HHS Public Access
Author manuscript
Stat Methods Med Res. Author manuscript; available in PMC 2022 October 01.

Published in final edited form as:
Stat Methods Med Res. 2021 October ; 30(10): 2239–2255. doi:10.1177/09622802211027593.

Additive rates model for recurrent event data with intermittently observed time-dependent covariates

Tianmeng Lyu[1], **Xianghua Luo**[2], **Chiung-Yu Huang**[3], **Yifei Sun**[4]

[1]Clinical Development & Analytics, Novartis Pharmaceuticals Corporation, East Hanover, NJ, USA

[2]Division of Biostatistics, School of Public Health, University of Minnesota, Minneapolis, MN, USA

[3]Department of Epidemiology and Biostatistics, University of California, San Francisco, CA, USA

[4]Department of Biostatistics, Mailman School of Public Health, Columbia University, New York, NY, USA

Abstract

Various regression methods have been proposed for analyzing recurrent event data. Among them, the semiparametric additive rates model is particularly appealing because the regression coefficients quantify the absolute difference in the occurrence rate of the recurrent events between different groups. Estimation of the additive rates model requires the values of time-dependent covariates being observed throughout the entire follow-up period. In practice, however, the time-dependent covariates are usually only measured at intermittent follow-up visits. In this paper, we propose to kernel smooth functions involving time-dependent covariates across subjects in the estimating function, as opposed to imputing individual covariate trajectories. Simulation studies show that the proposed method outperforms simple imputation methods. The proposed method is illustrated with data from an epidemiologic study of the effect of streptococcal infections on recurrent pharyngitis episodes.

Keywords

Kernel smoothing; recurrent events; time-dependent covariates; additive rates models; estimating equations

1 Introduction

Recurrent event data are frequently encountered in clinical and epidemiological studies, where each subject can experience events of interest repeatedly. Examples of recurrent

Corresponding author: Yifei Sun, Columbia University Mailman School of Public Health, 722 West 168th St R653, New York 10032-3727, NY, USA. ys3072@cumc.columbia.edu.

Declaration of conflicting interests
The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Supplemental material
Supplemental material for this article is available online. It includes the details of the testing procedure of the additivity assumption. The R package implementing the proposed method can be downloaded at www.github.com/TianmengL/rectime.

events include infections after hematopoietic cell transplantations,[1] repeated cardiovascular events,[2] and rehospitalizations of patients with psychiatric disorders.[3] In such studies, data on time-dependent covariates are often collected during the course of follow-up. Regression methods that can handle time-dependent covariates have been an important tool as investigators are often interested in evaluating the effect of variables that are evolving over time such as in studies of personalized medicine. Researchers may also be interested in utilizing updated information on risk factors during follow-up in dynamic prediction of event risk.

The motivating example of this research is an observational study conducted in India between 2002 and 2004 which aimed to evaluate the effects of time-varying streptococcal infections, including group A, C, G streptococcus, on the risk of the recurrent pharyngitis.[4] In this study, participants were examined weekly for the symptoms of pharyngitis and throat swabs were obtained to identify the status of streptococcal infections in symptomatic patients. In addition, monthly visits were scheduled to determine the carriage rate of streptococcal infections in this population. In the analysis, the infections of streptococcal groups were regarded as time-varying risk factors of the recurrent pharyngitis occurrence.

Regression methods for recurrent events are usually formulated based on either the conditional *intensity* or the marginal *rate* function of the counting process of recurrent events. Andersen and Gill[5] and Prentice et al.[6] proposed a proportional intensity model, which postulates a multiplicative covariate effect on the intensity function of the underlying counting process, that is the instantaneous risk of recurrent event conditional on the event history and covariate history. Alternatively, Pepe and Cai[7] and Lin et al.[8] proposed proportional rates models which are based on the marginal rate function. Although the proportional intensity or rate models have gained great popularity in applications, they assume the covariates to have multiplicative effects on the recurrent event risk. In applications, it is possible that the covariate effects add to, instead of multiplying, the baseline event risk. In this case, it would be more appropriate to use additive models such as the semiparametric additive rates model by Schaubel et al.[9] and the additive intensity model by Liu and Wu.[10] Moreover, the additive models can provide the risk difference estimates which are especially relevant and desired in epidemiological and clinical studies.

Although the aforementioned recurrent event models can naturally accommodate time-dependent covariates, their model estimation procedures require the values of time-dependent covariates to be continuously observed throughout the entire follow-up period for all subjects. In practice, however, the time-dependent covariates are often intermittently measured, rendering the existing model estimation procedures not readily applicable. A number of approaches to handle intermittently measured covariates have been discussed and reviewed in Andersen and Liestøl.[11] The first type of methods is a two-stage approach where the values of time-dependent covariates are estimated in the first stage and then the estimated covariate values are used in the regression model in the second stage. Simple methods for the first stage include carrying forward the last observed value or imputing the missing values between two observation times by linear interpolation. More complex methods such as parametric or non-parametric smoothing techniques,[12] random effects model,[13] and stochastic models[14,15] have also been considered.

The second type of methods involves jointly modeling longitudinally measured covariates and event times. When the event time is univariate, various joint models have been proposed including selection models, pattern mixture models, and shared parameter models. Readers are referred to Tsiatis and Davidian[16] and Rizopoulos[17] for comprehensive reviews. When the event time is recurrent, Henderson et al.[18] modeled the covariate and recurrent event processes jointly via a latent bivariate Gaussian process, whereas Li[19] considered a joint model of the recurrent event process and the longitudinal process for binary covariate specifically. Others considered joint models in the presence of a terminal event.[20-22] The estimation of the joint models could be computationally intensive, especially when the longitudinal covariates are multi-dimensional or include categorical variables. In addition, the validity of joint modeling relies on certain assumptions about the covariate model and the dependence structure of the repeatedly measured covariates, which may be difficult to verify. Misspecification of the model for longitudinal measurements will result in biased estimation of the event time model.

Recently, Cao et al.[23] and Li et al.[24] proposed kernel-weighted estimation procedures for the proportional hazards/rates models with time-dependent covariates. Specifically, Cao et al.[23] considered the case where data on covariates are not collected at failure times and proposed to smooth the partial likelihood to derive a consistent estimator with a convergence rate slower than root-n. Li et al.[24] focused on the setting of recurrent event data where, in addition to regular follow-up visits, covariate values are usually collected when an event occurs. As pointed out by these authors, measurements at event visits give a biased representation of the underlying covariate process of an individual. Hence, in the construction of kernel-smoothed pseudo-partial score functions, only covariate values measured at regular visits, whose timing is noninformative of the underlying recurrent event risk, are used to estimate the expected covariate value of individuals in a risk set. The estimated score function gives a consistent estimator with a root-n convergence rate.

In this paper, we propose a semiparametric estimator for the additive rates model with intermittently observed time-dependent covariates. Specifically, we kernel smooth functions of time-dependent covariates across subjects instead of smoothing individual covariate trajectories. Our proposed method is demonstrated to have better performance than simple covariate imputation methods such as the last covariate carried forward (LCCF) method through simulation studies. We also discuss a few practical issues including the situation when both time-dependent and time-independent covariates are present and the case when different time-dependent covariates are measured at different times.

The remainder of this paper is organized as follows. In Section 2, we first review the additive rates model and the estimation procedure[9] in the ideal case where time-dependent covariates are monitored continuously, then we present the proposed kernel smoothed estimator for the case where covariates are time-dependent and intermittently observed. Some extensions of the proposed method are discussed in Section 3. Section 4 compares the performance of the proposed estimator to the two simple approaches including the LCCF and linear interpolation methods with simulation studies. In Section 5, we present a real data analysis using the Indian pharyngitis data. Some concluding remarks are included in Section 6.

2 Model and the proposed estimator

Let $i = 1, \ldots, n$ index the n subjects in a study. Let $N_i^*(t)$ denote the number of events that subject i has experienced at or prior to time t in the absence of censoring. Denote by $\mathbf{Z}_i(t) = (Z_{i1}(t), \ldots, Z_{ip}(t))^\top$ a $p \times 1$ vector of possibly time-dependent covariates. The semiparametric additive rates model assumes that the rate function of $N_i^*(t)$ conditioning on the covariates at time t is

$$\lambda\{t \mid \mathbf{Z}_i(t)\} = \lambda_0(t) + \boldsymbol{\beta}^\top \mathbf{Z}_i(t)$$

where $\lambda_0(t)$ is an unspecified baseline rate function and $\boldsymbol{\beta} = (\beta_1, \ldots, \beta_p)^\top$ is a $p \times 1$ vector of regression parameters, whose jth component β_j is interpreted as the rate difference associated with one unit difference in $Z_{ij}(t)$. Let C_i denote the follow-up time for subject i and define $Y_i(t) = I(C_i \geq t)$. Let $N_i(t) = N_i^*(t \wedge C_i)$, where $t \wedge C_i = \min(t, C_i)$ be the number of observed events up to C_i. Let τ be a pre-specified time point such that the recurrent event process could potentially be observed beyond τ with a non-zero probability. The observed data $\{N_i(\cdot), \mathbf{Z}_i(\cdot), Y_i(\cdot)\}$, $i = 1, \ldots, n$, are assumed to be independent and identically distributed.

For model estimation, following Lin and Ying,[25] Schaubel et al.[9] considered the estimating function

$$\begin{aligned} U(\boldsymbol{\beta}) &= n^{-1}\sum_{i=1}^{n}\int_0^\tau \{\mathbf{Z}_i(t) - \bar{\mathbf{Z}}(t)\}\{dN_i(t) - Y_i(t)\boldsymbol{\beta}^\top \mathbf{Z}_i(t)dt\} \\ &= n^{-1}\sum_{i=1}^{n}\int_0^\tau \{\mathbf{Z}_i(t) - \bar{\mathbf{Z}}(t)\}dN_i(t) \\ &\quad - \left[n^{-1}\sum_{i=1}^{n}\int_0^\tau \{\mathbf{Z}_i(t) - \bar{\mathbf{Z}}(t)\}^{\otimes 2} Y_i(t)dt\right]\boldsymbol{\beta} \end{aligned} \tag{1}$$

where $\bar{\mathbf{Z}}(t) = \{n^{-1}\sum_{i=1}^{n} Y_i(t)\mathbf{Z}_i(t)\} / \{n^{-1}\sum_{i=1}^{n} Y_i(t)\}$, $\mathbf{z}^{\otimes 0} = 1$, $\mathbf{z}^{\otimes 1} = \mathbf{z}$, $\mathbf{z}^{\otimes 2} = \mathbf{z}\mathbf{z}^\top$. Solving $U(\boldsymbol{\beta}) = 0$ gives a simple closed-form solution

$$\widehat{\boldsymbol{\beta}} = \left[n^{-1}\sum_{i=1}^{n}\int_0^\tau \{\mathbf{Z}_i(t) - \bar{\mathbf{Z}}(t)\}^{\otimes 2} Y_i(t)dt\right]^{-1}\left[n^{-1}\sum_{i=1}^{n}\int_0^\tau \{\mathbf{Z}_i(t) - \bar{\mathbf{Z}}(t)\}dN_i(t)\right] \tag{2}$$

It is easy to see that the estimator in equation (2), in particular the denominator, requires the time-dependent covariates to be continuously observed throughout the follow-up period. In practice, however, time-dependent covariates are often only observed intermittently. For example, in the pharyngitis data that motivates this research, the bacterial infection status of patients was only identified monthly. In such case, the estimator in equation (2) is not evaluable with the observed data.

A simple method for handling intermittently observed time-dependent covariates is to impute unobserved values using the LCCF approach. Under LCCF, the last known value of the covariate of a subject is used forward in time until a new value is measured or the observation of this subject is censored. This method has been shown to yield biased estimation under the proportional rates model.[24] Another simple approach is to use linear interpolation to estimate covariate values between two observations within each subject. Instead of imputing missing values in the individual covariate trajectories, we propose a method focusing on smoothing the estimating function using the observed covariate information across subjects.

We first consider the simple case where all covariates in the model are time-dependent and observed at the same regular visits. More general cases such as when both time-dependent and -independent covariates are present in the model or when multiple time-dependent covariates are measured at different regular visits are discussed in Section 3. Let $O(\cdot)$ denote the counting process for the regular visits, where regular visits are referred to as prescheduled follow-up visits, and when a regular visit occurs, $O(\cdot)$ jumps by 1. In the Indian pharyngitis study, $O(t)$ is a function with unit steps at the monthly visits. Since the participants may be sick at a regular visit, we allow $O(\cdot)$ and $N(\cdot)$ to jump at the same time. We assume that the process $O(\cdot)$ is independent of $\mathbf{Z}(\cdot)$ and C. The rate function of $O(\cdot)$ is denoted by $m(t)$, that is, $\mathrm{E}\{dO(t)\} = m(t)dt$.

Let $S^{(k)}(t) = n^{-1}\sum_{i=1}^{n} Y_i(t)\mathbf{Z}_i(t)^{\otimes k}$, $k = 0, 1, 2$. It is easy to show that the estimating function in equation (1) can be re-expressed as

$$U(\boldsymbol{\beta}) = n^{-1}\sum_{i=1}^{n}\int_0^{\tau}\mathbf{Z}_i(t)dN_i(t) - \int_0^{\tau}\frac{S^{(1)}(t)}{S^{(0)}(t)}\left\{n^{-1}\sum_{i=1}^{n} dN_i(t)\right\} - \left(\int_0^{\tau}\left[n^{-1}\sum_{i=1}^{n} Y_i(t)\frac{S^{(2)}(t)}{S^{(0)}(t)} - n^{-1}\sum_{i=1}^{n} Y_i(t)\left\{\frac{S^{(1)}(t)}{S^{(0)}(t)}\right\}^{\otimes 2}\right]dt\right)\boldsymbol{\beta} \quad (3)$$

Thus, $\widehat{\boldsymbol{\beta}}$ can be expressed as

$$\widehat{\boldsymbol{\beta}} = \left(\int_0^{\tau}\left[n^{-1}\sum_{i=1}^{n} Y_i(t)\frac{S^{(2)}(t)}{S^{(0)}(t)} - n^{-1}\sum_{i=1}^{n} Y_i(t)\left\{\frac{S^{(1)}(t)}{S^{(0)}(t)}\right\}^{\otimes 2}\right]dt\right)^{-1} \times\left[n^{-1}\sum_{i=1}^{n}\int_0^{\tau}\mathbf{Z}_i(t)dN_i(t) - \int_0^{\tau}\frac{S^{(1)}(t)}{S^{(0)}(t)}\left\{n^{-1}\sum_{i=1}^{n} dN_i(t)\right\}\right] \quad (4)$$

As can be seen from equation (4), the estimator $\widehat{\boldsymbol{\beta}}$ is a functional of the empirical processes $n^{-1}\sum_{i=1}^{n}\mathbf{Z}_i(t)dN_i(t)$, $n^{-1}\sum_{i=1}^{n} dN_i(t)$, and $S^{(k)}(t)$, $k = 0, 1, 2$. We assume that the covariates of subject i, $\mathbf{Z}_i(t)$, are observed at this subject's event times (that is, where $N_i(t)$ jumps), which is typically satisfied in recurrent event data, such as in the Indian pharyngitis data example. Hence, the empirical processes $n^{-1}\sum_{i=1}^{n}\mathbf{Z}_i(t)dN_i(t)$ and $n^{-1}\sum_{i=1}^{n} dN_i(t)$ can be computed based on the observed data. However, we note that the processes $S^{(k)}(t)/S^{(0)}(t)$, k = 1, 2 cannot be evaluated when the time-dependent covariates are not continuously observed.

In what follows, we show how to approximate the ratios $S^{(k)}(t)/S^{(0)}(t)$, k = 1, 2, using intermittently observed time-dependent covariate data. Let $s^{(k)}(t)$ denote the expectation of $S^{(k)}(t)$: $s^{(k)}(t) \equiv \mathrm{E}\{S^{(k)}(t)\} = \mathrm{E}\{Y_i(t)\mathbf{Z}_i(t)^{\otimes k}\}$. We aim to find a consistent estimator of $s^{(k)}(t)/s^{(0)}(t)$ to approximate $S^{(k)}(t)/S^{(0)}(t)$. We propose to apply the kernel smoothing method to estimate $s^{(k)}(t)/s^{(0)}(t)$ as follows. Define the kernel smoothed process

$$\widehat{S}_h^{(k)}(t) = n^{-1}\sum_{i=1}^{n}\int_0^{\tau} K_h(t-u)\,Y_i(u)\mathbf{Z}_i(u)^{\otimes k} dO_i(u),\ t \in [h, \tau - h]$$

for k = 0, 1, 2, where $K_h(t) = K(t/h)/h$, h is the bandwidth, $0 < h < \tau/2$, and $K(t)$ is a second-order kernel function with support [−1, 1]. In order to avoid the bias in the boundary region, we let $\widehat{S}_h^{(k)}(t) = \widehat{S}_h^{(k)}(h)$ for $t \in [0, h)$, $\widehat{S}_h^{(k)}(t) = \widehat{S}_h^{(k)}(\tau - h)$ for $t \in (\tau - h, \tau]$. One can prove that $\widehat{S}_h^{(k)}(t)$ converges in probability to the limit $s^{(k)}(t) \cdot m(t)$, where $m(t) = \mathrm{E}\{dO_i(t)\}/dt$ is the rate function of the observation process and usually considered as a nuisance. Therefore, by kernel smoothing all $S^{(k)}(t)$, including $S^{(0)}(t)$, we can construct $\widehat{\xi}_h^{(k)}(t) \equiv \widehat{S}_h^{(k)}(t) / \widehat{S}_h^{(0)}(t)$, which converges in probability to $\{s^{(k)}(t)m(t)\}/\{s^{(0)}(t)m(t)\} = s^{(k)}(t)/s^{(0)}(t)$ as $n \to \infty$. Note that although $S^{(0)}(t)$ can be calculated directly from the observed data, we apply the same kernel smoothing technique on it as for $S^{(1)}(t)$ and $S^{(2)}(t)$ to circumvent the estimation of the nuisance $m(t)$. We also note that, in the construction of the kernel smoothed estimator $\widehat{\xi}_h^{(k)}(t)$, we only utilize the covariate values measured at regular visits (through $O_i(t)$). The covariate values measured at the event times (i.e. when $dN_i(t) = 1$) are only used in the evaluation of $n^{-1}\sum_{i=1}^{n}\mathbf{Z}_i(t)dN_i(t)$ in the estimating function (3).

Now, we replace $S^{(k)}(t)/S^{(0)}(t)$ with $\widehat{\xi}_h^{(k)}(t)$ in equation (3) to obtain the following kernel estimating function

$$\begin{aligned}\widehat{U}_h(\boldsymbol{\beta}) = {} & n^{-1}\sum_{i=1}^{n}\int_0^{\tau}\mathbf{Z}_i(t)dN_i(t) - \int_0^{\tau}\widehat{\xi}_h^{(1)}(t)\left\{n^{-1}\sum_{i=1}^{n}dN_i(t)\right\} \\ & - \left[\int_0^{\tau}\left\{n^{-1}\sum_{i=1}^{n}Y_i(t)\widehat{\xi}_h^{(2)}(t) - n^{-1}\sum_{i=1}^{n}Y_i(t)\widehat{\xi}_h^{(1)}(t)^{\otimes 2}\right\}dt\right]\boldsymbol{\beta}\end{aligned} \tag{5}$$

Solving $\widehat{U}_h(\boldsymbol{\beta}) = 0$ leads to the proposed estimator of $\boldsymbol{\beta}$

$$\begin{aligned}\widehat{\boldsymbol{\beta}}_h = {} & \left[\int_0^{\tau}\left\{n^{-1}\sum_{i=1}^{n}Y_i(t)\widehat{\xi}_h^{(2)}(t) - n^{-1}\sum_{i=1}^{n}Y_i(t)\widehat{\xi}_h^{(1)}(t)^{\otimes 2}\right\}dt\right]^{-1} \\ & \times\left[n^{-1}\sum_{i=1}^{n}\int_0^{\tau}\mathbf{Z}_i(t)dN_i(t) - \int_0^{\tau}\widehat{\xi}_h^{(1)}(t)\left\{n^{-1}\sum_{i=1}^{n}dN_i(t)\right\}\right]\end{aligned} \tag{6}$$

We summarize the large sample property of $\hat{\boldsymbol{\beta}}_h$ in the following theorem. The detailed proof is given in Appendix 1.

Theorem 1. Let $\boldsymbol{\beta}_0$ denote the true parameter. Under regularity conditions (1) to (9) in Appendix 1, as $n \to \infty$, $\sqrt{n}(\hat{\boldsymbol{\beta}}_h - \boldsymbol{\beta}_0)$ converges in distribution to a normal random variable with zero mean and variance Σ, where Σ is defined in Appendix 1, on the condition that $h = O(n^{-v})$, $1/4 < v < 1/2$.

The asymptotic variance of $\hat{\beta}$ involves unknown nuisance functions such as $s^{(1)}(t)$, $s^{(2)}(t)$, and $m(t)$. To estimate the variance based on the asymptotic variance formula, these unknown functions need to be nonparametrically estimated using kernel smoothing. Hence, bootstrap is recommended for variance estimation because of its better finite-sample performance.

For the estimation of the baseline mean function $\mu_0(t) = \int_0^t \lambda_0(u)du$, following Schaubel et al.,[9] we have

$$
\begin{aligned}
\hat{\mu}_0(t, \boldsymbol{\beta}) &= \int_0^t \frac{\sum_{i=1}^n Y_i(u)\{dN_i(u) - \boldsymbol{\beta}^\top \mathbf{Z}_i(u)du\}}{\sum_{i=1}^n Y_i(u)} \\
&= \sum_{i=1}^n \int_0^t \frac{1}{Y_\cdot(t)} Y_i(u)dN_i(u) - \int_0^t \boldsymbol{\beta}^\top \frac{S^{(1)}(u)}{S^{(0)}(u)} du
\end{aligned}
$$

with $Y_\cdot(t) = \sum_{i=1}^n Y_i(t)$. As discussed before, $S^{(1)})(t)/S^{(0)}(t)$ cannot be evaluated directly with observed data. We consider the following estimator

$$
\hat{\mu}_{0,h}(t, \hat{\boldsymbol{\beta}}_h) = \sum_{i=1}^n \int_0^t \frac{1}{Y_\cdot(u)} Y_i(u)dN_i(u) - \int_0^t \hat{\boldsymbol{\beta}}_h^\top \hat{\xi}_h^{(1)}(u)\, du
$$

Note that $\hat{\mu}_{0,h}(t, \hat{\boldsymbol{\beta}}_h)$ may not give a nondecreasing function because the increment could be negative, especially for the time interval without observed recurrent events. To ensure monotonicity, we propose to estimate the baseline mean function by $\tilde{\mu}_{0,h}(t, \hat{\boldsymbol{\beta}}_h) = \max_{0 \le u \le t} \hat{\mu}_{0,h}(u, \hat{\boldsymbol{\beta}}_h)$.

3 Extensions of the proposed estimator

3.1 Estimation when both time-dependent and -independent covariates are present

Thus far, our discussions focus on the estimation of models with time-dependent covariates only. In practice, however, it is common to collect data on both time-dependent and time-independent covariates. One may be interested in the effect of a time-dependent covariate adjusting for baseline variables or vice versa, for example, the effect of a time-varying biomarker adjusting for sex or the effect of a randomized treatment adjusting for a time-varying adjuvant treatment. Note that the proposed method in Section 2 can be applied to the scenario where both time-dependent and -independent covariates are present. However, instead of kernel smoothing, it is more natural to estimate the mean covariate processes that only involve time-independent covariates with their simple empirical averages. In this section, we present a more appropriate method to deal with the two types of covariates.

Let $\mathbf{Z}_i(t) = (Z_{i1}(t),\ldots, Z_{ip}(t))^\top$ denote the vector of time-dependent covariates and $\mathbf{W}_i = (W_{i1}, \ldots, W_{iq})^\top$ the vector of time-independent covariates. Then the additive rates model can be expressed as

$$\lambda\{t \mid \mathbf{Z}_i(t), \mathbf{W}_i\} = \lambda_0(t) + \boldsymbol{\beta}^\top \mathbf{Z}_i(t) + \boldsymbol{\gamma}^\top \mathbf{W}_i$$

where $\boldsymbol{\beta}$ and $\boldsymbol{\gamma}$ are $p \times 1$ and $q \times 1$ vectors of parameters for the time-dependent covariates and the time-independent covariates, respectively.

The estimating function for $(\boldsymbol{\beta}, \boldsymbol{\gamma})$ is given by

$$U(\boldsymbol{\beta},\boldsymbol{\gamma}) = n^{-1}\sum_{i=1}^{n}\int_0^\tau (\mathbf{Z}_i(t)^\top, \mathbf{W}_i^\top)^\top dN_i(t) - \int_0^\tau (\bar{\mathbf{Z}}(t)^\top, \bar{\mathbf{W}}(t)^\top)^\top \left\{n^{-1}\sum_{i=1}^{n} dN_i(t)\right\}$$
$$-\left[\int_0^\tau n^{-1}\sum_{i=1}^{n} Y_i(t)\left\{\frac{n^{-1}\sum_{i=1}^{n} Y_i(t)(\mathbf{Z}_i(t)^\top, \mathbf{W}_i^\top)^{\top\otimes 2}}{S^{(0)}(t)} - \left(\bar{\mathbf{Z}}(t)^\top, \bar{\mathbf{W}}(t)^\top\right)^{\top\otimes 2}\right\}dt\right](\boldsymbol{\beta}^\top, \boldsymbol{\gamma}^\top)^\top$$

where $\bar{\mathbf{W}}(t) = \{n^{-1}\sum_{i=1}^{n} Y_i(t)\mathbf{W}_i\} / \{n^{-1}\sum_{i=1}^{n} Y_i(t)\}$. For $k = 1, 2$, define $S_z^{(k)}(t) = n^{-1}\sum_{i=1}^{n} Y_i(t)\mathbf{Z}_i(t)^{\otimes k}$, $S_w^{(k)}(t) = n^{-1}\sum_{i=1}^{n} Y_i(t)\mathbf{W}_i^{\otimes k}$, and

$$S^{(2)}(t) = \begin{pmatrix} S_z^{(2)}(t) & S_{zw}^{(2)}(t) \\ S_{wz}^{(2)}(t) & S_w^{(2)}(t) \end{pmatrix}$$

where $S_{zw}^{(2)}(t) = n^{-1}\sum_{i=1}^{n} Y_i(t)\mathbf{Z}_i(t)\mathbf{W}_i^\top$, and $S_{wz}^{(2)}(t) = n^{-1}\sum_{i=1}^{n} Y_i(t)\mathbf{W}_i\mathbf{Z}_i(t)^\top$. The estimating function can be reexpressed as

$$\begin{aligned} U(\boldsymbol{\beta},\boldsymbol{\gamma}) &= n^{-1}\sum_{i=1}^{n}\int_0^\tau (\mathbf{Z}_i(t)^\top, \mathbf{W}_i^\top)^\top dN_i(t) \\ &- \int_0^\tau \left(\frac{S_z^{(1)}(t)^\top}{S^{(0)}(t)}, \bar{\mathbf{W}}(t)^\top\right)^\top \left\{n^{-1}\sum_{i=1}^{n} dN_i(t)\right\} \\ &\quad - \left[\int_0^\tau n^{-1}\sum_{i=1}^{n} Y_i(t)\left\{\frac{S^{(2)}(t)}{S^{(0)}(t)} - \left(\frac{S_z^{(1)}(t)^\top}{S^{(0)}(t)}, \bar{\mathbf{W}}(t)^\top\right)^{\top\otimes 2}\right\}dt\right](\boldsymbol{\beta}^\top, \boldsymbol{\gamma}^\top)^\top \end{aligned} \tag{7}$$

Note that when the time-dependent covariates $\mathbf{Z}_i(t)$ are observed intermittently, a few quantities in equation (7) are not evaluable: $S_z^{(k)}(t) / S^{(0)}(t)$, $k = 1, 2$, $S_{wz}^{(2)}(t) / S^{(0)}$ and

$S_{zw}^{(2)}(t)\,/\,S^{(0)}$, whereas the values of $S_w^{(k)}(t)$, $k = 1, 2$, and $\bar{\mathbf{W}}(t)$ are known for all time t since $\mathbf{W}_i$ are time-independent. We can use the kernel smoothed processes $\hat{S}_h^{(k)}(t)$ defined in Section 2 to replace $S_z^{(k)}(t)$ for $k = 1, 2$ and the same $\hat{S}_h^{(0)}(t)$ to replace $S^{(0)}(t)$. Further, we propose the kernel smoothed processes $\hat{S}_{zw,h}^{(2)}(t) = n^{-1}\sum_{i=1}^{n}\int_0^{\tau}K_h(t-u)\,Y_i(u)\mathbf{Z}_i(u)\mathbf{W}_i^{\top}dO_i(u)$ and $\hat{S}_{wz,h}^{(2)}(t) = n^{-1}\sum_{i=1}^{n}\int_0^{\tau}K_h(t-u)\,Y_i(u)\mathbf{W}_i\mathbf{Z}_i(u)^{\top}dO_i(u)$, $t \in [h, \tau - h]$. Similar boundary corrections as described in Section 2 are applied. Then, we propose the estimating function, $\hat{U}_h(\boldsymbol{\beta}, \boldsymbol{\gamma})$, by replacing the non-observable quantities in equation (7) specified above with their kernel smoothed counterparts.

3.2 Estimation when multiple time-dependent covariates are measured on different schedules

In the previous sections, we assume that multiple time-dependent covariates are observed simultaneously at the same visits (i.e. synchronized). In practice, it is possible that these covariates are measured on different schedules. For example, in social behavioral studies, in order to prevent survey fatigue and maintain a high retention rate, different surveys may be delivered at different visits. In this section, we discuss how to extend the proposed method to accommodate multiple time-dependent covariates measured on different schedules. For ease of discussion, we assume that there are two time-dependent covariates ($p = 2$), $Z_{i1}(t)$ and $Z_{i2}(t)$. The proposed method can be easily extended to the case where $p > 2$.

Let $O_{i1}(t)$ and $O_{i2}(t)$ denote the bivariate observation process that counts the cumulative number of measurements of $Z_{i1}(t)$ and $Z_{i2}(t)$, respectively. We assume that $\{O_1(\cdot), O_2(\cdot)\}$ is independent of $\{\mathbf{Z}(\cdot), C\}$, $\mathrm{E}\{dO_{i1}(u)dO_{i2}(w)\} = m_{12}(u, w)dudw$, $\mathrm{E}\{dO_{ig}(u)\} = m_g(u)du$ for $g = 1, 2$. For $k = 0, 1, 2$, define $\hat{S}_{g,h}^{(k)}(t) = n^{-1}\sum_{i=1}^{n}\int_0^{\tau}K_h(t-u)\,Y_i(u)Z_{ig}(u)^k dO_{ig}(u)$, which consistently estimate $\mathrm{E}\{Y(t)Z_g^k(t)\}m_g(t)$. It is easy to see that $S^{(1)}(t)/S^{(0)}(t)$ in equation (4) can be replaced by $\tilde{S}^{(1)}(t)\,/\,\tilde{S}^{(0)}(t) = \left(\hat{S}_{1,h}^{(1)}(t)\,/\,\hat{S}_{1,h}^{(0)}(t), \hat{S}_{2,h}^{(1)}(t)\,/\,\hat{S}_{2,h}^{(0)}(t)\right)^{\top}$. Moreover, the matrix $S^{(2)}(t)/S^{(0)}(t)$ in equation (4) is

$$\begin{pmatrix} \sum_{i=1}^{n} Y_i(t)Z_{i1}(t)^2 \,/\, \sum_{i=1}^{n} Y_i(t) & \sum_{i=1}^{n} Y_i(t)Z_{i1}(t)Z_{i2}(t) \,/\, \sum_{i=1}^{n} Y_i(t) \\ \sum_{i=1}^{n} Y_i(t)Z_{i1}(t)Z_{i2}(t) \,/\, \sum_{i=1}^{n} Y_i(t) & \sum_{i=1}^{n} Y_i(t)Z_{i2}(t)^2 \,/\, \sum_{i=1}^{n} Y_i(t) \end{pmatrix}$$

As before, for $g = 1, 2$, the diagonal entries $\sum_{i=1}^{n} Y_i(t)Z_{ig}(t)^2 \,/\, \sum_{i=1}^{n} Y_i(t)$ can be replaced by the kernel type estimators $\hat{S}_{g,h}^{(2)}(t)\,/\,\hat{S}_{g,h}^{(0)}(t)$. The off-diagonal entries involve both $Z_1(t)$ and $Z_2(t)$, and thus we consider the following bivariate kernel type estimator $\hat{S}_{12,h}^{(2)}(t)\,/\,\hat{S}_{12,h}^{(0)}(t)$, where

$$\widehat{S}^{(2)}_{12,h}(t) = n^{-1}\sum_{i=1}^{n}\int_0^{\tau}\int_0^{\tau}K_h(t-u)K_h(t-w)Y_i(u \vee w)Z_{i1}(u)Z_{i2}(w)dO_{i1}(u)dO_{i2}(w),$$

and

$$\widehat{S}^{(0)}_{12,h}(t) = n^{-1}\sum_{i=1}^{n}\int_0^{\tau}\int_0^{\tau}K_h(t-u)K_h(t-w)Y_i(u \vee w)dO_{i1}(u)dO_{i2}(w)$$

We note that $\widehat{S}^{(2)}_{12,h}(t)$ consistently estimates $E\{Y(t)Z_1(t)Z_2(t)\}m_{12}(t, t)$ and $\widehat{S}^{(0)}_{12,h}(t)$ consistently estimates $E\{Y(t)\}m_{12}(t, t)$. Thus the off-diagonal entries can be replaced by $\widehat{S}^{(2)}_{12,h}(t) / \widehat{S}^{(0)}_{12,h}(t)$, which consistently estimates the population level quantity $E\{Y(t)Z_1(t)Z_2(t)\}/E\{Y(t)\}$. To sum up, define

$$\widetilde{S}^{(2)}(t) / \widetilde{S}^{(0)}(t) = \begin{pmatrix} \widehat{S}^{(2)}_{1,h}(t) / \widehat{S}^{(0)}_{1,h}(t) & \widehat{S}^{(2)}_{12,h}(t) / \widehat{S}^{(0)}_{12,h}(t) \\ \widehat{S}^{(2)}_{12,h}(t) / \widehat{S}^{(0)}_{12,h}(t) & \widehat{S}^{(2)}_{2,h}(t) / \widehat{S}^{(0)}_{2,h}(t) \end{pmatrix}$$

then $\boldsymbol{\beta}$ can be consistently estimated by

$$\widehat{\boldsymbol{\beta}}_h = \left(\int_0^{\tau}\left[n^{-1}\sum_{i=1}^{n}Y_i(t)\frac{\widetilde{S}^{(2)}(t)}{\widetilde{S}^{(0)}(t)} - n^{-1}\sum_{i=1}^{n}Y_i(t)\left\{\frac{\widetilde{S}^{(1)}(t)}{\widetilde{S}^{(0)}(t)}\right\}^{\otimes 2}\right]dt\right)^{-1} \times \left[n^{-1}\sum_{i=1}^{n}\int_0^{\tau}\mathbf{Z}_i(t)dN_i(t) - \int_0^{\tau}\frac{\widetilde{S}^{(1)}(t)}{\widetilde{S}^{(0)}(t)}\left\{n^{-1}\sum_{i=1}^{n}dN_i(t)\right\}\right]$$

4 Simulation

We conducted simulation studies to evaluate the performance of the proposed method. Under each simulation scenario, we generated 1000 data replicates with sample size 300 and 600. The resampling size was set to be 100 in the bootstrap method for variance estimation. The recurrent events were generated based on the following additive intensity model where the intensity of the recurrent event process for subject i is

$$\lambda\{t \mid Z_i(t), \gamma_i\} = \lambda_0(t) + \beta Z_i(t) + \gamma_i \tag{8}$$

The frailty variable γ_i was generated from a gamma distribution with mean 0.02 and variance 0.004. The baseline intensity function $\lambda_0(t) = 0.1I(t \le 10) + 0.3I(10 < t \le 20)$. Note that the intensity model in equation (8) implies the additive rates model $\lambda\{t \mid Z_i(t)\} = \lambda_0^*(t) + \beta Z_i(t)$, where the baseline rate function $\lambda_0^*(t) = 0.02 + \lambda_0(t)$.

In the first set of simulations, we considered a continuous time-dependent covariate defined by $Z_i(t) = b_{0i} + b_{1i}t$, where the random intercept b_{0i} was generated from a normal distribution with mean 1.5 and variance 0.05. The random slope b_{1i} was generated from either a zero mean or a non-zero mean (−0.05) normal distribution with variance 5×10^{-4}. The two cases are referred to as *without time trend* and *with time trend*, respectively. The regression coefficient β is set at 0.2.

In the second set of simulations, we considered a binary time-dependent covariate. First, we generated the baseline value, $Z_i(0)$ from a Bernoulli distribution with probability 0.2. Then the binary covariate process was generated from a multistate process which consists of two states, 0 and 1. The duration of state 0 of subject i was generated from an exponential distribution with rate function $1/\{\xi_i g(t)\}$, and the duration of state 1 was generated from an exponential distribution with rate $1/\xi_i$, where the subject-specific random effect ξ_i followed a gamma distribution with mean 1 and variance 0.25 and the function $g(t)$ was set such that the covariate was either *with a time trend*: $g(t) = 4\mathrm{I}(t \le 10) + 6\mathrm{I}(10 < t \le 20)$ or *without a time trend*: $g(t) = 4$ for $t \in [0, 20]$). The regression coefficient is set at $\beta = 0.5$.

In all settings, we let the covariates of a subject be observed at its own event times and each subject has a baseline visit at time 0. For each subject, the time of 20 follow-up visits (if there is no censoring) was generated based on a uniform distribution within each of 20 unit time intervals, (0, 1], (1, 2],…, (19, 20]. We allowed each visit to have a certain probability to be missing, p_m = 0%, 20%, 40%, and 60%. The censoring time was simulated from a uniform distribution on the interval [0, 20].

We applied the proposed method and two simple approaches, the LCCF method and the linear interpolation method, to the simulated data. For the proposed method, we used the Epanechnikov kernel function and a bandwidth selection procedure as follows. First, we define the averaged squared error as $ASE(h) = n^{-1}\sum_{i=1}^{n}\int_0^{\tau} Y_i(u)\left\{\hat{\xi}_h^{(1)}(u) - \xi^{(1)}(u)\right\}^2 dO_i(u)$. Since $ASE(h)$ involves the unknown quantity $\xi^{(1)}(\cdot)$, we define $CV(h)$ with the leave-one-out estimator as $CV(h) = n^{-1}\sum_{i=1}^{n}\int_0^{\tau} Y_i(u)\left\{Z_i(u) - \hat{\xi}_{h,-i}^{(1)}(u)\right\}^2 dO_i(u)$. It is easy to show that

$$CV(h) = n^{-1}\sum_{i=1}^{n}\int_0^{\tau} Y_i(u)\left\{Z_i(u) - \xi^{(1)}(u)\right\}^2 dO_i(u) + n^{-1}\sum_{i=1}^{n}\int_0^{\tau} Y_i(u)\left\{\hat{\xi}_{h,-i}^{(1)}(u) - \xi^{(1)}(u)\right\}^2 dO_i(u) - 2n^{-1}\sum_{i=1}^{n}\int_0^{\tau} Y_i(u)\left\{Z_i(u) - \xi^{(1)}(u)\right\}\left\{\hat{\xi}_{h,-i}^{(1)}(u) - \xi^{(1)}(u)\right\} dO_i(u)$$

Since the first item on the right-hand side does not involve h and the expectation of the third item is zero, minimizing $ASE(h)$ is on average equivalent to minimizing $CV(h)$. Using similar techniques as those in Chiang et al.,[26] it can be shown that the ASE converges to $O(h^4) + O(1/(nh))$, where the first term corresponds to squared bias and the second corresponds to variance. Thus, we can show that the optimal nonparametric convergence rate is $Cn^{-1/5}$ by following the same argument as in Newey et al.[27] We then determine the constant C by minimizing $CV(h)$ with $h = Cn^{-1/5}$. In Appendix 1 we show that the range of the bandwidth for $\hat{\boldsymbol{\beta}}_h$ is $h = O(n^{-v})$, where $1/4 < v < 1/2$, so after choosing the constant C in the first step, we use $h = Cn^{-1/3}$ for the estimation of $\boldsymbol{\beta}$.

In the simulation tables, we report the relative bias (Bias) and the Monte-Carlo empirical standard deviation of the point estimates (SD). For the proposed method, we also report the average standard errors (ASE) estimated by the bootstrap method and the coverage percentage (CP) of the 95% confidence intervals. Table 1 shows the simulation results when Z_i is continuous. For the scenarios with no time trend in the covariate, the LCCF method gives biased point estimates and the bias increases as the missing probability increases. The linear interpolation method and the proposed method give virtually unbiased point estimates. For the scenarios with time trend in the covariate, both the LCCF and linear interpolation methods give biased estimates, while the proposed method still provides virtually unbiased estimates. For the variance estimation of the proposed method, the ASEs are all close to the Monte Carlo SDs, and the coverage percentages are all close to 95%. As expected, the SDs (and ASEs) of the proposed method decrease as the sample size increases and increase as the missing rate increases. Table 2 shows the results when the time-dependent covariate is binary. The two simple methods provide biased estimations regardless of whether there is time trend or not in the covariate. The proposed estimator gives virtually unbiased estimates for all scenarios.

We examined the performance of the baseline mean function estimation using the first scenario in Tables 1 and 2. Figure 1 shows that the bias of the baseline mean function estimates from their corresponding true baseline function was negligible under the simulation scenarios.

We also conducted simulation studies to evaluate the performance of the two extensions of the proposed method described in Sections 3.1 and 3.2, namely (1) when both time-dependent and -independent covariates are present in the model and (2) when time-dependent covariates are measured on different time schedules. For the first extension, we simulated data with one continuous time-independent covariate W_i from a normal distribution with mean 1.5 and variance 0.05 and one binary time-dependent covariate $Z_i(t)$, in the same way as for the binary covariate with time trend described before. The simulation results (the top panel of Table 3) show that the proposed method described in Section 3.1 works well under various scenarios. Additional simulations were performed to compare the method proposed in Section 3.1 and the estimator without discriminating time-dependent and -independent covariates in equation (6). The results in Table 4 show that the former method was more efficient than the latter, especially in the estimation of the time-independent covariate's effect for our simulated data.

For the second extension, we simulated data with two time-dependent covariates, one binary and one continuous, following the same way as before, except that the measuring times of the two covariates were simulated separately. We explored situations where each covariate was either with or without time trend. The simulation results of the two extensions are presented in Table 3. It is shown that the extensions of the proposed method perform well under all scenarios.

5 Real data analysis

We applied the proposed method to a study investigating the effect of streptococci on the risk of pharyngitis.[4] Pharyngitis is an infection of the pharynx, the back of the throat, which is often due to viruses, but several bacteria which include group A streptococcus (GAS) are also a common cause of pharyngitis. The pharyngitis caused by GAS is also known as strep throat and is prevalent in children and usually occurs in late winter and early spring. Bacteria of other streptococcal groups including GCS and GGS may also cause pharyngitis, and thus it is of clinical interest to investigate the effect of these bacteria on the risk of pharyngitis. Between March 2002 and March 2004, 307 school children in a rural area near Vellore, India were recruited. During the follow-up time, cases of pharyngitis were identified weekly (referred to as 'event visits') and the streptococci status was also determined for those with pharyngitis at the time when pharyngitis was diagnosed. In addition, monthly visits were scheduled to monitor the streptococci status (referred to as 'regular visits'). The detailed design of this study can be found in Jose et al.[4] Note that although regular visits were scheduled on a monthly basis, the actual observation times were irregularly spaced across subjects to balance the workload. It is reasonable to assume that the regular observation process $O_i(t)$ is independent of the covariate processes $\mathbf{Z}_i(t)$ and the censoring time C_i.

The start time of the study, March 11, 2002, is used as the time origin of the recurrent event process of the occurrence of pharyngitis. By choosing calendar time as the time scale, we can avoid modeling the confounding effect of season which is a nuisance in this study. Note that 74 (out of 307) school children were recruited in the second year after 15 June 2003, for whom, the at-risk indicator $Y_i(t)$ is modified to reflect whether subject i has been enrolled in the study prior to time t and remained under observation at time t. A two-week rule was applied to determine an episode of pharyngitis, i.e. a pharyngitis event occurred within 14 days after a previous episode was considered as the same episode. During the two-year follow-up, 640 pharyngitis occurrences were identified and 2827 regular visits were recorded. Among throat cultures collected in the regular visits, about 11.43% of them were positive for GAS, 2.90% were positive for GCS and 15.32% were positive for GGS. Among the cultures collected at the event visits, about 17.19% of them were GAS positive, 4.69% GCS positive and 17.66% GGS positive. Since GAS, GCS and GGS all belong to the Streptococcus genus family, they are likely to be correlated. We applied McNemar's test for pairwise comparison using the measurements in the first regular visit of each child to test if these bacterial infections were correlated with each other. The results show that GCS was significantly correlated with both GAS and GGS but no significant correlation was observed between GAS and GGS. Thus to avoid collinearity, we fit the additive rates model with only GAS and GGS to explore their relationship with the occurrence of pharyngitis. The bandwidth parameter in the proposed estimator was selected to be 0.6 using the approach described in Section 4. The estimated rate difference for the time-dependent GAS and GGS status based on the proposed kernel method are 0.067 and 0.020, respectively, and their corresponding 95% confidence intervals are (0.028, 0.106) and (−0.013, 0.053). Thus, we conclude that positive GAS was associated with a higher risk of pharyngitis, while the GGS infection status was not significantly associated with the risk of pharyngitis. Figure 2 shows the estimated baseline mean function with point-wise 95% confidence bands.

Lastly, we extend the supremum test proposed in Lin et al.[28] to check the additive rates assumption, where the p-value is approximated by the empirical probability that the supremum statistics based on score processes simulated under the additive rates model is greater than the supremum statistic based on the observed score process. The supremum test yielded a p-value of 0.41, suggesting that the additivity assumption is reasonably met. Note that the Lin et al. approach[28] requires the values of covariates to be observed throughout the follow-up period. For ease of implementation, we used the last-covariate-carried-forward approach in the above model checking procedure. Details of the testing procedure can be found in the supplemental material.

6 Discussion

In this paper, we propose a kernel smoothed estimating function method to deal with intermittently measured time-dependent covariates in the additive rates model. Compared to the Cox-type models, the additive model is more appealing to practitioners when the rate difference is of primary interest or the proportional rates assumption is violated. In relation to the recent works on Cox-type model,[23,24] the proposed work offers an alternative tool to analyze recurrent event data with time-dependent covariates which are only intermittently observed. Moreover, when multiple time-dependent covariates are in presence and measured on different schedules, the methods in Cao et al.[23] and Li et al.[24] cannot be directly applied. In this case, we extend our proposed method by using multivariate kernels to obtain consistent estimates.

Even though joint models would be a useful alternative method for analyzing recurrent event data with time-dependent covariates, there has been no joint models with the additive rates submodel for the recurrent events existing in literature. Even if such a joint model exists, the proposed model would still be preferable since it does not require modeling the underlying covariate process, while the joint model would require a complete specification of the joint distribution of the recurrent event process and the covariate process. This could be challenging when both continuous and binary covariates are present. Instead, we apply nonparametric kernel smoothing method to approximate the mean covariate process to obtain consistent estimates, and hence is more robust against model misspecifications.

In the motivating example, the covariates were measured at both event visits and regular visits, which is typical in recurrent event data since the subjects are still at risk after an event occurs. If the covariates are not observed at the time of events, a double kernel approach similar to what was proposed for the proportional rates model in Cao et al.[23] can be extended to the additive rates model, but the convergence rate of the resulting estimator would be slower than the regular root-n rate. A less computationally intensive and simpler method is to carry forward the last observed value to replace the missing observation at event times in $\mathbf{Z}_i(t)dN_i(t)$ in equation (5) and keep the rest of the terms in the estimating function which involve kernel smoothing the same. The performance of these two approaches will be evaluated in future research.

As another future direction, we can apply the kernel smoothing method to deal with intermittently measured covariates in additive-multiplicative rates model. It is also of interest to investigate model checking procedures to determine whether a covariate has an additive or multiplicative effect. For example, Lin et al.[8] have proposed a standardized score-type process to check the multiplicative assumption for recurrent event data with continuously monitored time-dependent covariates. Research on checking the additive or multiplicative assumption for intermittently observed time-dependent covariates is warranted.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: Lyu and Luo were partially supported by the U.S. National Institutes of Health (R03MH112895). Luo was also supported by the U.S. National Institutes of Health (P30CA077598). Huang was supported by the U.S. National Institutes of Health (R01CA193888).

Appendix 1

1.1 Proof of consistency in Theorem 1

Similar to the proofs in Li et al.,[24] we impose the following assumptions:

1. $\{N_i(\cdot), O_i(\cdot), Y_i(\cdot), \mathbf{Z}_i(\cdot)\}$, $i = 1, \ldots, n$, are independent and identically distributed.
2. $N_i(\tau)$ is bounded. Define $\lambda^c(\cdot)$ as the rate function of $N_i(\cdot)$ and $\lambda^c(\cdot)$ is of bounded variation.
3. The true parameter $\boldsymbol{\beta}_0$ is in a compact set $\mathscr{B}$ in $\mathscr{R}^p$ and the baseline rate function $\lambda_0(t)$ is absolutely continuous.
4. For each element in the covariates $\mathbf{Z}_i(t)$, the covariate process $Z_{ij}(t)$ has uniformly bounded total variation, namely $\int_0^\tau |dZ_{ij}(t)| + |Z_{ij}(0)| \le c$ for some $c > 0$ for all i and j. Without loss of generality, we assume $Z_{ij}(t) \ge 0$.
5. The censoring time C_i is independent of $N_i^*(\cdot)$ conditional on $\mathbf{Z}_i(\cdot)$ with $G(\tau) = P(C_i \ge \tau) > 0$.
6. The function $s^{(k)}(t) = \mathrm{E}\{Y_i(t)\mathbf{Z}_i(t)^{\otimes k}\}$, $k = 0, 1, 2$ has bounded second derivatives for $t \in [0, \tau]$.
7. The observation time process $O_i(t)$ is independent of $\{N_i^*(\cdot), Y_i(\cdot), \mathbf{Z}_i(\cdot)\}$ and is bounded. Moreover, the covariate collection rate function $m(t)$, defined by $m(t)dt = \mathrm{E}\{dO_i(t)\}$, is positive and has bounded second derivative for $t \in [0, \tau]$.
8. The kernel function $K(\cdot)$ is a symmetric density function with bounded support which satisfies: $\int_{-1}^1 K(t)dt = 1$, $\int_{-1}^1 tK(t)dt = 0$ and $\int_{-1}^1 t^2K(t)dt$ is a positive constant.
9. The bandwidth $h = O(n^{-v})$, where $1/4 < v < 1/2$.

Define $\Psi(u) = \mathrm{E}\{Y(u)\mathbf{Z}(u)^{\otimes k}\}m(u)$, then $s^{(k)}(t)m(t) = \Psi(t)$. The expectation of the kernel smoothed processes $\mathrm{E}\{\hat{S}_h^{(k)}(t)\} = \int_0^\tau K_h(t-u)\mathrm{E}\{Y(u)\mathbf{Z}(u)^{\otimes k}\}m(u)du$, then we have

$$\begin{aligned}\mathrm{E}\{\hat{S}_h^{(k)}(t)\} &= \int_0^\tau K_h(t-u)\boldsymbol{\Psi}(u)du = \int_{(t-\tau)/h}^{t/h} K(\bar{u})\boldsymbol{\Psi}(t-h\bar{u})d\bar{u}\\ &= \boldsymbol{\Psi}(t)\int_{(t-\tau)/h}^{t/h} K(\bar{u})d\bar{u} - \int_{(t-\tau)/h}^{t/h} h\bar{u}K(\bar{u})\boldsymbol{\Psi}'(t)d\bar{u} + \int_{(t-\tau)/h}^{t/h} h^2\bar{u}^2K(\bar{u})d\bar{u}\boldsymbol{\Psi}''(t^*)\end{aligned}$$

It is easy to see that $\sup_{t\in[h,\tau-h]}|\mathrm{E}\{\hat{S}_h^{(k)}(t)\} - s^{(k)}(t)m(t)| = O(h^2)$ under the assumption 8. Also, it is straightforward to show that $\sup_{t\in[0,h)}|s^{(k)}(t)m(t) - s^{(k)}(h)m(h)|$ and $\sup_{t\in(\tau-h,\tau]}|s^{(k)}(t)m(t) - s^{(k)}(\tau-h)m(\tau-h)| = O(h)$.

Next, we show the convergence of $\hat{S}_h^{(k)}(t) - \mathrm{E}\{\hat{S}_h^{(k)}(t)\}$. We define $\hat{R}^{(k)}(t) = n^{-1}\sum_{i=1}^n\int_0^t Y_i(u)\,\mathbf{Z}_i(u)^{\otimes k}dO_i(u)$ and $r^{(k)}(t) = \mathrm{E}\{\int_0^t Y(u)\mathbf{Z}(u)^{\otimes k}dO(u)\}$, so $\hat{S}_h^{(k)}(t) = \int_0^\tau K_h(t-u)d\hat{R}^{(k)}(u)$ and $\mathrm{E}\{\hat{S}_h^{(k)}(t)\} = \int_0^\tau K_h(t-u)dr^{(k)}(u)$. Then we have

$$\sup_{t\in[h,\tau-h]}|\hat{S}_h^{(k)}(t) - \mathbf{E}\{\hat{S}_h^{(k)}(t)\}| \le h^{-1}\sup_{t\in[0,\tau]}|\hat{R}^{(k)}(t) - r^{(k)}(t)|\,V(K) \tag{9}$$

where $V(K)$ is the variation of the kernel function. Also, since the function classes $\mathscr{F}_k = \{\int_0^t Y(u)\mathbf{Z}(u)^{\otimes k}dO(u) : t\in[0,\tau]\}$ are monotone, by Theorem 2.14.9 in Van Der Vaart and Wellner,[29] $P\left(\sup_{t\in[0,\tau]}\sqrt{n}\,|\hat{R}^{(k)}(t) - r^{(k)}(t)| > x\right) \le c_k x^{v_k} e^{-b_k x^2}$, where c_k, v_k, b_k are constants. Therefore, for any ϵ, we have

$$\begin{aligned}&P\left(\sup_{t\in[0,\tau]}h^{-1}|\hat{R}^{(k)}(t) - r^{(k)}(t)| > \epsilon\right)\\ &= P\left(\sup_{t\in[0,\tau]}\sqrt{n}\,|\hat{R}^{(k)}(t) - r^{(k)}(t)| > \sqrt{n}h\epsilon\right)\\ &\le c_k(\sqrt{n}h\epsilon)^{v_k}e^{-b_k(\sqrt{n}h\epsilon)^2}.\end{aligned} \tag{10}$$

It follows from equations (9) and (10) that $\sup_{t\in[h,\tau-h]}|\hat{S}_h^{(k)}(t) - \mathrm{E}\{\hat{S}_h^{(k)}(t)\}|$ converges to 0 when $nh^2 \to \infty$. Previously we have shown that $\sup_{t\in[h,\tau-h]}|\mathrm{E}\{\hat{S}_h^{(k)}(t)\} - s^{(k)}(t)m(t)| = O(h^2)$, so the consistency of $\hat{S}_h^{(k)}(t)$ has been proved. By the law of large numbers, we know that $n^{-1}\sum_{i=1}^n N_i(t)$ converges to $\mathrm{E}\{N_i(t)\}$ and $n^{-1}\sum_{i=1}^n\int_0^\tau \mathbf{Z}_i(t)dN_i(t)$ converges to $\int_0^\tau \mathrm{E}\{\mathbf{Z}_i(t)dN_i(t)\}$. Thus, we show that $\hat{\boldsymbol{\beta}}$ converges in probability to $\boldsymbol{\beta}_0$.

1.2 Proof of asymptotic normality in Theorem 1

To establish asymptotic normality of $\sqrt{n}\widehat{\boldsymbol{\beta}}$, we first obtain the asymptotic i.i.d. representation of $\widehat{U}_h(\boldsymbol{\beta}_0)$. By equation (5), we have

$$\begin{aligned}\sqrt{n}\widehat{U}_h(\boldsymbol{\beta}_0) &= n^{-1/2}\sum_{i=1}^{n}\int_0^{\tau}\mathbf{Z}_i(t)dN_i(t) - n^{-1/2}\sum_{i=1}^{n}\int_0^{\tau}\hat{\xi}_h^{(1)}(t)dN_i(t)\\ &\quad - n^{-1/2}\sum_{i=1}^{n}\int_0^{\tau}Y_i(t)\hat{\xi}_h^{(2)}(t)dt\boldsymbol{\beta}_0 + n^{-1/2}\sum_{i=1}^{n}\int_0^{\tau}Y_i(t)\hat{\xi}_h^{(1)}(t)^{\otimes 2}dt\boldsymbol{\beta}_0\\ &\overset{def}{=} I_1 + I_2 + I_3 + I_4\end{aligned}$$

It can be shown that the second term can be expressed as

$$\begin{aligned}I_2 &= -n^{-1/2}\sum_{i=1}^{n}\int_0^{\tau}\hat{\xi}_h^{(1)}(t)dN_i(t)\\ &= -\sqrt{n}\int_0^{\tau}\hat{\xi}_h^{(1)}(t)d\left\{n^{-1}\sum_{i=1}^{n}N_i(t) - \mathrm{E}\{N_i(t)\}\right\} - \sqrt{n}\int_0^{\tau}\hat{\xi}_h^{(1)}(t)d\mathrm{E}\{N_i(t)\}\\ &= -\sqrt{n}\int_0^{\tau}\frac{s^{(1)}(t)}{s^{(0)}(t)}d\left\{n^{-1}\sum_{i=1}^{n}N_i(t) - \mathrm{E}\{N_i(t)\}\right\} - \sqrt{n}\int_0^{\tau}\hat{\xi}_h^{(1)}(t)d\mathrm{E}\{N_i(t)\} + o_p(1)\end{aligned}$$

Since $\lambda^c(t)$ is the rate function of $N_i(t)$, we have $\lambda^c(t)dt = d\mathrm{E}\{N_i(t)\}$. Moreover, it follows from

$$\begin{aligned}&\sqrt{n}\int_0^{\tau}\hat{\xi}_h^{(1)}(t)\lambda^c(t)dt - \sqrt{n}\int_0^{\tau}\frac{s^{(1)}(t)}{s^{(0)}(t)}\lambda^c(t)dt\\ &= n^{-1/2}\sum_{i=1}^{n}\left\{\int_0^{\tau}\frac{\lambda^c(t)}{s^{(0)}(t)m(t)}Y_i(t)\mathbf{Z}_i(t)dO_i(t) - \int_0^{\tau}\frac{s^{(1)}(t)\lambda^c(t)}{s^{(0)}(t)^2m(t)}Y_i(t)dO_i(t)\right\} + o_p(1)\end{aligned}$$

that $I_2 = n^{-1/2}\sum_{i=1}^{n}\phi_{2i} + o_p(1)$, with

$$\phi_{2i} = -\int_0^{\tau}\frac{s^{(1)}(t)}{s^{(0)}(t)}dN_i(t) - \int_0^{\tau}\frac{\lambda^c(t)}{s^{(0)}(t)m(t)}Y_i(t)\mathbf{Z}_i(t)dO_i(t) + \int_0^{\tau}\frac{s^{(1)}(t)\lambda^c(t)}{s^{(0)}(t)^2m(t)}Y_i(t)dO_i(t)$$

Next, we show that

$$
\begin{aligned}
I_3 &= -\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\int_0^{\tau} Y_i(t)\hat{\xi}_h^{(2)}(t)dt\boldsymbol{\beta}_0 \\
&= -\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\int_0^{\tau} Y_i(t)\frac{s^{(2)}(t)}{s^{(0)}(t)}dt\boldsymbol{\beta}_0 + \sqrt{n}\int_0^{\tau} s^{(0)}(t)\frac{s^{(2)}(t)}{s^{(0)}(t)}dt\boldsymbol{\beta}_0 - \sqrt{n}\int_0^{\tau}\hat{\xi}_h^{(2)}(t)s^{(0)}(t)dt\boldsymbol{\beta}_0 + o_p(1) \\
&= -\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\left\{\int_0^{\tau} Y_i(t)\frac{s^{(2)}(t)}{s^{(0)}(t)}dt - \int_0^{\tau}\frac{1}{m(t)}Y_i(t)\mathbf{Z}_i(t)^{\otimes 2}dO_i(t) + \int_0^{\tau}\frac{s^{(2)}(t)}{s^{(0)}(t)m(t)}Y_i(t)dO_i(t)\right\}\boldsymbol{\beta}_0 + o_p(1) \\
&\overset{def}{=} \frac{1}{\sqrt{n}}\sum_{i=1}^{n}\phi_{3i}(\boldsymbol{\beta}_0) + o_p(1)
\end{aligned}
$$

and

$$
\begin{aligned}
I_4 &= \frac{1}{\sqrt{n}}\sum_{i=1}^{n}\int_0^{\tau} Y_i(t)\hat{\xi}_h^{(1)}(t)^{\otimes 2}dt\boldsymbol{\beta}_0 \\
&= \frac{1}{\sqrt{n}}\sum_{i=1}^{n}\int_0^{\tau} Y_i(t)\frac{s^{(1)}(t)^{\otimes 2}}{s^{(0)}(t)^2}dt\boldsymbol{\beta}_0 - \sqrt{n}\int_0^{\tau} s^{(0)}(t)\frac{s^{(1)}(t)^{\otimes 2}}{s^{(0)}(t)^2}dt\boldsymbol{\beta}_0 + \sqrt{n}\int_0^{\tau} s^{(0)}(t)\hat{\xi}_h^{(1)}(t)^{\otimes 2}dt\boldsymbol{\beta}_0 + o_p(1) \\
&= \frac{1}{\sqrt{n}}\sum_{i=1}^{n}\left\{\int_0^{\tau} Y_i(t)\frac{s^{(1)}(t)^{\otimes 2}}{s^{(0)}(t)^2}dt + \int_0^{\tau}\frac{2s^{(1)}(t)}{s^{(0)}(t)m(t)}Y_i(t)\mathbf{Z}_i(t)^{\top}dO_i(t) - \int_0^{\tau}\frac{2s^{(1)}(t)^{\otimes 2}}{s^{(0)}(t)^2 m(t)}Y_i(t)dO_i(t)\right\}\boldsymbol{\beta}_0 + o_p(1) \\
&\overset{def}{=} \frac{1}{\sqrt{n}}\sum_{i=1}^{n}\phi_{4i}(\boldsymbol{\beta}_0) + o_p(1
\end{aligned}
$$

)

Thus, we have $\sqrt{n}\hat{U}_h(\boldsymbol{\beta}_0) = n^{-1/2}\sum_{i=1}^{n}\phi_i(\boldsymbol{\beta}_0) + o_p(1)$, where $\phi_i(\boldsymbol{\beta}_0) = \int_0^{\tau}\mathbf{Z}_i(t)dN_i(t) + \phi_{2i}(\boldsymbol{\beta}_0) + \phi_{3i}(\boldsymbol{\beta}_0) + \phi_{4i}(\boldsymbol{\beta}_0)$. Therefore, $\sqrt{n}(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta}_0)$ converges in distribution to a normal random variable with mean zero and variance $\Sigma = A(\boldsymbol{\beta}_0)^{-1}\ V(\boldsymbol{\beta}_0)\{A(\boldsymbol{\beta}_0)^{-1}\}^{\top}$, where $A(\boldsymbol{\beta}_0) = \int_0^{\tau}s^{(0)}(t)\left[s^{(2)}(t)/s^{(0)}(t) - \{s^{(1)}(t)/s^{(0)}(t)\}^{\otimes 2}\right]dt$ and $V(\boldsymbol{\beta}_0) = E\{\phi_1(\boldsymbol{\beta}_0)\phi_1(\boldsymbol{\beta}_0)^{\top}\}$.

References

1. Barker JN, Hough RE, van Burik JAH, et al. Serious infections after unrelated donor transplantation in 136 children: impact of stem cell source. Biol Blood Marrow Transplant 2005; 11: 362–370. [PubMed: 15846290]
2. Kohli P, Wallentin L, Reyes E, et al. Reduction in first and recurrent cardiovascular events with ticagrelor compared with clopidogrel in the PLATO Study. Circulation 2013; 127: 673–680. [PubMed: 23277305]
3. Eaton WW, Mortensen PB, Herrman H, et al. Long-term course of hospitalization for schizophrenia: Part I. Risk for rehospitalization. Schizophr Bull 1992; 18: 217–228. [PubMed: 1621069]
4. Jose JJM, Brahmadathan KN, Abraham VJ, et al. Streptococcal group A, C, and G pharyngitis in school children: a prospective cohort study in Southern India. Epidemiol Infect 2018; 146: 848–853. [PubMed: 29616606]
5. Andersen PK and Gill RD. Cox's regression model for counting processes: a large sample study. Ann Stat 1982; 10: 1100–1120.
6. Prentice RL, Williams BJ and Peterson AV. On the regression analysis of multivariate failure time data. Biometrika 1981; 68: 373–379.

7. Pepe MS and Cai J. Some graphical displays and marginal regression analyses for recurrent failure times and time dependent covariates. J Am Stat Assoc 1993; 88: 811–820.

8. Lin DY, Wei LJ, Yang I, et al. Semiparametric regression for the mean and rate functions of recurrent events. J Royal Stat Soc Ser B 2000;62:711–730.

9. Schaubel DE, Zeng D and Cai J. A semiparametric additive rates model for recurrent event data. Lifetime Data Anal 2006; 12: 389–406. [PubMed: 17031499]

10. Liu YY and Wu YS. Semiparametric additive intensity model with frailty for recurrent events. Acta Math Sin Engl Ser 2011; 27: 1831–1842.

11. Andersen PK and Liestøl K. Attenuation caused by infrequently updated covariates in survival analysis. Biostat 2003; 4: 633–649.

12. Raboud J, Reid N, Coates RA, et al. Estimating risks of progressing to AIDS when covariates are measured with error. J Royal Stat Soc Ser A 1993; 156: 393–406.

13. Dafni UG and Tsiatis AA. Evaluating surrogate markers of clinical outcome when measured with error. Biometrics 1998; 54: 1445–1462. [PubMed: 9883544]

14. Boscardin WJ, Taylor JMG and Law N. Longitudinal models for AIDS marker data. Stat Methods Med Res 1998; 7: 13–27. [PubMed: 9533259]

15. Bycott P and Taylor J. A comparison of smoothing techniques for CD4 data measured with error in a time-dependent Cox proportional hazards model. Stat Med 1998; 17: 2061–2077. [PubMed: 9789914]

16. Tsiatis AA and Davidian M. Joint modeling of longitudinal and time-to-event data: an overview. Stat Sin 2004; 14: 809–834.

17. Rizopoulos D Joint models for longitudinal and time-to-event data: with applications in R. Boca Raton, FL: CRC Press, Taylor & Francis Group, LLC., 2012.

18. Henderson R, Diggle P and Dobson A. Joint modelling of longitudinal measurements and event time data. Biostat 2000; 1: 465–480.

19. Li S Joint modeling of recurrent event processes and intermittently observed time-varying binary covariate processes. Lifetime Data Anal 2016; 22: 145–160. [PubMed: 25573223]

20. Liu L and Huang X. Joint analysis of correlated repeated measures and recurrent events processes in the presence of death, with application to a study on acquired immune deficiency syndrome. J Royal Stat Soc Ser C 2009; 58: 65–81.

21. Kim S, Zeng D, Chambless L, et al. Joint models of longitudinal data and recurrent events with informative terminal event. Stat Biosci 2012; 4: 262–281. [PubMed: 23227131]

22. Cai Q, Wang M-C and Chan KCG. Joint modeling of longitudinal, recurrent events and failure time data for survivor’s population. Biometrics 2017; 73: 1150–1160. [PubMed: 28334426]

23. Cao H, Churpek MM, Zeng D, et al. Analysis of the proportional hazards model with sparse longitudinal covariates. J Am Stat Assoc 2015; 110: 1187–1196. [PubMed: 26576066]

24. Li S, Sun Y, Huang C-Y, et al. Recurrent event data analysis with intermittently observed time-varying covariates. Stat Med 2016; 35: 3049–3065. [PubMed: 26887664]

25. Lin DY and Ying Z. Semiparametric analysis of the additive risk model. Biometrika 1994; 81: 61–71.

26. Chiang C-T, Wang M-C and Huang C-Y. Kernel estimation of rate function for recurrent event data. Scand J Stat 2005; 32: 77–91.

27. Newey WK, Hsieh F and Robins JM. Twicing kernels and a small bias property of semiparametric estimators. Econometrica 2004; 72: 947–962.

28. Lin DY, Wei LJ and Ying Z. Checking the Cox model with cumulative sums of martingale-based residuals. Biometrika 1993; 80: 557–572.

29. Van Der Vaart AW and Wellner JA. Weak convergence and empirical processes. New York, NY: Springer-Verlag; 1996.

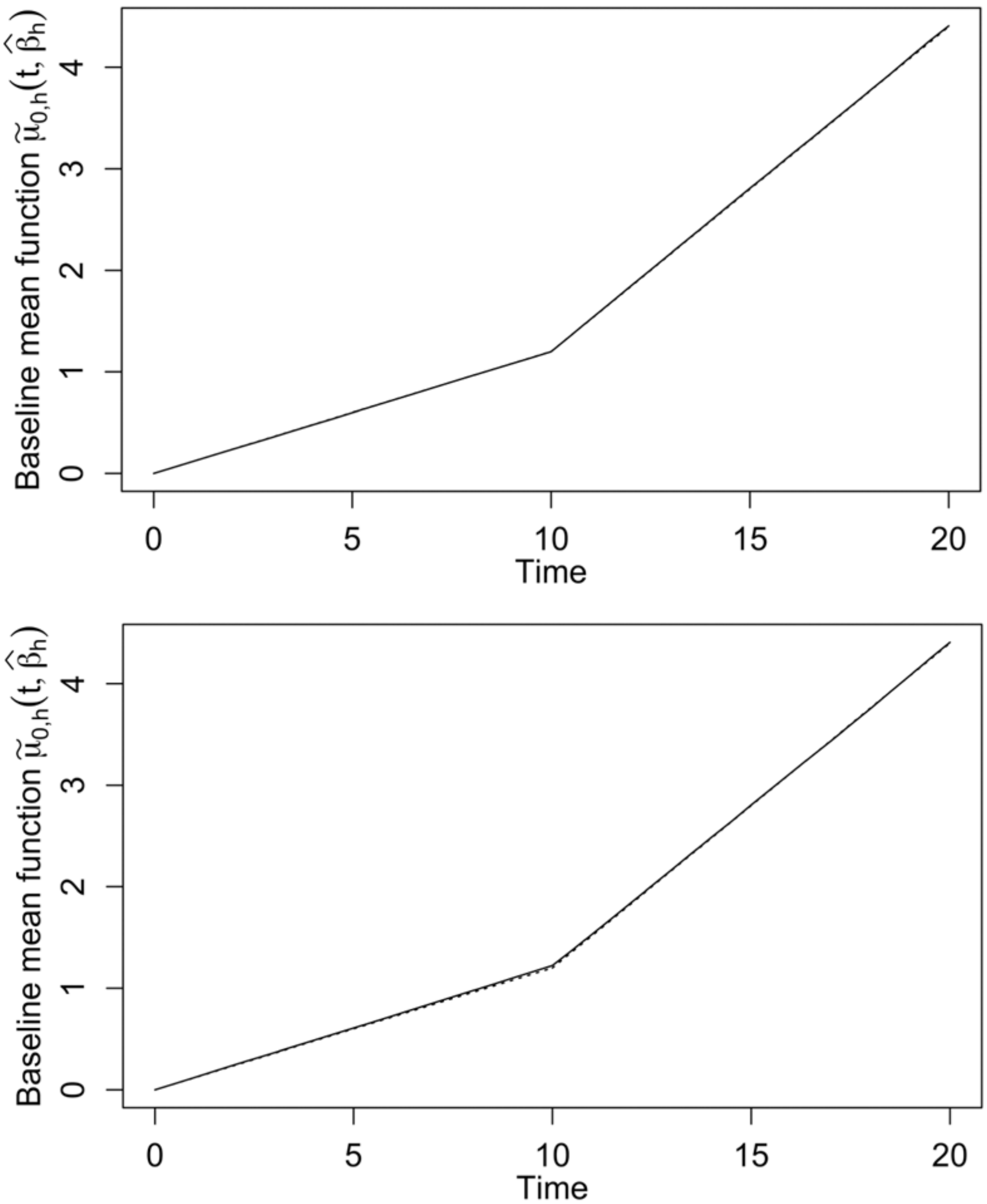

Figure 1.

Estimation of the baseline mean function for simulated data: (a) continuous time-dependent covariate scenario, (b) binary time-dependent covariate scenario. The solid line is the mean of 1000 estimated baseline mean functions and the dotted line is the true baseline mean function.

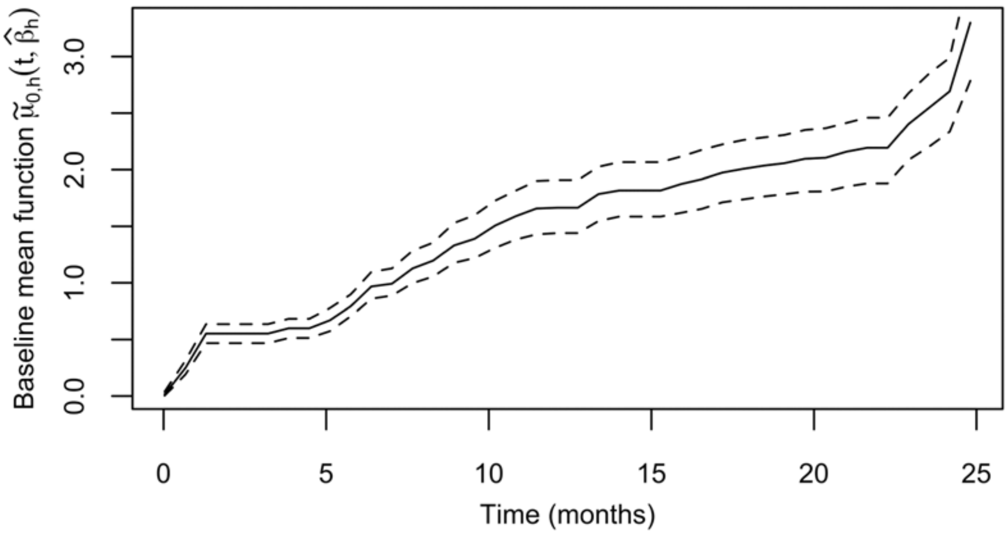

Figure 2.
Estimation of the baseline mean function for Indian pharyngitis data. Time 0 is the start time of the study, 11 March 2002. The dashed lines are the 95% point-wise confidence bands based on the bootstrap samples.

Table 1.

Simulation results for the model with a continuous time-dependent covariate using the last covariate carried forward (LCCF) method, linear interpolation method (Linear), and the proposed kernel smoothing method (Proposed).

p_m	LCCF		Linear		Proposed			
	Bias	**SD**	**Bias**	**SD**	**Bias**	**SD**	**ASE**	**CP**
$n = 300$								
Without time trend								
0%	0.037	0.049	0.005	0.047	−0.001	0.048	0.049	0.940
20%	0.043	0.049	0.005	0.047	−0.004	0.051	0.051	0.945
40%	0.051	0.050	0.006	0.047	0.003	0.055	0.056	0.946
60%	0.060	0.050	0.007	0.047	0.001	0.063	0.064	0.952
With time trend								
0%	−0.567	0.047	−0.020	0.045	0.016	0.046	0.044	0.943
20%	−0.728	0.049	−0.038	0.045	0.019	0.048	0.047	0.941
40%	−0.952	0.050	−0.071	0.045	0.024	0.051	0.051	0.948
60%	−1.277	0.053	−0.134	0.045	0.036	0.058	0.058	0.950
$n = 600$								
Without time trend								
0%	0.036	0.036	0.003	0.035	0.001	0.035	0.034	0.943
20%	0.042	0.036	0.004	0.035	0.003	0.037	0.036	0.945
40%	0.050	0.036	0.005	0.035	0.005	0.040	0.039	0.938
60%	0.059	0.037	0.006	0.035	0.002	0.046	0.045	0.944
With time trend								
0%	−0.565	0.033	−0.022	0.031	0.008	0.032	0.031	0.942
20%	−0.724	0.034	−0.040	0.031	0.013	0.033	0.033	0.935
40%	−0.946	0.035	−0.072	0.031	0.016	0.035	0.036	0.947
60%	−1.271	0.037	−0.136	0.031	0.016	0.041	0.041	0.942

Note: p_m is the missing probability of the covariate values at regular visits; Bias is the relative bias computed by dividing the difference of the mean of the 1000 estimated parameters and the true value by the true value; SD is the standard deviation of the 1000 estimated values; ASE is the mean of the 1000 estimated standard errors by bootstrap method; CP is the proportion of 95% confidence intervals covering the true value.

Table 2.

Simulation results for the model with a binary time-dependent covariate using the last covariate carried forward (LCCF) method, linear interpolation method (Linear), and the proposed kernel smoothing method (Proposed).

	LCCF		Linear		Proposed			
p_m	Bias	SD	Bias	SD	Bias	SD	ASE	CP
$n = 300$								
Without time trend								
0%	−0.095	0.033	0.153	0.042	0.003	0.041	0.039	0.933
20%	−0.141	0.033	0.107	0.041	0.004	0.043	0.042	0.941
40%	−0.206	0.032	0.031	0.039	0.006	0.046	0.046	0.949
60%	−0.299	0.031	−0.094	0.037	0.008	0.053	0.053	0.945
With time trend								
0%	−0.108	0.033	0.148	0.041	−0.006	0.040	0.040	0.944
20%	−0.155	0.032	0.104	0.040	−0.004	0.042	0.042	0.952
40%	−0.221	0.031	0.031	0.038	−0.003	0.045	0.046	0.953
60%	−0.319	0.031	−0.096	0.038	−0.002	0.054	0.053	0.938
$n = 600$								
Without time trend								
0%	−0.096	0.023	0.151	0.029	0.001	0.028	0.028	0.933
20%	−0.142	0.023	0.106	0.029	0.002	0.031	0.029	0.930
40%	−0.206	0.022	0.029	0.027	0.003	0.032	0.032	0.940
60%	−0.300	0.021	−0.095	0.026	0.004	0.037	0.036	0.941
With time trend								
0%	−0.106	0.024	0.150	0.029	−0.003	0.028	0.028	0.941
20%	−0.154	0.023	0.106	0.028	−0.002	0.029	0.030	0.945
40%	−0.222	0.022	0.030	0.027	−0.001	0.031	0.032	0.958
60%	−0.319	0.022	−0.096	0.026	−0.001	0.037	0.037	0.944

Note: p_m is the missing probability of the covariate values at regular visits; Bias is the relative bias computed by dividing the difference of the mean of the 1000 estimated parameters and the true value by the true value; SD is the standard deviation of the 1000 estimated values; ASE is the mean of the 1000 estimated standard errors by bootstrap method; CP is the proportion of 95% confidence intervals covering the true value.

Table 3.

Simulation results for the extensions of the proposed method: (a) both time-dependent and time-independent covariates are present, where β is the coefficient for the binary, time-dependent covariate Z with time trend, and γ is for the continuous, time-independent covariate W; (b) two time-dependent covariates with different observation time schedules, where β_1 is the coefficient for the continuous covariate Z_1 and β_2 is for the binary covariate Z_2.

(a)

	β				γ			
p_m	**Bias**	**SD**	**ASE**	**CP**	**Bias**	**SD**	**ASE**	**CP**
n = 300								
0%	−0.004	0.050	0.052	0.949	−0.015	0.066	0.065	0.952
20%	−0.000	0.055	0.056	0.949	−0.016	0.066	0.066	0.951
40%	−0.000	0.061	0.063	0.944	−0.017	0.068	0.067	0.952
60%	0.006	0.076	0.075	0.940	−0.013	0.070	0.070	0.951
n = 600								
0%	−0.002	0.036	0.036	0.942	0.000	0.047	0.046	0.944
20%	−0.002	0.039	0.039	0.946	0.000	0.047	0.046	0.944
40%	−0.001	0.043	0.044	0.947	−0.000	0.048	0.047	0.940
60%	0.004	0.053	0.052	0.945	0.002	0.049	0.049	0.948

(b)

Time trend		β_1				β_2			
Z_1	Z_2	**Bias**	**SD**	**ASE**	**CP**	**Bias**	**SD**	**ASE**	**CP**
n = 300									
No	No	0.013	0.058	0.056	0.944	−0.000	0.052	0.052	0.962
Yes	No	0.016	0.057	0.059	0.952	0.011	0.049	0.048	0.943
No	Yes	0.014	0.056	0.055	0.949	−0.010	0.055	0.053	0.928
Yes	Yes	0.030	0.056	0.058	0.950	0.006	0.050	0.049	0.938
n = 600									
No	No	−0.003	0.040	0.039	0.941	0.002	0.036	0.036	0.946
Yes	No	0.009	0.044	0.045	0.960	0.007	0.034	0.034	0.950
No	Yes	0.016	0.040	0.039	0.936	−0.008	0.036	0.037	0.952
Yes	Yes	0.001	0.043	0.044	0.958	0.003	0.036	0.035	0.929

Note: p_m is the missing probability of the covariate values at regular visits; Bias is the relative bias computed by dividing the difference of the mean of the 1000 estimated parameters and the true value by the true value; SD is the standard deviation of the 1000 estimated values; ASE is the mean of the 1000 estimated standard errors by bootstrap method; CP is the proportion of 95% confidence intervals covering the true value.

Author Manuscript

Author Manuscript

Author Manuscript

Table 4.

Results of simulations when both time-dependent and time-independent covariates are present, using the extension of the proposed method described in Section 3.1 (Extension) and the estimator without discriminating time-dependent and -independent covariates (No Extension).

	β				γ			
	Extension		**No Extension**		**Extension**		**No Extension**	
p_m	**Bias**	**SD**	**Bias**	**SD**	**Bias**	**SD**	**Bias**	**SD**
$n = 300$								
0%	−0.004	0.050	−0.003	0.051	−0.015	0.066	−0.013	0.066
20%	−0.000	0.055	0.000	0.055	−0.016	0.066	−0.016	0.071
40%	−0.000	0.061	0.000	0.061	−0.017	0.068	−0.016	0.079
60%	0.006	0.076	0.006	0.076	−0.013	0.070	−0.009	0.094
$n = 600$								
0%	−0.002	0.036	−0.002	0.036	0.000	0.047	−0.001	0.047
20%	−0.002	0.039	−0.002	0.039	0.000	0.047	0.004	0.051
40%	−0.001	0.043	−0.001	0.043	−0.000	0.048	0.004	0.056
60%	0.004	0.053	0.004	0.053	0.002	0.049	−0.001	0.062

Note: β is the coefficient for the (binary) time-dependent covariate Z; γ is for the (continuous) time-independent covariate W; p_m is the missing probability of the covariate values at regular visits; Bias is the relative bias computed by dividing the difference of the mean of the 1000 estimated parameters and the true value by the true value; SD is the standard deviation of the 1000 estimated values.