The Canadian Journal of Statistics
Vol. 43, No. 4, 2015, Pages 519–533
La revue canadienne de statistique

Robust estimation for longitudinal data with informative observation times

Kin-yat LIU[1] and Xingqiu ZHAO[1,2]*

[1]*Department of Applied Mathematics, The Hong Kong Polytechnic University, Hung Hom, Hong Kong*
[2]*The Hong Kong Polytechnic University Shenzhen Research Institute, Shenzhen, China*

Key words and phrases: Estimating equation; informative observation process; longitudinal data; model checking; robust estimation.

MSC 2010: Primary 62J05, 62G08; secondary 62E20

Abstract: In this paper we focus on regression analysis of irregularly observed longitudinal data that often occur in medical follow-up studies and observational investigations. The analysis of these data involves two processes. One is the underlying longitudinal response process of interest and the other is the observation process that controls observation times. Most of the existing methods, however, rely on some restrictive models or assumptions such as the Poisson assumption. For this we propose a class of more flexible joint models and a robust estimation approach for regression analysis of longitudinal data with related observation times. The asymptotic properties of the proposed estimators are established and a model checking procedure is also presented. The numerical studies indicate that the proposed methods work well for practical situations. *The Canadian Journal of Statistics* 43: 519–533; 2015 © 2015 Statistical Society of Canada

Résumé: Les auteurs s'intéressent à la régression pour des données longitudinales observées de façon irrégulière, une situation fréquente dans les études comportant un suivi médical et dans les études observationnelles. De telles données émergent de deux processus : le processus de réponses longitudinales sous-jacent et le processus qui contrôle les temps d'observation. La plupart des méthodes existantes se basent sur des hypothèses ou un modèle restrictifs telle que l'hypothèse de Poisson. Les auteurs proposent une classe plus flexible de modèles conjoints et une approche robuste d'estimation de la régression pour des données longitudinales et leur temps d'observation. Ils établissent les propriétés asymptotiques de l'estimateur proposé et présentent des procédures de vérification des hypothèses. Les auteurs présentent également des études numériques qui indiquent que les méthodes proposées fonctionnent bien dans des situations pratiques. *La revue canadienne de statistique* 43: 519–533; 2015 © 2015 Société statistique du Canada

1. INTRODUCTION

The analysis of longitudinal data has become more and more important as such data frequently occur in medical follow-up studies and observational investigations. For the analysis of longitudinal data, various methods have been developed, mostly under the fixed observation scheme or the assumption that the longitudinal response process and the observation process are independent completely, or given the covariates, that is, the observation process is non-informative. For example, Diggle, Liang & Zeger (1994) provided an excellent summary on commonly used methods such as the generalized estimating equation and random-effect modelling approaches. Lin & Ying (2001) and Welsh, Lin, & Carroll (2002) discussed general semiparametric regression analysis of longitudinal data assuming the response variable, observation times, and censoring times are independent given the covariates.

* *Author to whom correspondence may be addressed.*
E-mail: xingqiu.zhao@polyu.edu.hk

© 2015 Statistical Society of Canada / Société statistique du Canada

In practice, the longitudinal response and observation processes may still be related even given the covariates, which means that the observation times may contain potential information about the response variable of interest, that is, the observation times may be informative. For example, informative observation processes may be hospitalization times of subjects in the study (Wang Qin, & Chiang, 2001), or clinical visit times of patients in medical follow-up studies (Sun & Wei, 2000; Zhang, 2002; Liu, Huang, & O'Quigley, 2008). In a bladder cancer follow-up study conducted by the Veterans Administration Cooperative Urological Research Group (Byar, 1980), some patients had significantly more clinical visits than others, which implies that the visit times may be related to the occurrence rate of bladder tumours. In the medical cost data, some patients were more likely to visit hospital, and paid more for their visits, which indicates that the patients' visiting times are informative on medical costs. Lipsitz et al. (2002) also presented a set of longitudinal data from a study of children with acute lymphoblastic leukemia where the response and observation processes may be correlated. However there is limited research that addresses the analysis of such correlated longitudinal data. Sun et al. (2005) discussed conditional semiparametric models that allow observation times to be correlated with the longitudinal process by treating the observation process as covariates; Sun, Sun, & Liu (2007) proposed a joint model for the longitudinal process and the observation process through a shared latent variable; Liang, Lu, & Ying (2009) presented a joint model through two random effects and made a valid inference by specifying the relationship between the random effects and using a parametric distribution assumption for a random effect; Zhu et al. (2011) proposed a joint modelling approach through a shared unobserved random variable; Zhu et al. (2011) proposed a general joint model for longitudinal and observation processes through a latent variable and an unspecified link function. All these methods require a common and key assumption that the observation process is a mixed Poisson process, which may not be true in practice. There is limited research on non-Poisson observation processes. For example, Lin, Scharfstein, & Rosenheck (2004) studied a class of marginal regression models for longitudinal responses, where an intensity model with respect to the visit process was required to be correctly specified. Liu, Huang, & O'Quigley (2008) proposed a joint Gaussian random effects model and analyzed medical cost data. Song, Mu, & Sun (2012) developed a joint modelling approach for the two correlated longitudinal and observation processes through two dependent variables. However these methods mentioned above require estimation of observation process models for making a valid inference on the longitudinal response process. The aim of this paper is to consider more general joint models for longitudinal data with dependent observation times and to develop a robust estimation approach.

The remainder of this work is organized as follows. In Section 2 we will begin by introducing the notation and assumptions and then present the models that will be used below. A robust estimation procedure is also presented in Section 2 for the parameters of interest and the asymptotic properties of the resulting estimators are established. In Section 3 a model checking procedure is provided. Section 4 reports some simulation results obtained for assessing the finite sample properties of the proposed estimates and diagnostic tests, and an illustrative example is given in Section 5. Section 6 concludes with some discussion and remarks.

2. JOINT MODELLING AND ESTIMATION PROCEDURE

Consider a longitudinal study that consists of n independent subjects and let $Y_i(t)$ denote the longitudinal response variable of interest before or at time t for subject i. Suppose that for each subject, there exists a p-dimensional vector of covariates denoted by X_i. Given X_i and an unobserved positive random variable Z_i that is independent of X_i, the mean function of $Y_i(t)$ has the form

$$E\{Y_i(t)|X_i, Z_i\} = \mu_0(t) + X_i'\beta + g(Z_i). \tag{1}$$

Here, $\mu_0(t)$ is a completely unknown continuous baseline mean function, β is a vector of unknown regression parameters, and $g(\cdot)$ is a completely unspecified link function. For identifiability reasons we assume that $E(g(Z_i)) = 0$.

For subject i, suppose that $Y_i(\cdot)$ is observed only at finite time points $T_{i1} < \cdots < T_{iK_i}$, where K_i denotes the potential number of observation times. That is, only the values of $Y_i(t)$ at these observation times are known and we have panel data on the $Y_i(t)$. Let C_i denote the follow-up time associated with subject i and thus $Y_i(t)$ is observed only at these T_{ij} with $T_{ij} \le C_i$, $i = 1, \ldots, n$. Define $\tilde{O}_i(t) = O_i(\min(t, C_i))$, where $O_i(t) = \sum_{j=1}^{K_i} I(T_{ij} \le t)$, $i = 1, \ldots, n$. Then $\tilde{O}_i(t)$ is a point process characterizing the ith subject's observation process and jumps only at the observation times.

For the observation process we will assume that $O_i(t)$ satisfies the following rate function model:

$$E\{dO_i(t)|X_i, Z_i\} = Z_i h(X_i) d\Lambda_0(t), \tag{2}$$

where $h(\cdot)$ is a completely unspecified positive function and $\Lambda_0(\cdot)$ is a completely unknown continuous baseline function. Under model (2), one does not need the Poisson assumption anymore. In the following, it will be assumed that given (X_i, Z_i), $Y_i(t)$, and $O_i(t)$ are independent. Also C_i is independent of $\{Y_i, O_i, X_i, Z_i\}$ and $\{Y_i(t), O_i(t), C_i, X_i, 0 \le t \le \tau\}_{i=1}^n$ are independent and identically distributed, where τ denotes the length of the study. Our focus is to estimate the regression parameter β.

Remark 1. *The $g(\cdot)$ is an unknown link function that is used to characterize the association between the two processes $Y_i(t)$ and $O_i(t)$. To see this we suppose that $Y_i(t)$ follows a semiparametric random effects model*

$$Y_i(t) = \mu_0(t) + X_i'\beta + V_i + \varepsilon_i(t),$$

where V_i is a random variable of subject-specific effect, and $\varepsilon_i(t)$ is a zero mean measurement error process. Taking the conditional expectation of $Y_i(t)$ given X_i and Z_i we obtain model (1) with $g(Z_i) = E(V_i|Z_i)$.

To estimate β, note that if the latent variables Z_i are known, model (1) would become the usual linear mean model. Unfortunately, the Z_i's are unknown in practice. One natural approach is to estimate the Z_i first and then treat them as known. In the following we take a different approach motivated by that proposed in Sun & Wei (2000) among others.

Specifically, define

$$\bar{Y}_i = \sum_{j=1}^{m_i} Y_i(T_{ij}) I(T_{ij} \le \tau) = \int_0^\tau Y_i(t) d\tilde{O}_i(t),$$

where $m_i = \tilde{O}_i(C_i)$, the total number of observations on subject i, $i = 1, \ldots, n$. Then we have

$$\begin{aligned} E(\bar{Y}_i|X_i) &= E(Z_i)E(\Lambda_0(C_i))h(X_i)(X_i'\beta) \\ &\quad + h(X_i)\int_0^\tau [E(Z_i)\mu_0(t) + E\{g(Z_i)Z_i\}]P(C_i \ge t)d\Lambda_0(t). \end{aligned} \tag{3}$$

Note that Equation (3) involves unknown parameters and unspecified functions. However the existing estimation approaches (Lawless & Nadeau, 1995; Sun & Wei, 2000; Wang, Qin, & Chiang, 2001) are no longer applicable to model (2) as $h(\cdot)$ is completely unspecified. To solve this problem we consider the conditional mean of m_i given X_i. Under model (2) we have

$$E(m_i|X_i) = E(Z_i)E\{\Lambda_0(C_i)\}h(X_i). \tag{4}$$

Thus, combining (3) and (4) yields

$$E(\bar{Y}_i|X_i) = E(m_i|X_i)(X_i'\beta + \theta), \tag{5}$$

where

$$\theta = \frac{\int_0^\tau [E(Z_i)\mu_0(t) + E\{g(Z_i)Z_i\}]P(C_i \geq t)d\Lambda_0(t)}{E(Z_i)E(\Lambda_0(C_i))}$$

is an unknown parameter.

Remark 2. *Here, θ is a nuisance parameter. As our main focus is on estimation of covariate effects on the response process we can estimate θ to avoid estimating two unknown functions $\mu_0(t)$ and $\Lambda_0(t)$ and an unknown association term $E(g(Z_i)Z_i)$ between the longitudinal response and observation processes. If we specify $E(Z_i) = 1$, then*

$$\theta = \frac{\int_0^\tau \mu_0(t)P(C_i \geq t)d\Lambda_0(t)}{E(\Lambda_0(C_i))} + Cov(g(Z_i), Z_i),$$

where the first component of θ depends on the baseline mean functions of two processes and the survival distribution of the follow-up time and can be regarded as the integrated baseline mean function of the longitudinal response process, and the second component of θ characterizes the relation between the two processes. When the longitudinal response and observation processes are independent we have $g(Z_i) = 0$, and so the second component of θ is 0. The proposed method in this paper provides a robust solution for estimating covariate effects on the response process no matter whether the two processes are related or not, while estimation of $Cov(g(Z_i), Z_i)$ is beyond the scope of this paper.

For estimation of β, motivated by Equation (5) we propose to use the following class of estimating functions

$$U(\beta_1) = \sum_{i=1}^{n} W_i X_{1i}\{\bar{Y}_i - m_i X_{1i}'\beta_1\} = 0, \tag{6}$$

where the W_i are some weights that could depend on X_i, $X_{1i}' = (X_i', 1)$ and $\beta_1' = (\beta', \theta)$.

Remark 3. *Clearly, when W_i depends on X_i only, $E\{U(\beta_{10})\} = 0$ for the true value β_{10} of β_1. Naturally, one may ask whether the weight can be taken as $W_i = m_i$. To answer this question we consider a simple case where the longitudinal response and observation processes are independent and m_i given X_i and Z_i follows a Poisson distribution. By direct calculations we have*

$$E\left[m_i X_{1i}\{\bar{Y}_i - m_i X_{1i}'\beta_{10}\}\right] = -E(Z_i)E\{\Lambda_0(C_i)\}E[h(X_i)X_{1i}X_{1i}'\beta_{10}] \neq 0.$$

Hence, the estimating equation with $W_i = m_i$ is biased for this case.

Let $\hat{\beta}_1 = (\hat{\beta}', \hat{\theta})'$ denote the solution to Equation (6). Then we have

$$\hat{\beta}_1 = \left[\sum_{i=1}^{n} W_i m_i X_{1i}X_{1i}'\right]^{-1} \sum_{i=1}^{n} W_i X_{1i}\bar{Y}_i. \tag{7}$$

Then we will show in Appendix A.1 that under some regularity conditions, the estimator $\hat{\beta}_1$ is a consistent estimator of true parameter β_{10} and $\sqrt{n}(\hat{\beta}_1 - \beta_{10})$ has asymptotically a normal distribution with mean zero and covariance matrix that can be consistently estimated by $\hat{\Gamma}^{-1}\hat{\Sigma}\hat{\Gamma}^{-1}$, where

$$\hat{\Gamma} = \frac{1}{n}\sum_{i=1}^{n} W_i m_i X_{1i} X_{1i}' \quad \text{and} \quad \hat{\Sigma} = \frac{1}{n}\sum_{i=1}^{n} \hat{\phi}_i \hat{\phi}_i'$$

with $\hat{\phi}_i = W_i X_{1i}\{\bar{Y}_i - m_i X_{1i}'\hat{\beta}_1\}$.

3. MODEL CHECKING PROCEDURE

In practice, in addition to the estimation of β, one may also be interested in checking the adequacy of model (1). Following Lin et al. (2000) we develop a model checking procedure using certain cumulative sums of the residuals. For this, define

$$\mathcal{A}(t) = \frac{\int_0^t [E(Z_i)\mu_0(u) + E\{g(Z_i)Z_i\}]P(C_i \geq u)d\Lambda_0(u)}{E(Z_i)E(\Lambda_0(C_i))}.$$

As $E\left[\int_0^t \{Y_i(u) - \beta_0' X_i\} d\tilde{O}_i(u) | X_i\right] = E(m_i|X_i)\mathcal{A}(t)$ we can estimate $\mathcal{A}(t)$ by

$$\hat{\mathcal{A}}(t) = \frac{\sum_{i=1}^n \int_0^t \{Y_i(u) - \hat{\beta}' X_i\} d\tilde{O}_i(u)}{\sum_{i=1}^n m_i}.$$

Furthermore, for each i, $i = 1, \ldots, n$, define the residual

$$\hat{R}_i(t) = \int_0^t \{Y_i(u) - \hat{\beta}' X_i\} d\tilde{O}_i(u) - m_i \hat{\mathcal{A}}(t).$$

To check the functional form for the jth component of X formally we consider the process

$$\mathcal{F}_j(x) = \frac{1}{\sqrt{n}}\sum_{i=1}^{n} I(X_{ji} \leq x)\tilde{R}_i,$$

where $\tilde{R}_i = \hat{R}_i(\tau)$.

Let

$$S_0 = \frac{1}{n}\sum_{i=1}^{n} m_i, \qquad S_j(x) = \frac{1}{n}\sum_{i=1}^{n} I(X_{ji} \leq x) m_i,$$

and

$$B_j(t, x) = \frac{1}{n}\sum_{i=1}^{n} \left\{ I(X_{ji} \leq x) - \frac{S_j(x)}{S_0} \right\} X_i \tilde{O}_i(t).$$

To apply the statistic $\mathcal{F}_j(x)$, we show in Appendix A.2 that its null distribution can be approximated by the zero-mean Gaussian process

$$\hat{\mathcal{F}}_j(x) = \frac{1}{\sqrt{n}}\sum_{i=1}^{n} \left\{ I(X_{ji} \leq x) - \frac{S_j(x)}{S_0} \right\} \tilde{R}_i G_i - B_j(\tau, x)' \frac{1}{\sqrt{n}}\sum_{i=1}^{n} \hat{d}_i G_i,$$

where $\hat{d}_i$ is the vector $\hat{\Gamma}^{-1}\hat{\phi}_i$ without the last entry and $(G_1, \ldots, G_n)$ are independent standard normal variables independent of the data. This suggests that one can approximate the distribution of $\mathcal{F}_j(x)$ by the empirical distribution of a large number of realizations of $\hat{\mathcal{F}}_j(x)$ given by repeatedly generating the standard normal random sample $(G_1, \ldots, G_n)$ given the observed data. For checking the functional form of the jth component of covariates in model (1) we can apply the supremum test statistic $\sup_x |\mathcal{F}_j(x)|$, where the P-value can be obtained by comparing the observed value of $\sup_x |\mathcal{F}_j(x)|$ to a large number of realizations of $\sup_x |\hat{\mathcal{F}}_j(x)|$.

Finally, to test the goodness-of-fit of model (1) we propose to apply the statistic

$$\phi(t, x) = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} I(X_i \le x)\hat{R}_i(t),$$

where the event $I(X_i \le x)$ means that each of the components of X_i is not larger than the corresponding component of x. It is easy to see that $\phi(t, x)$ is the cumulative sum of $\hat{R}_i(t)$ over the values of X_i. In Appendix A.2 we show that the null distribution of $\phi(t, x)$ can be approximated by the zero-mean Gaussian process $\hat{\phi}(t, x)$, which is obtained from the expression for $\hat{\mathcal{F}}_j(x)$ by replacing $I(X_{ji} \le x)$ with $I(X_i \le x)$, $\tilde{R}_i$ with $\hat{R}_i(t)$, and $B_j(\tau, x)$ with $B(t, x)$. Graphical and numerical procedures can be conducted in the same fashion as for $\mathcal{F}_j(x)$. Thus for checking the overall fit of models (1) and (2) based on $\phi(t, x)$, the P-value of the omnibus test can be obtained by comparing the observed value of $\sup_{t,x} |\phi(t, x)|$ to a large number of realization of $\sup_{t,x} |\hat{\phi}(t, x)|$.

4. SIMULATION STUDIES

We conducted three simulation studies to assess the performance of the proposed inference procedure with the focus on estimation of regression parameter β. The purpose of the first one was to evaluate the finite sample properties of the proposed estimator, whereas in the second study we compared the proposed estimator to that given in Zhu et al. (2011). In the third simulation study we evaluated the sizes and powers of the diagnostic tests when the sample size is finite. In particular, to demonstrate the robustness of the proposed approach we considered linear and non-linear forms of the link function g and different patterns of the observation scheme in each simulation study.

For the first study we considered the situation where there exist two covariates, X_{1i} and X_{2i}, which follow the Bernoulli distribution with success probability 0.5 and the uniform distribution over interval [0, 1], respectively. The latent variables Z_i were generated from the gamma distribution with shape parameter 10 and scale parameter 10 (equivalently, mean 100 and variance 1,000). We considered two cases of $g(Z_i)$. In the first case we took $g(Z_i) = \rho\{Z_i - E(Z_i)\}/\sqrt{\text{Var}(Z_i)}$. In the second case we took $g(Z_i) = \rho\{\log(Z_i) - E(\log(Z_i))\}/\sqrt{\text{Var}(\log(Z_i))}$. Here ρ characterizes the relationship between the observation process and the longitudinal response process. When $\rho > 0$, the two processes are positively correlated; when $\rho = 0$, the two processes have no correlation given the covariates; when $\rho < 0$, the two processes are negatively correlated. Here, three situations with $\rho = -0.5$, 0, and 0.5 were considered. The follow-up time C_i was generated from the uniform distribution over $[\tau/2, \tau]$ with $\tau = 18$.

With respect to the observation process $O_i(t)$, three set-ups were considered as follows:

(i) Given X_i, Z_i, and C_i, the number of observation times m_i was assumed to follow the Poisson distribution with mean

$$\Lambda(C_i|X_i, Z_i) = Z_i\Lambda_0(C_i)\exp(X_i'\gamma) = \frac{Z_iC_i\exp(X_i'\gamma)}{\tau},$$

$i = 1, 2, \ldots, n$. The observation times $(T_{i1}, \ldots, T_{im})$ were taken to be the order statistics of a random sample of size m_i from the uniform distribution over $(0, C_i)$.

(ii) Given X_i, Z_i, and C_i, the number of observation times m_i was assumed to follow the Poisson distribution with mean

$$\Lambda(C_i|X_i, Z_i) = Z_i \Lambda_0(C_i) \exp(X_i'\gamma) = \frac{Z_i C_i(C_i/2 + 1) \exp(X_i'\gamma)}{\tau(\tau/2 + 1)},$$

$i = 1, 2, \ldots, n$. The observation times $(T_{i1}, \ldots, T_{im})$ were taken to be the order statistics of a random sample of size m_i from the cumulative distribution function

$$\frac{t^2/2 + t}{C_i^2/2 + C_i} I(0 \le t \le C_i).$$

(iii) Given X_i, Z_i, and C_i, the interarrival times were assumed to follow the Weibull distribution with shape parameter 0.5 and scale parameter $\tau/(Z_i C_i \exp(X_i'\gamma))$. The m_i was the number of observation times which were less than C_i.

For the response variable, it was assumed that

$$Y_i(t) = \mu_0(t) + X_i'\beta + g(Z_i) + \varepsilon_i(t),$$

where $\mu_0(t) = \sin(t)$ and $\varepsilon_i(t) \sim N(0, 0.1)$. We took $\gamma = (1, 1)'$ and $\beta = (1, 1)'$, representing the covariate effects on the observation scheme and the response variable, respectively. For each setting we considered $n = 120$ and 240. All the results reported here were based on 1,000 Monte Carlo replications.

Tables 1–3 present the simulation results obtained on estimation of β_1 and β_2 with weight $W_i = 1$ for sample size $n = 120$ and 240. All tables include the estimated bias (BIAS) given by the average of proposed estimates of β minus the true value, the sample standard error (SSE) of the proposed estimates, the mean of the estimated standard error (ESE), and the empirical 95% coverage probabilities (CP). These results indicate that the proposed estimate seems to be unbiased and the proposed variance estimation procedure provides reasonable estimates. Also the results on the empirical coverage probabilities indicate that the normal approximation seems to be appropriate.

To further investigate the robustness of the proposed estimation procedure and also why one may need to use the proposed estimator instead of the estimators developed under the restricted models such as that given in Zhu et al. (2011) we conducted a simulation study to compare the performance of the proposed estimators and those presented in Zhu et al. (2011).

For this second study we considered the situation where there exists one covariate. The covariate X_i and the latent variable Z_i were generated from the Bernoulli distribution with success probability 0.5 and the gamma distribution with shape parameter 5 and scale parameter 5 (equivalently, mean 25 and variance 125), respectively. The two forms of $g(Z_i)$, the generation of the follow-up time C_i, the three set-ups for the observation process $O_i(t)$ were as defined in the first simulation study. For the longitudinal response variable, it was assumed that

$$Y_i(t) = \mu_0(t) + X_i\beta + g(Z_i) + \varepsilon_i(t),$$

where $\mu_0(t) = \sin(t)$ or $\log(1 + t)$ and $\varepsilon_i(t) \sim N(0, 0.1)$. We took $\beta = 1$. For each setting we considered $n = 100$ and 200, and $\rho = 0.5$, 0, and -0.5.

Table 4 provides the BIAS of the estimates proposed here and in Zhu et al. (2011), and the relative efficiency (RE) which is the ratio of the mean squared error of the estimator given in Zhu et al. (2011) to that of the proposed estimator, where ZTS denotes the estimator presented in Zhu

TABLE 1: Estimation results of β_1 and β_2 with a homogeneous Poisson observation process.

	Linear $g(Z_i)$				Non-linear $g(Z_i)$			
	$n = 120$		$n = 240$		$n = 120$		$n = 240$	
	β_1	β_2	β_1	β_2	β_1	β_2	β_1	β_2
$\rho = 0.5$								
BIAS	−0.0047	−0.0073	−0.0023	0.0080	0.0018	−0.0029	−0.0032	0.0018
SSE	0.1128	0.2103	0.0807	0.1575	0.0933	0.1707	0.0723	0.1307
ESE	0.1111	0.2054	0.0808	0.1500	0.0933	0.1713	0.0704	0.1309
CP	0.9410	0.9370	0.9430	0.9480	0.9450	0.9460	0.9480	0.9490
$\rho = 0$								
BIAS	0.0002	0.0006	0.0003	−0.0004	−0.0002	−0.0009	0.0000	0.0012
SSE	0.0140	0.0243	0.0098	0.0177	0.0142	0.0262	0.0100	0.0178
ESE	0.0142	0.0248	0.0100	0.0177	0.0141	0.0247	0.0101	0.017
CP	0.9450	0.9450	0.9500	0.9490	0.9520	0.9370	0.9510	0.9420
$\rho = -0.5$								
BIAS	−0.0003	0.0000	−0.0013	0.0019	−0.0061	0.0122	−0.0004	0.0065
SSE	0.1180	0.2136	0.0841	0.1509	0.0944	0.1711	0.0677	0.1368
ESE	0.1123	0.2066	0.0809	0.1502	0.0889	0.1656	0.0690	0.1291
CP	0.9360	0.9340	0.9390	0.9440	0.9300	0.9430	0.9460	0.9370

et al. (2011). It can be seen from the table that the proposed estimator seems to be unbiased and more efficient for all the situations considered here, whereas the ZTS estimator is clearly biased for the third case due to the misspecification of the observation process. A possible reason is that the ZTS method is a two-step estimation procedure depending on estimation of observation process model parameters. In general, the proposed estimation procedure does not rely on the form of the link function g or the pattern of observation process $O_i(t)$, and thus it is more robust.

In the third study we examined the finite sample properties of the proposed diagnostic tests. For the purpose we considered the situation where there exists one covariate X_i which follows the uniform distribution over $\{1, 2, 3, 4, 5\}$. For the latent variable Z_i, the follow-up time C_i, the link function g, and the observation process $O_i(t)$, the same set-ups were considered as in the second simulation study.

In order to evaluate the performance of the diagnostic tests with respect to the functional form of the covariate and the goodness-of-fit of models, the longitudinal response variables were assumed to satisfy, respectively,

$$Y_i(t) = \mu_0(t) + X_i^{\nu}\beta + g(Z_i) + \varepsilon_i(t),$$

and

$$Y_i(t) = \mu_0(t) + (X_i\beta)^{\nu} + g(Z_i) + \varepsilon_i(t),$$

where $\mu_0(t) = \sin(t)$ and $\varepsilon_i(t) \sim N(0, 0.25)$. We took $\beta = 1$. For each setting we considered $n = 100$ and 200, $\rho = 0.5$, and $\nu = 1, 1.2, 1.4, \ldots, 2.8$. Again, all the results reported here were based on 1,000 Monte Carlo replications.

TABLE 2: Estimation results of β_1 and β_2 with a non-homogeneous Poisson observation process.

	Linear $g(Z_i)$				Non-linear $g(Z_i)$			
	$n = 120$		$n = 240$		$n = 120$		$n = 240$	
	β_1	β_2	β_1	β_2	β_1	β_2	β_1	β_2
$\rho = 0.5$								
BIAS	−0.0021	0.0009	−0.0019	0.0037	0.0001	−0.0091	0.0002	0.0066
SSE	0.1174	0.2234	0.0873	0.1567	0.0994	0.1806	0.0789	0.1467
ESE	0.1150	0.2101	0.0834	0.1553	0.0959	0.1772	0.0774	0.1443
CP	0.9440	0.9330	0.9400	0.9510	0.9510	0.9410	0.9500	0.9430
$\rho = 0$								
BIAS	−0.0005	0.0005	−0.0003	0.0009	0.0004	0.0000	−0.0003	0.0003
SSE	0.0138	0.0237	0.0102	0.0173	0.0138	0.0240	0.0096	0.0168
ESE	0.0137	0.0232	0.0097	0.0166	0.0136	0.0234	0.0097	0.0166
CP	0.9400	0.9420	0.9440	0.9450	0.9450	0.9390	0.9500	0.9470
$\rho = -0.5$								
BIAS	0.0038	−0.0024	0.0019	0.0018	0.0007	0.0037	0.0005	−0.0046
SSE	0.1225	0.2307	0.0859	0.1621	0.1001	0.1940	0.0722	0.1330
ESE	0.1167	0.2127	0.0835	0.1552	0.0997	0.1840	0.0708	0.1323
CP	0.9400	0.9330	0.9470	0.9460	0.9450	0.9380	0.9510	0.9520

Tables 5 and 6 include the empirical sizes and powers of the proposed tests. It can be seen from the tables that the estimated sizes of the tests are close to the nominal size 5%, and the estimated powers are reasonable and increase when v increases or n increases, as expected. These simulation results suggest that the null distributions of the proposed test statistics are well approximated and the tests have good power properties.

5. AN ILLUSTRATIVE EXAMPLE

In this section we apply our proposed methodology in the previous sections to a set of longitudinal data from a bladder cancer study conducted by the Veterans Administration Cooperative Urological Research Group (Byar, 1980; Andrews & Herzberg, 1985; Sun & Wei, 2000; Wellner & Zhang, 2000; Zhang, 2002). In the study, all the patients had superficial bladder tumours, and there were three treatment groups: placebo, thiotepa, or pyridoxine. At the beginning of the study, the patients were randomly assigned to one of the treatment groups. During the study, many patients had multiple recurrences of the bladder tumours and all recurrences between visits were recorded and the new tumours were removed at clinical visits. From the observed data we found that the number of visits and visit time points varied from patient to patient. In addition, for each patient, the number of initial tumours and the size of the largest initial tumour were reported as two important baseline covariates.

For the analysis we took $Y_i(t)$ to be the logarithm of the number of observed tumours at time t, plus 1 to avoid 0, $i = 1, \ldots, 116$. We set the first component of X_i to be 1 if the ith patient was given the pyridoxine treatment and 0 otherwise, the second component of X_i to be 1 if the ith patient was given the thiotepa treatment and 0 otherwise, the third and the fourth components of X_i to be the

TABLE 3: Estimation results of β_1 and β_2 with a non-Poisson observation process.

	Linear $g(Z_i)$				Non-linear $g(Z_i)$			
	$n = 120$		$n = 240$		$n = 120$		$n = 240$	
	β_1	β_2	β_1	β_2	β_1	β_2	β_1	β_2
$\rho = 0.5$								
BIAS	0.0021	−0.0178	−0.0039	−0.0030	−0.0028	−0.0102	0.0003	0.0022
SSE	0.1162	0.2246	0.0836	0.1544	0.1099	0.2071	0.0705	0.1303
ESE	0.1158	0.2115	0.0840	0.1548	0.1078	0.1993	0.0698	0.1295
CP	0.9380	0.9360	0.9560	0.9460	0.9440	0.9340	0.9490	0.9430
$\rho = 0$								
BIAS	0.0003	0.0012	0.0002	0.0008	0.0001	−0.0007	−0.0005	0.0004
SSE	0.0123	0.0221	0.0085	0.0155	0.0127	0.0230	0.0089	0.0158
ESE	0.0124	0.0217	0.0088	0.0156	0.0124	0.0220	0.0088	0.0155
CP	0.9470	0.9440	0.9560	0.9550	0.9470	0.9310	0.9500	0.9460
$\rho = -0.5$								
BIAS	0.0009	−0.0071	−0.0022	0.0045	0.0023	0.0069	−0.0001	−0.0033
SSE	0.1204	0.2263	0.0848	0.1606	0.1012	0.1880	0.0671	0.1276
ESE	0.1169	0.2152	0.0836	0.1541	0.0991	0.1828	0.0673	0.1247
CP	0.9440	0.9330	0.9530	0.9480	0.9430	0.9340	0.9460	0.9440

number of initial tumours and the size of the largest initial tumour of the ith patient, respectively. Suppose that the longitudinal process of the bladder tumours $Y_i(t)$ and the clinical visit process were described by models (1) and (2), respectively. The application of the proposed estimation procedure with $W_i = 1$ gave $\hat{\beta} = (-0.0085, -0.3232, 0.0729, -0.0102)'$ with the estimated standard errors $(0.1437, 0.0928, 0.0269, 0.0267)'$, and thus P-values $(0.9530, 0.0005, 0.0069, 0.7016)'$. These results indicate that the thiotepa treatment significantly reduced the occurrence rate of the bladder tumours and the number of initial tumours has a significant positive effect on the tumour recurrence rate. However the pyridoxine treatment and the size of the largest initial tumour did not have significant effects on the occurrence rate of the bladder tumours. Sun, Sun, & Liu (2007) applied their method to analyze the same data with the focus on the placebo and the thiotepa groups and obtained that the thiotepa treatment had a significant effect in reducing the recurrence of bladder tumours, but they did not detect the effect of the initial number of bladder tumours on the recurrence rate of the bladder tumour. The reason for this difference between the two application results may be due to the misspecification of the link function between the longitudinal response process and the observation process in Sun, Sun, & Liu (2007). Liang, Lu, & Ying (2009) and Zhu et al. (2011) also analyzed a subset of bladder tumour data and found that both the treatment indicator and the initial tumour number had significant effects on tumour recurrence rate. These results are consistent with those obtained by our proposed approach.

We furthermore applied the model-checking procedures given in Section 3 to the data. For each of the four covariates we computed the proposed test statistics and obtained $\sup_x |\mathcal{F}_1(x)| = 2.650$ with the P-value of 0.441, $\sup_x |\mathcal{F}_2(x)| = 3.210$ with the P-value of 0.121, $\sup_x |\mathcal{F}_3(x)| = 2.451$

TABLE 4: Estimation results of β based on the proposed method and ZTS method.

		$n = 100$			$n = 200$		
	$\mu_0(t)$	ZTS	Proposed	RE	ZTS	Proposed	RE
$\rho = 0.5$							
Homogeneous Poisson	$\sin(t)$	0.0029	−0.0020	1.1179	−0.0065	0.0002	1.0815
	$\log(1+t)$	0.0059	−0.0017	2.0788	−0.0035	0.0008	1.9560
Non-homogeneous Poisson	$\sin(t)$	0.0270	0.0020	1.0519	0.0299	0.0054	1.0978
	$\log(1+t)$	0.0301	0.0005	1.6661	0.0345	0.0050	1.7591
Non-Poisson	$\sin(t)$	−0.3469	0.0034	3.0331	−0.3465	0.0034	3.1248
	$\log(1+t)$	−0.3376	−0.0011	2.8982	−0.3491	0.0024	2.7552
$\rho = 0$							
Homogeneous Poisson	$\sin(t)$	−0.0003	0.0007	1.8728	0.0001	0.0015	1.5488
	$\log(1+t)$	−0.0026	0.0010	11.4730	0.0005	−0.0004	10.9011
Non-homogeneous Poisson	$\sin(t)$	0.0333	0.0001	1.5154	0.0322	0.0005	1.5624
	$\log(1+t)$	0.0335	−0.0005	5.1719	0.0308	−0.0007	5.2976
Non-Poisson	$\sin(t)$	−0.3508	0.0005	8.5140	−0.3427	−0.0006	9.2894
	$\log(1+t)$	−0.3501	−0.0003	4.1698	−0.3408	−0.0023	3.8843
$\rho = -0.5$							
Homogeneous Poisson	$\sin(t)$	0.0026	−0.0016	1.0547	−0.0051	−0.0022	1.1340
	$\log(1+t)$	0.0009	−0.0053	1.6433	−0.0037	−0.0013	1.8634
Non-homogeneous Poisson	$\sin(t)$	0.0254	−0.0026	1.1616	0.0275	0.0067	1.1144
	$\log(1+t)$	0.0282	−0.0048	1.6780	0.0292	0.0063	1.4315
Non-Poisson	$\sin(t)$	−0.3551	−0.0008	2.6833	−0.3390	−0.0007	2.6811
	$\log(1+t)$	−0.3349	−0.0057	2.4942	−0.3455	0.0022	2.4725

with the P-value of 0.725, and $\sup_x |\mathcal{F}_4(x)| = 3.039$ with the P-value of 0.349. All four P-values suggest that the linear form of the covariates is appropriate.

To assess the overall fit of model (1) we calculated the proposed test statistics and obtained $\sup_{t,x} |\phi(t, x)| = 3.133$ with the P-value of 0.486. This suggests that the model seems to be appropriate for fitting the bladder cancer data considered here.

6. CONCLUDING REMARKS

This paper investigated regression analysis of irregularly observed longitudinal data when the observation process may be related to the underlying recurrent event process of interest. For the problem, a class of joint models for the two processes has been proposed through a completely unspecified link function of a latent variable and a corresponding estimating equation-based estimation procedure has been developed for the covariate effect on the recurrent event process. The resulting estimator is consistent and is asymptotically normally distributed. The simulation studies indicate that the estimation approach and model checking procedure perform well and are robust with respect to different forms of the link function and patterns of the observation process.

TABLE 5: Empirical sizes and powers of diagnostic tests for the functional form of a covariate.

	$n = 100$						$n = 200$					
	Linear $g(Z_i)$			Non-linear $g(Z_i)$			Linear $g(Z_i)$			Non-linear $g(Z_i)$		
v	I	II	III	I	II	III	I	II	III	I	II	III
1.0	0.056	0.050	0.047	0.055	0.052	0.047	0.059	0.053	0.044	0.055	0.055	0.052
1.2	0.138	0.132	0.079	0.121	0.137	0.108	0.197	0.181	0.191	0.248	0.237	0.190
1.4	0.299	0.270	0.257	0.269	0.308	0.255	0.511	0.527	0.529	0.660	0.652	0.569
1.6	0.388	0.448	0.338	0.349	0.420	0.334	0.818	0.814	0.764	0.879	0.866	0.764
1.8	0.452	0.525	0.426	0.412	0.512	0.428	0.945	0.955	0.890	0.973	0.967	0.883
2.0	0.542	0.613	0.514	0.508	0.610	0.514	0.991	0.996	0.967	0.995	0.995	0.957
2.2	0.661	0.698	0.620	0.618	0.695	0.611	0.999	1.000	0.994	0.999	1.000	0.984
2.4	0.722	0.801	0.682	0.718	0.810	0.700	0.999	0.999	0.996	1.000	1.000	0.997
2.6	0.812	0.880	0.798	0.815	0.873	0.782	1.000	1.000	1.000	1.000	1.000	1.000
2.8	0.884	0.917	0.835	0.875	0.918	0.844	1.000	1.000	1.000	1.000	1.000	1.000

Note: I, II, and III refer to three setups with respect to observation processes.

The proposed inference procedure has two key advantages. One is that it allows the correlation between the recurrent event process of interest and the observation process in a general format. This is very important as the format of the relationship between the two processes is generally unknown in practice and could be very complicated and thus a flexible model may be preferred. Another is that the proposed approach does not require any estimation for the observation process model. Compared to the existing estimation procedures such as that presented in Zhu et al. (2011), our method is more robust.

TABLE 6: Empirical sizes and powers of diagnostic tests for the goodness-of-fit.

	$n = 100$						$n = 200$					
	Linear $g(Z_i)$			Non-linear $g(Z_i)$			Linear $g(Z_i)$			Non-linear $g(Z_i)$		
v	I	II	III	I	II	III	I	II	III	I	II	III
1.0	0.051	0.054	0.044	0.053	0.052	0.043	0.050	0.052	0.048	0.054	0.053	0.050
1.2	0.145	0.142	0.079	0.136	0.126	0.095	0.179	0.187	0.175	0.227	0.209	0.155
1.4	0.282	0.317	0.244	0.298	0.303	0.266	0.514	0.512	0.557	0.657	0.650	0.540
1.6	0.365	0.392	0.355	0.359	0.400	0.358	0.793	0.808	0.752	0.854	0.871	0.753
1.8	0.430	0.513	0.443	0.415	0.501	0.425	0.947	0.955	0.894	0.965	0.976	0.885
2.0	0.514	0.606	0.496	0.501	0.609	0.505	0.993	0.995	0.965	0.992	0.994	0.961
2.2	0.648	0.695	0.627	0.603	0.698	0.591	1.000	1.000	0.980	0.998	0.999	0.992
2.4	0.755	0.796	0.722	0.713	0.784	0.703	0.999	1.000	0.995	1.000	1.000	0.997
2.6	0.795	0.887	0.776	0.801	0.855	0.790	1.000	1.000	0.999	1.000	1.000	0.999
2.8	0.863	0.917	0.832	0.868	0.905	0.851	1.000	1.000	1.000	1.000	1.000	1.000

Note: I, II, and III refer to three setups with respect to observation processes.

Note that we took the weight function $W(\cdot)$ as 1 in our simulations and applications for simplicity. Theoretically, the weight function would be chosen by minimizing the variance of the estimator. However the selection problem of weight functions is complicated (Lin & Ying, 2001), and it seems impossible to find the optimal weight based on our flexible models where the dependence structure on the underlying longitudinal process is completely unspecified. Further, a more efficient inference procedure needs to be developed.

Also note that covariates were assumed to be time-invariant. When longitudinal data involve both time-independent and time-dependent covariates we propose the following joint models for the underlying recurrent event process and observation process

$$E\{Y_i(t)|X_i(t), Z_i\} = \mu_0(t) + X_i'(t)\beta(t) + g(Z_i),$$

and

$$E\{dO_i(t)|X_i(t), Z_i\} = Z_i h(X_i(t))d\Lambda_0(t),$$

and extend the proposed approach to these situations for future research.

APPENDIX

Proofs. In this appendix we will sketch the proofs for the consistency and asymptotic normality of the proposed estimate $\hat{\beta}_1$ and also for the asymptotic properties of the goodness-of-fit test statistic $\phi(t, x)$. For this we will employ the notation defined in the previous sections and assume that $P(C_i \geq \tau) > 0$. Define $\Gamma = E\{W_i m_i X_{1i} X_{1i}'\}$ and assume that Γ is positive definite.

A.1 Asymptotic Normality of $\hat{\beta}_1$

First we will show the consistency of $\hat{\beta}_1$. For this, note the two facts:

(i)

$$\frac{U(\beta_{10})}{n} = \frac{\sum \phi_i}{n} \xrightarrow{p} 0$$

where $\phi_i = W_i X_{1i}\{\bar{Y}_i - m_i \beta_{10}' X_{1i}\}$;

(ii)

$$\frac{1}{n}\frac{\partial}{\partial \beta_1} U(\beta_1) = -\frac{1}{n}\sum_{i=1}^{n} W_i m_i X_{1i} X_{1i}'$$

converges uniformly to a negative matrix $-E\{W_i m_i X_{1i} X_{1i}'\}$ over β_1 for any value of β_{10}.

Therefore, the solution $\hat{\beta}_1$ of the estimating equation $U(\beta_1) = 0$ is unique with a large n. Note that $n^{-1}\sum_{i=1}^{n} W_i X_{1i}\bar{Y}_i$ converges in probability to $E(W_i X_{1i}\bar{Y}_i) = \Gamma\beta_{10}$ by (5). Then it follows from the expression of $\hat{\beta}_1$ in (7) that $\hat{\beta}_1$ is a consistent estimator of β_{10}.

Now we turn to prove the asymptotic normality of the proposed estimator $\hat{\beta}_1$. For this, note that by the Taylor series expansion we have

$$\begin{aligned}\sqrt{n}(\hat{\beta}_{1n} - \beta_{10}) &= \frac{\Gamma^{-1}}{\sqrt{n}} U(\beta_{10}) + o_p(1) \\ &= \frac{\Gamma^{-1}}{\sqrt{n}} \sum_{i=1}^{n} \phi_i + o_p(1)\end{aligned}$$

where $\phi_i = W_i X_{1i}\{\bar{Y}_i - m_i \beta_{10}' X_{1i}\}$. It thus follows that $\sqrt{n}(\hat{\beta}_{1n} - \beta_{10})$ has an asymptotic normal distribution with mean zero and covariance matrix $\Gamma^{-1}\Sigma(\Gamma^{-1})'$ that can be consistently estimated by $\hat{\Gamma}^{-1}\hat{\Sigma}(\hat{\Gamma}^{-1})'$ where $\Sigma = E\{\phi_i \phi_i'\}$, and $\hat{\Gamma}$ and $\hat{\Sigma}$ are given as

$$\hat{\Gamma} = \frac{1}{n}\sum_{i=1}^{n} W_i m_i X_{1i} X_{1i}' \quad \text{and} \quad \hat{\Sigma} = \frac{1}{n}\sum_{i=1}^{n} \hat{\phi}_i \hat{\phi}_i'$$

with $\hat{\phi}_i = W_i X_{1i}\{\bar{Y}_i - m_i X_{1i}' \hat{\beta}_1\}$.

A.2 Proof of the Asymptotic Property of $\phi(t, x)$

In the following we will sketch the proof for the weak convergence of $\phi(t, x)$ under models (1) and (2); the weak convergence of $\mathcal{F}_j(x)$ can be derived similarly.

Assume that the limits of S_0, $S(x)$, and $B(t, x)$ exist and are denoted by s_0, $s(x)$, and $b(t, x)$, respectively. Define

$$R_i(t) = \int_0^t \{Y_i(u) - \beta_0' X_i\} d\tilde{O}_i(u) - m_i \mathcal{A}(t).$$

To prove the weak convergence of $\phi(t, x)$, using Lemma A.1 of Lin & Ying (2001) we have

$$\phi(t, x) = \frac{1}{\sqrt{n}}\sum_{i=1}^{n}\left\{ I(X_i \le x) - \frac{s(x)}{s_0}\right\} R_i(t) - b(t, x)\sqrt{n}(\hat{\beta} - \beta_0) + o_p(1).$$

The tightness of the first term on the right-hand side of the above follows from the arguments given in Appendix A.5 of Lin et al. (2000). The second term is also tight because $\sqrt{n}(\hat{\beta} - \beta_0)$ converges in distribution and $b(t, x)$ is a deterministic function. Thus $\phi(t, x)$ is tight. Let d_i be the vector $\Gamma^{-1}\phi_i$ without the last entry. Then we can further write $\phi(t, x)$ as

$$\phi(t, x) = \frac{1}{\sqrt{n}}\sum_{i=1}^{n}\left\{ I(X_i \le x) - \frac{s(x)}{s_0}\right\} R_i(t) - b(t, x)\frac{1}{\sqrt{n}}\sum_{i=1}^{n} d_i + o_p(1).$$

It thus follows from the multivariate central limit theorem and the tightness of $\phi(t, x)$ that $\phi(t, x)$ converges weakly to a zero-mean Gaussian process that can be approximated by the zero-mean Gaussian process

$$\hat{\phi}(t, x) = \frac{1}{\sqrt{n}}\sum_{i=1}^{n}\left\{ I(X_i \le x) - \frac{S(x)}{S_0}\right\} \hat{R}_i(t) - B(t, x)\frac{1}{\sqrt{n}}\sum_{i=1}^{n} \hat{d}_i.$$

Thus, using the simulation approach presented in Lin et al. (2000), the null distribution of $\phi(t, x)$ can be approximated by that of $\hat{\phi}(t, x)$. □

ACKNOWLEDGEMENTS

The authors would like to thank the Editor-in-Chief, Dr. David A. Stephens, the Associate Editor, and the two reviewers for their constructive and insightful comments and suggestions that greatly improved the paper. Zhao's research is partly supported by the Research Grant Council of Hong Kong (504011, 503513), The Hong Kong Polytechnic University, and the National Natural Science Foundation of China (No. 11371299).

BIBLIOGRAPHY

Andrews, D. F. & Herzberg, A. M. (1985). *Data: A Collection of Problems from Many Fields for the Student and Research Worker*, Springer-Verlag, New York.

Byar, D. P. (1980). The Veterans Administration study of chemoprophylaxis for recurrent stage I bladder tumors: Comparisons of placebo, pyridoxine, and topical thiotepa. In *Bladder Tumors and Other Topics in Urological Oncology*, Pavane-Macaluso, M., Smith P. H., & Edsmyr F., editors. Plenum, New York, pp. 363–370.

Diggle, P. J., Liang, K. Y., & Zeger, S. L. (1994). *The Analysis of Longitudinal Data*, Oxford University Press, Oxford.

Lawless, J. F. & Nadeau, C. (1995). Some simple robust methods for the analysis of recurrent events. *Technometrics*, 37, 158–168.

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

Lin, D. Y. & Ying, Z. (2001). Semiparametric and nonparametric regression analysis of longitudinal data. *Journal of the American Statistical Association*, 96, 103–126.

Lin, D. Y., Wei, L. J., Yang, I., & Ying, Z. (2000). Semiparametric regression for the mean and rate function of recurrent events. *Journal of the Royal Statistical Society, Series B*, 69, 711–730.

Lin, H., Scharfstein, D. O., & Rosenheck, R. A. (2004). Analysis of longitudinal data with irregular, outcome-dependent follow-up. *Journal of the Royal Statistical Society, Series B*, 66, 791–813.

Lipsitz, S. R., Fitzmaurice, G. M., Ibrahim, J. G., Gelber, R., & Lipshultz, S. (2002). Parameter estimation in longitudinal studies with outcome-dependent follow-up. *Biometrics*, 58, 621–630.

Liu, L., Huang, X., & O'Quigley, J. (2008). Analysis of longitudinal data in the presence of information observational times and a dependent terminal event, with application to medical cost data. *Biometrics*, 64, 950–958.

Song, X., Mu, X., & Sun, L. (2012). Regression analysis of longitudinal data with time-dependent covariates and informative observation times. *Scandinavian Journal of Statistics*, 39, 248–258.

Sun, J., Park, D., Sun, L., & Zhao, X. (2005). Semiparametric regression analysis of longitudinal data with informative observation times. *Journal of the American Statistical Association*, 100, 882–889.

Sun, J., Sun, L., & Liu, D. (2007). Regression analysis of longitudinal data in the presence of informative observation and censoring times. *Journal of the American Statistical Association*, 102, 1397–1406.

Sun, J. & Wei, L. J. (2000). Regression analysis of panel count data with covariate-dependent observation and censoring times. *Journal of the Royal Statistical Society, Series B*, 62, 293–302.

Wang, M. C., Qin, J., & Chiang, C. T. (2001). Analyzing recurrent event data with informative censoring. *Journal of the American Statistical Association*, 96, 1057–1065.

Welsh, A. H., Lin, X., & Carroll, R. J. (2002). Marginal longitudinal nonparametric regression: Locality and efficiency of spline and kernel Methods. *Journal of the American Statistical Association*, 97, 482–493.

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

Zhao, X., Tong, X., & Sun, L. (2012). Joint analysis of longitudinal data with dependent observation times. *Statistica Sinica*, 22, 317–336.

Zhu, L., Sun, J., Tong, X., & Pounds, S. (2011). Regression analysis of longitudinal data with informative observation times and application to medical cost data . *Statistics in Medicine*, 30, 1429–1440.

Received 27 September 2014
Accepted 10 August 2015