, Q_n)$ independently from the standard normal distribution, then calculate

$n^{-1} \sum_{i=1}^n \widehat{S}_1(R_i, M_i, X_i, x) Q_i$ and $n^{-1} \sum_{i=1}^n \widehat{S}_2(R_i, M_i, X_i, y) Q_i$. The simulated process after repeating this procedure many times approximates the limiting Gaussian process. The proof is similar to those used for Theorem 3.

With Theorem 4, we can construct any summary statistics of the proposed processes $(\mathcal{T}_1(x), \mathcal{T}_2(y))^T$, for example, the maximal norm or L_2 -norm of these two processes. Then by simulating the limiting process under the null hypothesis that correct link functions are used, we are able to construct a valid test for testing the null hypothesis.

6 Simulation studies

Two simulation studies are conducted to examine the finite-sample performance of the proposed approach. In both simulation studies, we consider two independent covariates X_1 and X_2 where X_1 is from the uniform distribution in $[0, 1]$ and X_2 is a Bernoulli variable with equal probabilities at 0 and 1. The recurrent events are generated from the following intensity function

$$\lambda_k(t|X_1, X_2, \xi) = 0.5\xi \exp\{\beta_1 X_1 + \beta_2 X_2\},$$

where ξ denotes the subject-specific frailty and follows Gamma(2,2), $\beta_1 = -1$, and $\beta_2 = 1$. At each occurred event, the marker process $m(t)$ satisfies

$$m(t) = g_m(\gamma_1 X_1 + \gamma_2 X_2) e^{\sin(t) + \varepsilon \xi^\rho},$$

where ε follows the standard normal distribution, $\gamma_1 = 0.5$, $\gamma_2 = 0.8$, and ρ describes the association between the marker and the events due to the frailty. As discussed before, the above two models imply models (1) and (2). We consider $\rho = 0$ and $\rho = 1/2$. The link function g_m is taken to be the exponential link in the first simulation study, i.e., $g_m(x) = \exp\{x\}$, and in the second simulations study, $g_m(x) = (1 + x)^2$. The censoring time is generated from the uniform distribution in $[0, 3]$ so $\tau = 3$, which yields an average of 3.5 events per subject. The true value for $E[M(3)]$ is 55.421 in the first simulation study and 78.983 in the second one.

Table 1 summarizes the simulation results from 1000 repetitions with different sample sizes. From Table 1, we conclude that the proposed method performs well in finite sample: the bias of the estimators for β 's is almost negligible while the relative bias for γ 's is within 5% for sample size 200 and decreases significantly for sample size 400; the estimated standard errors agree well with the empirical standard errors with the maximal difference less than 5% with sample size 400; the confidence interval for estimating these parameters has proper coverage. Table 1 also shows that the estimates for the cumulative marker have very small bias and the inference seems to be accurate. In the simulation studies, we also evaluate the performance of the proposed goodness-of-fit test statistics. The Type I error rates for $\max_x |\mathcal{T}_1(x)|$ are all close to 0.05 for each setting in Table 1; the type I error rates for $\max_x |\mathcal{T}_2(x)|$ is conservatively around 0.03 for sample size 200 but range between 0.045 and 0.052 when the sample size increases to 400. This suggests that the proposed tests are valid with sample size 400. We also conduct simulations where the baseline function is $0.5 \exp(0.2t)$ (results not shown). The performance is also good and the conclusions are similar.

Additionally, we conduct simulation studies to examine the robustness of our estimators when the link function in the model for the recurrent marker is misspecified. Particularly, we generate the recurrent events and the marker data as before; however, the true link function for generating the marker at each event is $g_m(x) = (1 + x)^2$ while we misspecify it as the exponential link in estimation. Table 2 summarizes the results from 1000 repetitions. From the table, we find that the misspecification of the link function causes a large bias in estimating the effects of the covariates on the marker, especially for the dichotomous covariate X_2 . However, even if we use the wrong link function, the estimates for the cumulative marker have negligible bias and their inference is also accurate. For this simulation study, we also calculate the power based on $\mathcal{T}_2$ and it turns out to be low, suggesting that the misspecified links, although affect the regression parameter estimates, have little influence on the overall fit of the data. This may also explain why the estimates for the cumulative marker are robust to the misspecification of the link function.

7 Application

We apply our method to analyze the VACT data. Besides the covariate for treatment assignment, other covariates include children's baseline age and gender. The average number of days of follow-up is 331 days, with 83.7% of the children having daily continuous information for one year. A total of 1063 children had at least one diarrheal

episode during their follow-up period in the study and the average number of episodes is 5.9. The first half in Table 3 summarizes this information in both treatment group and placebo group. To assess the severity, we calculate the severity score at each episode as the severity scale (1,2,3) multiplied by the number of duration days in this episode. The summary information for this severity score is reported in the second half of Table 3.

We use models (1) and (2) for the recurrent events and the severity scores respectively. Particularly, we use the exponential link in model (1) while consider either exponential or square links in model (2). The results given in Table 4 show that using vitamin A supplement may reduce the risk of recurrent diarrhea and such reduction is marginally significant; using vitamin A supplement did not reduce the diarrheal severity. The older children tended to have lower risk of both diarrheal recurrence and less severe diarrhea. There appears to be no difference between boys and girls. Finally, we also calculate the estimates for the total severity measurement in treatment group and placebo group. Under the exponential link, the two estimates are respectively 53.96 and 60.12. The test based on the asymptotic normal distribution yields a p-value of 0.27. Thus, it indicates that, overall, the vitamin A supplement does not significantly reduce the total severity. These results are very close to the results when the square link function is used. Figure 1 plots the estimated cumulative severity score as a function of time in the two groups along with the pointwise confidence bands. It shows that overall, the treatment group tends to have less severe episodes as compared to the placebo group; however, such difference is not statistically significant.

To compare the exponential link versus the square link, we apply the goodness-of-fit test proposed in Sect. 5. The resulting processes are given in Fig. 2. Especially, subplot (a) presents the estimated process $\mathcal{T}_2(x)$ versus the percentiles of the linear predictors $X^T \gamma$ along with 10 simulated curves from the limiting distribution when the exponential link is used for fitting marker data; while subplot (b) presents the same processes under the square link. Note that the value for the estimate of γ is different with different link functions which results in different scale in x-axes. The figure shows that the patterns under these two different links are very similar. We further calculate the p-values based on the maximal-norm and find that the p-value associated with the exponential link is 0.66 and the p-value associated with the square link is 0.75. This indicates that there is no strong evidence to reject either link, i.e., essentially they both fit data well.

8 Remarks

We have proposed a semiparametric proportional means model for marker data contingent on recurrent events. Our estimates provide insight on the effects of risk factors on both the recurrent event process and the marker associated with the recurrent event. We have applied this approach to the Vitamin A Community Trial. We note that many other medical applications can fall under this framework.

In some applications, when one is particularly interested in prediction, frailty model can be useful. Under such situation, random frailty can be introduced in modelling recurrent events and markers to account for within-subject variation. Particularly, one can model the intensity of recurrent events and the conditional distribution of recurrent markers at each event. Inference for this semiparametric model is feasible but requires more delicate justification.

Further generalization of our setting will include informative censoring and informative marker data. Additionally, using more general transformation models as in Lin et al. (2001) for recurrent events and recurrent markers is also possible.

Acknowledgments

This research is partly supported by National Institutes of Health NHLBI Grant R01 HL-57444.

Appendix

Proof of Theorem 1

For convenience, we introduce $\mathcal{P}_n$ and $\mathcal{P}$ as the empirical measure and the true measure respectively. The empirical process $\mathcal{G}_n$ is defined as $\sqrt{n}(\mathcal{P}_n - \mathcal{P})$. Then Eq. 3 can be expressed as

$$\mathcal{P}_n \left\{ \int R(t) \left[X(t) \frac{g_r'(X(t)^T \beta)}{g_r(X(t)^T \beta)} - \frac{\mathcal{P}_n\{R(t)g_r'(X(t)^T \beta)X(t)\}}{\mathcal{P}_n\{R(t)g_r(X(t)^T \beta)\}} \right] dN(t) \right\} = 0.$$

In any neighborhood of β_0 , by the Glivenko–Cantelli theorem (van der Vaart and Wellner 1996, Chap. 2.3), with probability one, the left-hand side uniformly converges to

$$\mathcal{P} \left\{ \int R(t) \left[X(t) \frac{g_r'(X(t)^T \beta)}{g_r(X(t)^T \beta)} - \frac{\mathcal{P}\{R(t)g_r'(X(t)^T \beta)X(t)\}}{\mathcal{P}\{R(t)g_r(X(t)^T \beta)\}} \right] dN(t) \right\}.$$

Since

$$\begin{aligned} & \mathcal{P} \left\{ \int R(t) \left[X(t) \frac{g_r'(X(t)^T \beta)}{g_r(X(t)^T \beta)} - \frac{\mathcal{P}\{R(t)g_r'(X(t)^T \beta)X(t)\}}{\mathcal{P}\{R(t)g_r(X(t)^T \beta)\}} \right] dN(t) \right\} \\ &= \Sigma_{11}(\beta - \beta_0) + o(|\beta - \beta_0|) \end{aligned}$$

and by the assumptions, Σ_{11} is non-singular, we have

$$\begin{aligned} & \sup_{|\beta - \beta_0| = \varepsilon} \left| \mathcal{P}_n \left\{ \int R(t) \left[X(t) \frac{g_r'(X(t)^T \beta)}{g_r(X(t)^T \beta)} - \frac{\mathcal{P}_n\{R(t)g_r'(X(t)^T \beta)X(t)\}}{\mathcal{P}_n\{R(t)g_r(X(t)^T \beta)\}} \right] dN(t) \right\} \right. \\ & \quad \left. - \Sigma_{11}(\beta - \beta_0) \right| \\ & \leq \inf_{|\beta - \beta_0| = \varepsilon} |\Sigma_{11}(\beta - \beta_0)|. \end{aligned}$$

By the degree theory (Deming 1985, Chap. 1), the inequality implies

$$\mathcal{P}_n \left\{ \int R(t) \left[X(t) \frac{g_r'(X(t)^T \beta)}{g_r(X(t)^T \beta)} - \frac{\mathcal{P}_n\{R(t)g_r'(X(t)^T \beta)X(t)\}}{\mathcal{P}_n\{R(t)g_r(X(t)^T \beta)\}} \right] dN(t) \right\}$$

has the same number of zero-solution as $\Sigma_{11}(\beta - \beta_0)$ within $|\beta - \beta_0| \leq \varepsilon$ which has only one solution. That is, there exists $\hat{\beta}$ solving Eq. 3 and $|\hat{\beta} - \beta_0| \leq \varepsilon$ when n is larger enough. We have proved the consistency of $\hat{\beta}$.

From Eq. 3, using the Donsker property of $g_r(X(t)^T \beta)$, we obtain

$$\begin{aligned}
0 = & \sqrt{n} \mathcal{P} \left\{ \int R(t) \left[X(t) \frac{g_r'(X(t)^T \beta)}{g_r(X(t)^T \beta)} - \frac{\mathcal{P}\{R(t)g_r'(X(t)^T \beta)X(t)\}}{\mathcal{P}\{R(t)g_r(X(t)^T \beta)\}} \right] dN(t) \right\} \\
& + \mathcal{G}_n \left\{ \int R(t) \left[X(t) \frac{g_r'(X(t)^T \beta_0)}{g_r(X(t)^T \beta_0)} - \frac{\mathcal{P}\{R(t)g_r'(X(t)^T \beta_0)X(t)\}}{\mathcal{P}\{R(t)g_r(X(t)^T \beta_0)\}} \right] dN(t) \right\} \\
& - \mathcal{G}_n \left\{ \int R(t)g_r'(X(t)^T \beta_0)X(t) \frac{\mathcal{P}\{R(t)dN(t)\}}{\mathcal{P}\{R(t)g_r(X(t)^T \beta_0)\}} \right\} \\
& + \mathcal{G}_n \left\{ \int R(t)g_r(X(t)^T \beta_0) \frac{\mathcal{P}\{R(t)dN(t)\}\mathcal{P}\{R(t)g_r'(X(t)^T \beta_0)X(t)\}}{\mathcal{P}\{R(t)g_r(X(t)^T \beta_0)\}^2} \right\} + o_p(1).
\end{aligned}$$

Therefore,

$$\sqrt{n}(\widehat{\beta} - \beta_0) = [\Sigma_{11} + o(\widehat{\beta} - \beta_0)]^{-1} \mathcal{G}_n S_\beta + o_p(1),$$

where S_β is defined before. As the result, $\widehat{\beta}$ is an asymptotically linear estimator for β_0 with influence function $\Sigma_{11}^{-1} S_\beta$.

Similarly, we can show that there exists a solution $\widehat{\gamma}$ to Eq. 5 and $\widehat{\gamma}$ converges to γ_0 almost surely. From the Donsker property of the arguments, Eq. 5 implies

$$\begin{aligned}
0 = & \sqrt{n} \mathcal{P} \left\{ \int R(t) \left[X(t) \frac{g_m'(X(t)^T \gamma)}{g_m(X(t)^T \gamma)} - \frac{\mathcal{P}\{R(t)g_r(X(t)^T \beta)g_m'(X(t)^T \gamma)X(t)\}}{\mathcal{P}\{R(t)g_r(X(t)^T \beta)g_m(X(t)^T \gamma)\}} \right] dM(t) \right\} \\
& + \mathcal{G}_n \left\{ \int R(t) \left[X(t) \frac{g_m'(X(t)^T \gamma_0)}{g_m(X(t)^T \gamma_0)} - \frac{\mathcal{P}\{R(t)g_r(X(t)^T \beta_0)g_m'(X(t)^T \gamma_0)X(t)\}}{\mathcal{P}\{R(t)g_r(X(t)^T \beta_0)g_m(X(t)^T \gamma_0)\}} \right] dM(t) \right\} \\
& - \mathcal{G}_n \left\{ \int R(t)g_r(X(t)^T \beta_0)g_m'(X(t)^T \gamma_0)X(t) \frac{\mathcal{P}\{R(t)dM(t)\}}{\mathcal{P}\{R(t)g_r(X(t)^T \beta_0)g_m(X(t)^T \gamma_0)\}} \right\} \\
& + \mathcal{G}_n \left\{ \int R(t)g_r(X(t)^T \beta_0)g_m(X(t)^T \gamma_0) \right. \\
& \quad \times \left. \frac{\mathcal{P}\{R(t)dM(t)\}\mathcal{P}\{R(t)g_r(X(t)^T \beta_0)g_m'(X(t)^T \gamma_0)X(t)\}}{\mathcal{P}\{R(t)g_r(X(t)^T \beta_0)g_m(X(t)^T \gamma_0)\}^2} \right\} \\
& + o_p(1).
\end{aligned}$$

The Taylor expansion of the first term of the right-hand side yields

$$\sqrt{n}\Sigma_{22}(\widehat{\gamma} - \gamma_0) + \sqrt{n}\Sigma_{21}(\widehat{\beta} - \beta_0) = \mathcal{G}_n S_\gamma + o_p(1)$$

Therefore, $\widehat{\gamma}$ is an asymptotically linear estimator for γ_0 with influence function

$$\Sigma_{22}^{-1} \{S_\gamma - \Sigma_{21}\Sigma_{11}^{-1}S_\beta\}.$$

Hence, we conclude that $\sqrt{n}(\widehat{\beta} - \beta_0, \widehat{\gamma} - \gamma_0)$ converges in distribution to normal distribution with mean zero and covariance equal to

$$\mathcal{P} \left[\begin{pmatrix} \Sigma_{11}^{-1} S_\beta \\ \Sigma_{22}^{-1} \{S_\gamma - \Sigma_{21}\Sigma_{11}^{-1}S_\beta\} \end{pmatrix}^{\otimes 2} \right].$$

Clearly, a consistent estimator for this asymptotic covariance can be given by replacing $\mathcal{P}$ with $\mathcal{P}_n$ in the expressions of Σ_{11} , Σ_{12} , Σ_{22} , S_β and S_γ .

Proof of Theorem 2

We wish to derive the asymptotic distribution for $\hat{M}(s)$. Following the same argument as before, we obtain

$$\begin{aligned} & \sqrt{n} \{ \hat{M}(s) - E[M(s)] \} \\ &= \mathcal{G}_n \left[\int_0^s g_r(X(t)^T \beta_0) g_m(X(t)^T \gamma_0) \frac{\mathcal{P}\{R(t)dM(t)\}}{\mathcal{P}\{R(t)g_r(X(t)^T \beta_0)g_m(X(t)^T \gamma_0)\}} \right] \\ &+ \mathcal{G}_n \left[\int_0^s R(t)dM(t) \frac{\mathcal{P}\{g_r(X(t)^T \beta_0)g_m(X(t)^T \gamma_0)\}}{\mathcal{P}\{R(t)g_r(X(t)^T \beta_0)g_m(X(t)^T \gamma_0)\}} \right] \\ &- \mathcal{G}_n \left[\int_0^s R(t)g_r(X(t)^T \beta_0)g_m(X(t)^T \gamma_0) \right. \\ &\quad \times \left. \frac{\mathcal{P}\{R(t)dM(t)\}\mathcal{P}\{g_r(X(t)^T \beta_0)g_m(X(t)^T \gamma_0)\}}{\mathcal{P}\{R(t)g_r(X(t)^T \beta_0)g_m(X(t)^T \gamma_0)\}^2} \right] \\ &+ \nabla_\beta \left[\int_0^s \frac{\mathcal{P}\{R(t)dM(t)\}\mathcal{P}\{g_r(X(t)^T \beta_0)g_m(X(t)^T \gamma_0)\}}{\mathcal{P}\{R(t)g_r(X(t)^T \beta_0)g_m(X(t)^T \gamma_0)\}} \right] \mathcal{G}_n \Sigma_{11}^{-1} S_\beta \\ &+ \nabla_\gamma \left[\int_0^s \frac{\mathcal{P}\{R(t)dM(t)\}\mathcal{P}\{g_r(X(t)^T \beta_0)g_m(X(t)^T \gamma_0)\}}{\mathcal{P}\{R(t)g_r(X(t)^T \beta_0)g_m(X(t)^T \gamma_0)\}} \right] \\ &\quad \times \mathcal{G}_n \{ \Sigma_{22}^{-1} S_\gamma - \Sigma_{22}^{-1} \Sigma_{21} \Sigma_{11}^{-1} S_\beta \} + o_p(1), \end{aligned}$$

where $o_p(1)$ is in the metric space $l^\infty[0, \tau]$. Hence, the asymptotic Gaussian property for $\hat{M}(s)$ thus follows.

Additionally, we obtain the influence function for $\hat{M}(s)$ as

$$\begin{aligned} & \mathcal{J}(R, M, X, s; \mathcal{P}, \beta_0, \gamma_0) \\ &= \int_0^s g_r(X(t)^T \beta_0) g_m(X(t)^T \gamma_0) \frac{\mathcal{P}\{R(t)dM(t)\}}{\mathcal{P}\{R(t)g_r(X(t)^T \beta_0)g_m(X(t)^T \gamma_0)\}} \\ &+ \int_0^s R(t)dM(t) \frac{\mathcal{P}\{g_r(X(t)^T \beta_0)g_m(X(t)^T \gamma_0)\}}{\mathcal{P}\{R(t)g_r(X(t)^T \beta_0)g_m(X(t)^T \gamma_0)\}} \\ &- \int_0^s R(t)g_r(X(t)^T \beta_0)g_m(X(t)^T \gamma_0) \frac{\mathcal{P}\{R(t)dM(t)\}\mathcal{P}\{g_r(X(t)^T \beta_0)g_m(X(t)^T \gamma_0)\}}{\mathcal{P}\{R(t)g_r(X(t)^T \beta_0)g_m(X(t)^T \gamma_0)\}^2} \\ &+ \nabla_\beta \left[\int_0^s \frac{\mathcal{P}\{R(t)dM(t)\}\mathcal{P}\{g_r(X(t)^T \beta_0)g_m(X(t)^T \gamma_0)\}}{\mathcal{P}\{R(t)g_r(X(t)^T \beta_0)g_m(X(t)^T \gamma_0)\}} \right] \Sigma_{11}^{-1} S_\beta \\ &+ \nabla_\gamma \left[\int_0^s \frac{\mathcal{P}\{R(t)dM(t)\}\mathcal{P}\{g_r(X(t)^T \beta_0)g_m(X(t)^T \gamma_0)\}}{\mathcal{P}\{R(t)g_r(X(t)^T \beta_0)g_m(X(t)^T \gamma_0)\}} \right] \\ &\quad \times \{ \Sigma_{22}^{-1} S_\gamma - \Sigma_{22}^{-1} \Sigma_{21} \Sigma_{11}^{-1} S_\beta \}. \end{aligned}$$

Proof of Theorem 3

By Theorem 3.6.13 (van der Vaart and Wellner 1996), since $\mathcal{Q}(R, M, X, s; \mathcal{P}_n, \hat{\beta}, \hat{\gamma})$ belongs to a Donsker class and $\text{Var}(Q) = 1$, conditional on the observed data,

$$\sqrt{n}(\hat{M}_Q(s) - \hat{M}(s)) = n^{-1/2} \sum_{i=1}^n Q_i \mathcal{J}(R_i, M_i, X_i, s; \mathcal{P}_n, \hat{\beta}, \hat{\gamma}) \rightarrow_d \mathcal{G}.$$

Proof of Theorem 4

Following the same expansions as given in proving Theorems 2 and 3, we obtain

$$\begin{aligned}
 & \sqrt{n} \mathcal{T}_1(x) \\
 &= \mathcal{G}_n \left[\int R(t) I(X(t)^T \beta_0 \leq x) \left(dN(t) - g_r(X(t)^T \beta_0) d\Lambda_0(t) \right) \right] \\
 &= \mathcal{G}_n \left[\int R(t) dN(t) \frac{E[R(t) I(X(t)^T \beta_0 \leq x) g_r(X(t)^T \beta_0)]}{E[R(t) g_r(X(t)^T \beta_0)]} \right] \\
 &+ \mathcal{G}_n \left[\int R(t) g_r(X(t)^T \beta_0) \frac{E[R(t) I(X(t)^T \beta_0 \leq x) g_r(X(t)^T \beta_0)] E[R(t) dN(t)]}{E[R(t) g_r(X(t)^T \beta_0)]^2} \right] \\
 &- \nabla_{\beta} \Big|_{\beta=\beta_0} \left[\int \frac{E[R(t) I(X(t)^T \beta_0 \leq x) g_r(X(t)^T \beta)] E[R(t) dN(t)]}{E[R(t) g_r(X(t)^T \beta)]} \right] \\
 &\times \sqrt{n}(\widehat{\beta} - \beta_0) + o_p(1).
 \end{aligned}$$

Similarly,

$$\begin{aligned}
 & \sqrt{n} \mathcal{T}_2(y) \\
 &= \mathcal{G}_n \left[\int R(t) I(X(t)^T \gamma_0 \leq y) \left(dM(t) - g_r(X(t)^T \beta_0) g_m(X(t)^T \gamma_0) dV_0(t) \right) \right] \\
 &- \mathcal{G}_n \left[\int R(t) dM(t) \frac{E[R(t) I(X(t)^T \gamma_0 \leq y) g_r(X(t)^T \beta_0) g_m(X(t)^T \gamma_0)]}{E[R(t) g_r(X(t)^T \beta_0) g_m(X(t)^T \gamma_0)]} \right] \\
 &+ \mathcal{G}_n \left[\int R(t) g_r(X(t)^T \beta_0) g_m(X(t)^T \gamma_0) \right. \\
 &\quad \times \left. \frac{E[R(t) I(X(t)^T \gamma_0 \leq y) g_r(X(t)^T \beta_0) g_m(X(t)^T \gamma_0)] E[R(t) dM(t)]}{E[R(t) g_r(X(t)^T \beta_0) g_m(X(t)^T \gamma_0)]^2} \right] \\
 &- \nabla_{\beta} \Big|_{\beta=\beta_0} \left[\int \frac{E[R(t) I(X(t)^T \gamma_0 \leq y) g_r(X(t)^T \beta) g_m(X(t)^T \gamma_0)] E[R(t) dM(t)]}{E[R(t) g_r(X(t)^T \beta) g_m(X(t)^T \gamma_0)]} \right] \\
 &\quad \times \sqrt{n}(\widehat{\beta} - \beta_0) \\
 &- \nabla_{\gamma} \Big|_{\gamma=\gamma_0} \left[\int \frac{E[R(t) I(X(t)^T \gamma_0 \leq y) g_r(X(t)^T \beta_0) g_m(X(t)^T \gamma)] E[R(t) dM(t)]}{E[R(t) g_r(X(t)^T \beta_0) g_m(X(t)^T \gamma)]} \right] \\
 &\sqrt{n}(\widehat{\gamma} - \gamma_0) + o_p(1).
 \end{aligned}$$

After plugging the expansion of $\sqrt{n}(\widehat{\beta} - \beta_0)$ and $\sqrt{n}(\widehat{\gamma} - \gamma_0)$ into the last expression, we obtain

$$\sqrt{n} \mathcal{T}_1(x) = \mathcal{G}_n S_1(R, M, X, x) + o_p(1), \quad \sqrt{n} \mathcal{T}_2(y) = \mathcal{G}_n S_2(R, M, X, y) + o_p(1),$$

where

$$\begin{aligned}
 S_1(R, M, X, x) &= \int R(t) I(X(t)^T \beta_0 \leq x) \left(dN(t) - g_r(X(t)^T \beta_0) d\Lambda_0(t) \right) \\
 &- \int R(t) dN(t) \frac{E[R(t) I(X(t)^T \beta_0 \leq x) g_r(X(t)^T \beta_0)]}{E[R(t) g_r(X(t)^T \beta_0)]} \\
 &+ \int R(t) g_r(X(t)^T \beta_0) \frac{E[R(t) I(X(t)^T \beta_0 \leq x) g_r(X(t)^T \beta_0)] E[R(t) dN(t)]}{E[R(t) g_r(X(t)^T \beta_0)]^2} \\
 &- \nabla_{\beta} \Big|_{\beta=\beta_0} \left[\int \frac{E[R(t) I(X(t)^T \beta_0 \leq x) g_r(X(t)^T \beta)] E[R(t) dN(t)]}{E[R(t) g_r(X(t)^T \beta)]} \right] \Sigma_{11}^{-1} S_{\beta}
 \end{aligned}$$

and

$$\begin{aligned}
S_2(R, M, X, y) = & \int R(t)I(X(t)^T \gamma_0 \leq y) (dM(t) - g_r(X(t)^T \beta_0)g_m(X(t)^T \gamma_0)dV_0(t)) \\
& - \int R(t)dM(t) \frac{E[R(t)I(X(t)^T \gamma_0 \leq y)g_r(X(t)^T \beta_0)g_m(X(t)^T \gamma_0)]}{E[R(t)g_r(X(t)^T \beta_0)g_m(X(t)^T \gamma_0)]} \\
& + \int R(t)g_r(X(t)^T \beta_0)g_m(X(t)^T \gamma_0) \\
& \times \frac{E[R(t)I(X(t)^T \gamma_0 \leq y)g_r(X(t)^T \beta_0)g_m(X(t)^T \gamma_0)]E[R(t)dM(t)]}{E[R(t)g_r(X(t)^T \beta_0)g_m(X(t)^T \gamma_0)]^2} \\
& - \nabla_{\beta} \Big|_{\beta=\beta_0} \left[\int \frac{E[R(t)I(X(t)^T \gamma_0 \leq y)g_r(X(t)^T \beta)g_m(X(t)^T \gamma_0)]E[R(t)dM(t)]}{E[R(t)g_r(X(t)^T \beta)g_m(X(t)^T \gamma_0)]} \right] \\
& \times \Sigma_{11}^{-1} S_{\beta} \\
& - \nabla_{\gamma} \Big|_{\gamma=\gamma_0} \left[\int \frac{E[R(t)I(X(t)^T \gamma_0 \leq y)g_r(X(t)^T \beta_0)g_m(X(t)^T \gamma)]E[R(t)dM(t)]}{E[R(t)g_r(X(t)^T \beta_0)g_m(X(t)^T \gamma)]} \right] \\
& \times \Sigma_{22}^{-1} (S_{\gamma} - \Sigma_{21} \Sigma_{11}^{-1} S_{\beta}).
\end{aligned}$$

It is easy to verify $(S_1(R, M, X, x), S_2(R, M, X, y))$ belong to some Donsker classes in $l^{\infty}(R^2)$. Therefore, $\sqrt{n}(\mathcal{T}_1, \mathcal{T}_2)$ weakly converges to a mean-zero Gaussian process in $l^{\infty}(R^2)$. Particularity, the asymptotic covariance of this process is given by

$$Cov((S_1(R, M, X, x_1), S_2(R, M, X, y_1))^T, (S_1(R, M, X, x_2), S_2(R, M, X, y_2))^T).$$

Clearly, both $S_1(R, M, X, x)$ and $S_2(R, M, X, y)$ are consistently estimated if we replace expectation with empirical expression in $S_1(R, M, X, x)$ and $S_2(R, M, X, y)$ and replace $(\beta_0, \gamma_0, \Lambda_0, V_0)$ with their corresponding estimates.

References

- Barreto ML, Santos LMP, Assis AMO, Araujo MPN, Farenzena GG, Santos PAB, Fiaccone RL. Effect of vitamin A supplementation on diarrhoea and acute lower-respiratory-tract infections in young children in Brazil. *Lancet* 1994;344:228–231. [PubMed: 7913157]
- Deming, K. Nonlinear functional analysis. New York: Springer; 1985.
- Diggle, P. Statistical analysis of spatial point patterns. 2nd edn.. London: Arnold; 2003.
- Herring A, Yang J. Bayesian modeling of multiple episode occurrence and severity with a terminating event. *Biometrics* 2007;63:381–388. [PubMed: 17688490]
- Hu XJ, Sun J, Wei LJ. Regression parameter estimation from panel counts. *Scand J Stat* 2003;30:25–43.
- Huang Y, Berry K. Semiparametric estimation of marginal mark distribution. *Biometrika* 2006;93:895–910.
- Lin DY, Ying Z. Semiparametric and nonparametric regression analysis of longitudinal data. *J Am Stat Assoc* 2001;96:103–113.
- Lin DY, Wei LJ, Yang I, Ying Z. Semiparametric regression for the mean and rate functions of recurrent events. *J Roy Stat Soc B* 2000;62:711–730.
- Lin DY, Wei LJ, Ying Z. Semiparametric transformation models for point processes. *J Am Stat Assoc* 2001;96:620–628.
- Martinussen T, Scheike TH. A nonparametric dynamic additive regression model for longitudinal data. *Ann Stat* 2000;28:1000–1025.
- Martinussen T, Scheike TH. Sample adjusted analysis of dynamic additive regression for longitudinal data. *Scand J Stat* 2001;28:303–323.
- Ratcliffe S, Guo W, Ten Have TR. Joint modeling of longitudinal and survival data via a common frailty. *Biometrics* 2004;60:892–899. [PubMed: 15606409]
- Sun Y, Wu H. Semiparametric time-varying coefficients regression model for longitudinal data. *Scand J Stat* 2003;32:21–47.

- Tsiatis AA, Deguttola V, Wulfsohn MS. Modeling the relationship of survival to longitudinal data measured with error Applications to survival and CD 4. counts in patients with AIDS. *J Am Stat Assoc* 1995;90:27–37.
- van der Vaart, AW.; Wellner, JA. Weak convergence and empirical processes. New York: Springer; 1996.
- Wu MC, Bailey KR. Estimation and comparison of changes in the presence of informative right censoring by modeling the censoring process. *Biometrics* 1989;44:175–188.
- Xu J, Zeger SL. Joint analysis of longitudinal data comprising repeated measures and times to events. *Appl Stat* 2001;50:375–387.

Fig. 1.

Estimated cumulative severity score over time in VACT study: the upper solid curve is the estimated cumulative severity score in the placebo group and the lower solid curve is the estimated cumulative severity score in the treatment group. The dashed curves are the 95% point-wise confidence bands

Fig. 2. Test for goodness-of-fit. **a** corresponds to the exponential link and **b** corresponds to the square link. The x -axes correspond to linear predictors. The solid lines are the test statistics and the dashed lines are the curves simulated from the limiting distributions

Table 1

Summary of simulation results with correctly specified models

ρ	n	Parameter	True value	Est.	Emp. Se.	Est. Se	CP
$g_m(x) = e^x$							
$\rho = 0$	200	β_1	-1	-0.997	0.268	0.256	0.93
		β_2	1	0.999	0.153	0.145	0.94
		γ_1	0.5	0.495	0.213	0.199	0.94
		γ_2	0.8	0.799	0.124	0.117	0.94
	400	$E[M(\tau)]$	55.421	55.550	7.872	8.044	0.95
		β_1	-1	-0.994	0.177	0.183	0.96
		β_2	1	1.003	0.103	0.103	0.95
		γ_1	0.5	0.504	0.151	0.143	0.94
		γ_2	0.8	0.807	0.856	0.835	0.93
		$E[M(\tau)]$	55.421	55.698	5.369	5.801	0.96
$\rho = 0.5$	200	β_1	-1	-0.997	0.268	0.256	0.93
		β_2	1	0.999	0.153	0.145	0.94
		γ_1	0.5	0.495	0.286	0.257	0.92
		γ_2	0.8	0.800	0.152	0.140	0.92
	400	$E[M(\tau)]$	130.123	130.779	23.283	22.364	0.93
		β_1	-1	-0.994	0.177	0.183	0.96
		β_2	1	1.003	0.103	0.103	0.95
		γ_1	0.5	0.504	0.199	0.188	0.93
		γ_2	0.8	0.808	0.102	0.101	0.94
		$E[M(\tau)]$	130.123	131.002	15.692	16.295	0.95
$g_m(x) = (1 + x)^2$							
$\rho = 0$	200	β_1	-1	-0.997	0.268	0.256	0.93
		β_2	1	0.999	0.153	0.145	0.94
	400	γ_1	0.5	0.509	0.226	0.213	0.93
		γ_2	0.8	0.804	0.147	0.139	0.94
		$E[M(\tau)]$	78.983	79.135	11.518	11.670	0.95

ρ	n	Parameter	True value	Est.	Emp. Se.	Est. Se	CP
$\rho = 0.5$	400	β_1	-1	-0.994	0.177	0.183	0.96
		β_2	1	1.003	0.103	0.103	0.95
		γ_1	0.5	0.514	0.161	0.154	0.94
		γ_2	0.8	0.812	0.104	0.100	0.94
	200	$E[M(\tau)]$	78.983	79.393	7.845	8.425	0.96
		β_1	-1	-0.997	0.268	0.256	0.93
		β_2	1	0.999	0.153	0.145	0.94
		γ_1	0.5	0.524	0.303	0.278	0.91
		γ_2	0.8	0.810	0.186	0.171	0.93
	400	$E[M(\tau)]$	185.444	186.286	33.983	32.449	0.93
		β_1	-1	-0.994	0.177	0.183	0.96
		β_2	1	1.003	0.103	0.103	0.95
		γ_1	0.5	0.505	0.217	0.203	0.94
		γ_2	0.8	0.811	0.128	0.122	0.94
		$E[M(\tau)]$	185.444	186.735	22.898	23.679	0.95

Table 2

Summary of simulation results with misspecified link function

ρ	n	Parameter	True value	Est.	Emp. Se.	Est. Se	CP
$\rho = 0$	200	γ_1	0.5	0.526	0.220	0.205	0.94
		γ_2	0.8	1.007	0.127	0.120	0.55
	400	$E[M(\tau)]$	78.983	79.139	11.151	11.675	0.95
		γ_1	0.5	0.533	0.155	0.147	0.94
		γ_2	0.8	1.013	0.088	0.086	0.31
		$E[M(\tau)]$	78.983	79.389	7.840	8.424	0.96
$\rho = 0.5$	200	γ_1	0.5	0.533	0.297	0.265	0.91
		γ_2	0.8	1.009	0.155	0.143	0.64
	400	$E[M(\tau)]$	185.444	186.319	34.007	32.469	0.92
		γ_1	0.5	0.544	0.205	0.194	0.93
		γ_2	0.8	1.015	0.104	0.103	0.42
		$E[M(\tau)]$	185.444	186.728	22.881	23.678	0.95

Table 3

Summary of diarrheal episodes and severity in VACT data

Number of episodes	Vitamin A group	Placebo group
<i>Recurrent diarrheal episodes</i>		
0–3	271	256
4–7	157	153
8–11	89	94
12–15	49	54
≥ 16	39	45
<i>Total severity scores</i>		
(0, 50)	419	395
[50, 100)	89	101
[100, 200)	72	69
[200, ∞)	25	37

Table 4

Analysis of VACT data

Covariates	Recurrent event			Recurrent marker		
	Estimate	Std error	P-value	Estimate	Std error	P-value
$g_m(x) = e^x$						
Treatment versus placebo	-0.0816	0.0459	0.075	-0.0649	0.0536	0.226
Male versus female	0.0267	0.0461	0.562	0.0142	0.0525	0.787
Age	-0.0303	0.0019	< 0.001	-0.0123	0.0021	< 0.001
$g_m(x) = (1 + x)^2$						
Treatment versus placebo	-0.0816	0.0459	0.075	-0.0265	0.0222	0.233
Male versus female	0.0267	0.0461	0.562	0.0045	0.0222	0.839
Age	-0.0303	0.0019	< 0.001	-0.0053	0.0008	< 0.001