Biometrics 67, 404–414
June 2011
DOI: 10.1111/j.1541-0420.2010.01458.x

Semiparametric Transformation Models with Time-Varying Coefficients for Recurrent and Terminal Events

Xingqiu Zhao,[1,*] Jie Zhou,[2,] and Liuquan Sun[2,***]**

[1]Department of Applied Mathematics, The Hong Kong Polytechnic University, Hong Kong, China
[2]Institute of Applied Mathematics, Chinese Academy of Sciences, Beijing 100190, China
[*]*email:* xingqiu.zhao@polyu.edu.hk
[**]*email:* zhoujie@amss.ac.cn
[***]*email:* slq@amt.ac.cn

Summary. In this article, we propose a family of semiparametric transformation models with time-varying coefficients for recurrent event data in the presence of a terminal event such as death. The new model offers great flexibility in formulating the effects of covariates on the mean functions of the recurrent events among survivors at a given time. For the inference on the proposed models, a class of estimating equations is developed and asymptotic properties of the resulting estimators are established. In addition, a lack-of-fit test is provided for assessing the adequacy of the model, and some tests are presented for investigating whether or not covariate effects vary with time. The finite-sample behavior of the proposed methods is examined through Monte Carlo simulation studies, and an application to a bladder cancer study is also illustrated.

Key words: Counting process; Estimating equation; Marginal model; Model checking; Recurrent events; Terminal event; Time-varying coefficients.

1. Introduction

In many longitudinal studies, the event of interest can be experienced more than once per subject. Medical examples of recurrent events include, among others, multiple infection episodes and tumor recurrences. Other examples include repeated breakdowns of a certain machinery in reliability testing and repeated purchases of a certain product in marketing research. The investigators are often interested in assessing the effects of covariates on certain features of the recurrent event process. In many situations, the follow-up of recurrent event could be stopped by a terminal event such as death, which precludes further recurrent events. For example, patients may experience recurrent hospitalizations that are terminated by death.

For recurrent event data in the absence of a terminal event, there are several estimating procedures proposed in the literature, including conditional models (Prentice, Williams, and Peterson, 1981; Andersen and Gill, 1982; Chang and Wang, 1999; Zeng and Lin, 2006), marginal intensity models (Wei, Lin, and Weissfeld, 1989; Lee, Wei, and Amato, 1992), marginal mean and rate models (Pepe and Cai, 1993; Lawless and Nadeau, 1995; Lin et al., 2000), and the frailty model approach (Nielsen et al., 1992; Murphy, 1994, 1995; Zeng and Lin, 2007). To allow nonproportional means and rates, some other semiparametric regression methods have been studied. For example, Lin, Wei, and Ying (1998) proposed an accelerated failure time model to formulate the effects of covariates on the mean function of the counting process. Lin, Wei, and Ying (2001) suggested a class of semiparametric transformation models for point processes with positive jumps of arbitrary sizes. Ghosh (2004) presented the accelerated rate model and Schaubel, Zeng, and Cai (2006) proposed a semiparametric additive rate model for counting processes. Cook and Lawless (2007) provided an excellent review of methods for recurrent event data.

Some efforts have been made recently on the analysis of recurrent events in the presence of a terminal event (e.g., Cook and Lawless, 1997; Ghosh and Lin, 2000, 2002, 2003; Huang and Wang, 2004; Liu, Wolfe, and Huang, 2004; Schaubel and Cai, 2005; Ye, Kalbfleish, and Schaubel, 2007). For example, Cook and Lawless (1997) and Ghosh and Lin (2002) studied the mean and rate functions of recurrent events among survivors at a given time. Ghosh and Lin (2003) presented an accelerated failure time model for recurrent events. Huang and Wang (2004) discussed joint models of the recurrent and terminal events, where the inference was focused on the frequency of recurrent events at the failure time of the terminal event. Ye et al. (2007) proposed a joint semiparametric model in which the correlation between the recurrent and terminal events is incorporated through the frailty.

The aforementioned semiparametric regression models assume that regression coefficients are constant over time. In reality, however, the regression parameters may vary over time, and it is important to know the temporal effects of the covariates on the recurrent event times. The time-varying coefficient models provide a nice graphical summary of time dynamics of covariates and further allow for inference about covariate effects. In addition, making inference on recurrent event processes for subjects who are currently alive is of interest in many studies, and some authors have studied the conditional

recurrent event rate given survival (e.g., Cook and Lawless, 1997; Liu et al., 2004; Schaubel and Cai, 2005; Ye et al., 2007; Pan and Schaubel, 2009). In this article, we propose a family of semiparametric transformation models with time-varying coefficients for the mean functions of the recurrent events among survivors at a given time. The new model offers great flexibility in formulating the effects of covariates on the conditional mean functions while leaving the stochastic structure completely unspecified. For the sake of exposition, we will focus on time-invariant covariates and address the time-dependent covariate issue in the discussion.

The remainder of the article is organized as follows. In Section 2, we introduce relevant notation and formulate the model. By using the technique of inverse probability weighting after adjusting possibly time-dependent covariates (Ghosh and Lin, 2002), an estimating procedure is proposed for the model parameters. Section 3 studies the asymptotic properties of the resulting estimators. In Section 4, we develop a technique for checking the adequacy of the general model, and provide some tests for investigating whether or not covariate effects vary with time. Section 5 reports some results from simulation studies conducted for evaluating the proposed methods. In Section 6, we apply the methodology to a data set from a bladder cancer study, and some concluding remarks are made in Section 7.

2. Model and Estimation Procedures

Let $N^*(t)$ denote the number of recurrent events that occur over the interval $[0, t]$, and X and Z be vectors of covariates of dimensions p and q, respectively. In most applications, the subject is followed for a limited period of time so that $N^*(\cdot)$ may not be fully observed. Let C denote the follow-up or censoring time, and let D be the time of the terminating event, where the terminal event may stop the further occurrence of recurrent events in that $N^*(t)$ is constant after D. It is assumed that $N^*(\cdot)$ is independent of C conditional on X and Z. Note that $N^*(\cdot)$ can only be observed up to C and that in general only the minimum of D and C is known. Write $T = D \wedge C, \Delta = I(D \leqslant C)$, and $N(t) = N^*(t \wedge C)$, where $a \wedge b = \min(a, b)$ and $I(\cdot)$ is the indicator function.

Let $E\{N^*(t) \,|\, X, Z, D \geqslant t\}$ denote the mean function of the recurrent events conditional on the terminal event not occurring before t. Our proposed transformation models with time-varying coefficients take the form

$$E\{N^*(t) \,|\, X, Z, D \geqslant t\} = g\{\beta_0(t)'X + \gamma_0' Z\}, \tag{1}$$

where $g(\cdot) \geqslant 0$ is pre-specified and assumed to be twice continuously differentiable, $\beta_0(t)$ is a p-dimensional vector of unknown time-varying regression coefficients, and γ_0 is a q-dimensional vector of unknown time-independent regression parameters. Model (1) is marginal in the sense that it does not condition on the event history. Here we allow the first component of X to be 1, which gives a baseline mean function. When there is no dependent terminal event and $X \equiv 1$, model (1) reduces to the transformation models studied by Lin et al. (2001). Model (1) defines a very rich family of models through the link function $g(\cdot)$. The exponential link function $g(x) = \exp(x)$ is an obvious choice. It also encompasses the Box–Cox transformations, in which $g(\cdot)$ is given by $g(x) = \{(x+1)^\rho - 1\}/\rho (\rho \geq 0)$, where $\rho = 0$ means that $g(x) = \log(x+1)$. Another useful class is the logarithmic transformations, which are given by $g(x) = \log(1 + rx)/r (r \geqslant 0)$, where $r = 0$ means that $g(x) = x$.

For a random sample of n subjects, the data consist of $\{N_i(t), T_i, \Delta_i, X_i, Z_i; t \leqslant T_i, i = 1, \ldots, n\}$. We specify the proportional hazards model for the time of the terminating event D_i as

$$\lambda(t \,|\, W_i) = \lambda_0(t) \exp(\alpha_0' W_i), \tag{2}$$

where $\lambda_0(t)$ is an unspecified baseline hazard function, W_i is a r-dimensional covariate vector that is a part of $(X_i', Z_i')'$, and α_0 is an r-dimensional vector of unknown parameters. It is assumed that D_i and C_i are conditionally independent given W_i. We denote the survival function of D_i as $S(t \,|\, W_i) = P\{D_i \geqslant t \,|\, W_i\}$. Define $Y_i(t) = I(T_i \geqslant t)$, and

$$M_i(t) = \frac{Y_i(t)}{S(t \,|\, W_i)}[N_i(t) - g\{\beta_0(t)'X_i + \gamma_0' Z_i\}], \quad i = 1, \ldots, n.$$

Under model (1), $M_i(t)$'s are zero-mean stochastic processes.

In practice, the survival function $S(t \,|\, W_i)$ is unknown, but it can be estimated by the fit of model (2). Specifically, let $\hat{\alpha}$ be the maximum partial likelihood estimator of α_0, which is defined as the solution to

$$U_\alpha(\alpha) = \sum_{i=1}^n \int_0^\tau \{W_i - \bar{W}(t; \alpha)\} \, dN_i^D(t) = 0, \tag{3}$$

where $N_i^D(t) = I(T_i \leqslant t, \Delta_i = 1), \bar{W}(t; \alpha) = S^{(1)}(t; \alpha)/S^{(0)}(t; \alpha)$,

$$S^{(k)}(t; \alpha) = n^{-1} \sum_{i=1}^n Y_i(t) W_i^{\otimes k} \exp(\alpha' W_i) \quad \text{for } k = 0, 1, 2,$$

and, for a vector $v, v^{\otimes 0} = 1, v^{\otimes 1} = v$ and $v^{\otimes 2} = vv'$. Let $\hat{\Lambda}_0(t)$ be the Breslow estimator of $\Lambda_0(t) = \int_0^t \lambda_0(u) \, du$, that is,

$$\hat{\Lambda}_0(t) = \sum_{i=1}^n \int_0^t \frac{dN_i^D(u)}{\sum_{j=1}^n Y_j(u) \exp(\hat{\alpha}' W_j)} \, .$$

Then $S(t \,|\, W_i)$ can be estimated by $\hat{S}(t \,|\, W_i) = \exp\{-\exp(\hat{\alpha}' W_i)\hat{\Lambda}_0(t)\}$. To estimate the regression coefficient function $\beta(t)$ and the regression parameter γ, using the generalized estimating equation approach (Liang and Zeger, 1986) and the inverse probability weighting technique, we propose the following two estimating functions for $\beta_0(t)$ and γ_0:

$$U_1(t, \beta, \gamma) = \sum_{i=1}^n \frac{Y_i(t)}{\hat{S}(t \,|\, W_i)} X_i \times [N_i(t) - g\{\beta(t)'X_i + \gamma' Z_i\}], \quad 0 \leqslant t \leqslant \tau, \tag{4}$$

and

$$U_2(\tau, \beta, \gamma) = \sum_{i=1}^n \int_0^\tau \frac{Y_i(t)}{\hat{S}(t \,|\, W_i)} Z_i \times [N_i(t) - g\{\beta(t)'X_i + \gamma' Z_i\}] \, dH(t), \tag{5}$$

where τ is a pre-specified constant such that $P(T_i \geqslant \tau) > 0$, and $H(t)$ is an increasing and known weight function on $[0, \tau]$. Let $\hat{\beta}(t)$ and $\hat{\gamma}$ denote the solutions to $U_1(t, \beta, \gamma) = 0$ and

$U_2(\tau, \beta, \gamma) = 0$. The estimate $\hat{\beta}(t)$ will be a piecewise constant function with jumps only at the observed event times.

Let $\dot{g}(x) = dg(x)/dx$. To solve the estimating equations (4) and (5) simultaneously, we proceed by a Taylor expansion of $g\{\beta(t)'X_i + \gamma' Z_i\}$ around the current value of estimates $\beta^{(k)}(t)$ and $\gamma^{(k)}$ to get approximated estimating equations

$$\begin{aligned} U_1(t, \beta, \gamma) \approx & \sum_{i=1}^{n} \frac{Y_i(t)}{\hat{S}(t \mid W_i)} X_i \\ & \times \big[N_i(t) - g\big\{ \beta^{(k)}(t)' X_i + \gamma^{(k)'} Z_i \big\} \\ & \quad - \dot{g}\big\{ \beta^{(k)}(t)' X_i + \gamma^{(k)'} Z_i \big\} X_i' \{\beta(t) - \beta^{(k)}(t)\} \\ & \quad - \dot{g}\big\{ \beta^{(k)}(t)' X_i + \gamma^{(k)'} Z_i \big\} Z_i' \{\gamma - \gamma^{(k)}\} \big], \quad (6) \end{aligned}$$

and

$$\begin{aligned} U_2(\tau, \beta, \gamma) \approx & \sum_{i=1}^{n} \int_0^{\tau} \frac{Y_i(t)}{\hat{S}(t \mid W_i)} Z_i \\ & \times \big[N_i(t) - g\big\{ \beta^{(k)}(t)' X_i + \gamma^{(k)'} Z_i \big\} \\ & \quad - \dot{g}\big\{ \beta^{(k)}(t)' X_i + \gamma^{(k)'} Z_i \big\} X_i' \{\beta(t) - \beta^{(k)}(t)\} \\ & \quad - \dot{g}\big\{ \beta^{(k)}(t)' X_i + \gamma^{(k)'} Z_i \big\} Z_i' \{\gamma - \gamma^{(k)}\} \big] \, dH(t). \quad (7) \end{aligned}$$

Define

$$E_{xx}^{(k)}(t) = n^{-1} \sum_{i=1}^{n} \frac{Y_i(t)}{\hat{S}(t \mid W_i)} \dot{g}\big\{ \beta^{(k)}(t)' X_i + \gamma^{(k)'} Z_i \big\} X_i X_i',$$

$$E_{zz}^{(k)}(t) = n^{-1} \sum_{i=1}^{n} \frac{Y_i(t)}{\hat{S}(t \mid W_i)} \dot{g}\big\{ \beta^{(k)}(t)' X_i + \gamma^{(k)'} Z_i \big\} Z_i Z_i',$$

$$E_{zx}^{(k)}(t) = n^{-1} \sum_{i=1}^{n} \frac{Y_i(t)}{\hat{S}(t \mid W_i)} \dot{g}\big\{ \beta^{(k)}(t)' X_i + \gamma^{(k)'} Z_i \big\} Z_i X_i',$$

and $E_{xz}^{(k)}(t) = E_{zx}^{(k)}(t)'$. Solving (6)for $\beta(t)$ and inserting it into (7), we get the $(k+1)$th iterative estimator for γ_0:

$$\begin{aligned} \gamma^{(k+1)} = & \gamma^{(k)} + n^{-1} (A^{(k)})^{-1} \\ & \times \sum_{i=1}^{n} \int_0^{\tau} \frac{Y_i(t)}{\hat{S}(t \mid W_i)} \big\{ Z_i - E_{zx}^{(k)}(t) E_{xx}^{(k)}(t)^{-1} X_i \big\} \\ & \times \big[N_i(t) - g\big\{ \beta^{(k)}(t)' X_i + \gamma^{(k)'} Z_i \big\} \big] \, dH(t), \quad (8) \end{aligned}$$

where

$$A^{(k)} = \int_0^{\tau} \big[E_{zz}^{(k)}(t) - E_{zx}^{(k)}(t) E_{xx}^{(k)}(t)^{-1} E_{xz}^{(k)}(t) \big] \, dH(t).$$

Using the updated version $\gamma^{(k+1)}$ and solving (6)for $\beta(t)$, we have

$$\begin{aligned} \beta^{(k+1)}(t) = & \beta^{(k)}(t) + n^{-1} E_{xx}^{(k)}(t)^{-1} \sum_{i=1}^{n} \frac{Y_i(t)}{\hat{S}(t \mid W_i)} X_i \\ & \times \big[N_i(t) - g\big\{ \beta^{(k)}(t)' X_i + \gamma^{(k)'} Z_i \big\} \\ & \quad - \dot{g}\big\{ \beta^{(k)}(t)' X_i + \gamma^{(k)'} Z_i \big\} Z_i' \{\gamma^{(k+1)} - \gamma^{(k)}\} \big]. \quad (9) \end{aligned}$$

This iteration is continued until convergence and the estimates $\hat{\beta}(t)$ and $\hat{\gamma}$ are obtained at convergence. For the convergence, also several criteria can be applied and in the numerical studies below, we used the absolute differences between the iterative estimates of the parameters. The algorithm converges most times in general, but nonconvergence could occur occasionally depending on the setups.

3. Asymptotic Results

In this section, we establish the asymptotic properties of the estimates given in the previous section. First we consider the existence, uniqueness, and strong consistency of $\hat{\beta}(t)$ and $\hat{\gamma}$. The results are summarized in the following theorem with the proof given in the Web Appendix.

THEOREM 1: *Under the regularity conditions (C1)–(C5) stated in the Web Appendix, $\hat{\beta}(t)$ and $\hat{\gamma}$ exist and are unique. Moreover, $\hat{\gamma}$ is strongly consistent to γ_0, and $\hat{\beta}(t) \to \beta_0(t)$ almost surely uniformly in $t \in [0, \tau]$.*

Let $\hat{E}_{xx}(t), \hat{E}_{zz}(t), \hat{E}_{xz}(t), \hat{E}_{zx}(t)$, and $\hat{A}$ denote the quantities defined in the previous section with all unknown parameters replaced by their estimates. Define

$$\hat{M}_i(t) = \frac{Y_i(t)}{\hat{S}(t \mid W_i)} [N_i(t) - g\{\hat{\beta}(t)' X_i + \hat{\gamma}' Z_i\}].$$

The asymptotic distributions of $\hat{\beta}(t)$ and $\hat{\gamma}$ are given in the next theorem.

THEOREM 2: *Under the regularity conditions (C1)–(C5) stated in the Web Appendix, we have*

(*i*) *$n^{1/2}(\hat{\gamma} - \gamma_0)$ is asymptotically normal with mean zero and a variance–covariance matrix that can be consistently estimated by $\hat{A}^{-1} \hat{\Sigma} \hat{A}^{-1}$, where $\hat{\Sigma} = n^{-1} \sum_{i=1}^{n} \hat{\xi}_i^{\otimes 2}$,*

$$\begin{aligned} \hat{\xi}_i = & \int_0^{\tau} \hat{M}_i(t) \big[Z_i - \hat{E}_{zx}(t) \hat{E}_{xx}(t)^{-1} X_i \big] \, dH(t) \\ & + \int_0^{\tau} \frac{\hat{Q}(t)}{S^{(0)}(t; \hat{\alpha})} \, d\hat{M}_i^D(t) \\ & + \hat{B} \hat{\Omega}^{-1} \int_0^{\tau} \{W_i - \bar{W}(t; \hat{\alpha})\} \, d\hat{M}_i^D(t), \end{aligned}$$

$$\begin{aligned} \hat{Q}(t) = & n^{-1} \sum_{i=1}^{n} \int_t^{\tau} \exp(\hat{\alpha}' W_i) \hat{M}_i(u) \\ & \times \big\{ Z_i - \hat{E}_{zx}(u) \hat{E}_{xx}(u)^{-1} X_i \big\} \, dH(u), \end{aligned}$$

$$\begin{aligned} \hat{B} = & n^{-1} \sum_{i=1}^{n} \int_0^{\tau} \hat{M}_i(t) \\ & \times \big\{ Z_i - \hat{E}_{zx}(t) \hat{E}_{xx}(t)^{-1} X_i \big\} \hat{V}_i(t)' \, dH(t), \end{aligned}$$

$$\hat{V}_i(t) = \int_0^t \exp(\hat{\alpha}' W_i) \{W_i - \bar{W}(u; \hat{\alpha})\} \, d\hat{\Lambda}_0(u),$$

and

$$\hat{\Omega} = n^{-1} \sum_{i=1}^{n} \left\{ \frac{S^{(2)}(t; \hat{\alpha})}{S^{(0)}(t; \hat{\alpha})} - \bar{W}(t; \hat{\alpha})^{\otimes 2} \right\} dN_i^D(t).$$

(*ii*) $n^{1/2}\{\hat{\beta}(t)-\beta_0(t)\}$ $(0 \leqslant t \leqslant \tau)$ *converges weakly to a zero-mean Gaussian process whose covariance function at* (s,t) *can be consistently estimated by*

$$\hat{\Gamma}(s,t) = n^{-1}\sum_{i=1}^{n}\hat{\phi}_i(s)\hat{\phi}_i(t)',$$

where

$$\hat{\phi}_i(t) = \hat{E}_{xx}(t)^{-1}\Bigg[\int_0^t \frac{\hat{R}(t)}{S^{(0)}(u;\hat{\alpha})}\, d\hat{M}_i^D(u) + \hat{P}(t)\hat{\Omega}^{-1}\int_0^\tau \left\{W_i - \bar{W}(u;\hat{\alpha})\right\} d\hat{M}_i^D(u) + X_i\hat{M}_i(t) - \hat{E}_{xz}(t)\hat{A}^{-1}\hat{\xi}_i\Bigg], \tag{10}$$

$$\hat{R}(t) = n^{-1}\sum_{i=1}^{n}\exp(\hat{\alpha}'W_i)\hat{M}_i(t)X_i,$$

and

$$\hat{P}(t) = n^{-1}\sum_{i=1}^{n}M_i(t)X_i\hat{V}_i(t)'.$$

The asymptotic normality for $\hat{\beta}(t)$, together with the consistent covariance estimator $\hat{\Gamma}$, enables us to construct pointwise confidence intervals for $\beta_0(t)$. To construct simultaneous confidence bands for $\beta_0(t)$ over a time interval $[a,b]$ of interest, where $0 < a < b \leqslant \tau$, we need to evaluate the distribution of the supremum of a related process over $[a,b]$. It is not possible to evaluate such distributions analytically because the limiting process of $n^{1/2}\{\hat{\beta}(t)-\beta_0(t)\}$ does not have an independent increment structure. To handle this problem, we can use a resampling scheme to approximate the distribution of $n^{1/2}\{\hat{\beta}(t)-\beta_0(t)\}$. Define $\hat{\Xi}(t) = n^{-1/2}\sum_{i=1}^{n}\hat{\phi}_i(t)G_i$, where $(G_1,\ldots,G_n)$ are independent standard normal variables that are independent of the data $\{N_i(t), T_i, \Delta_i, X_i, Z_i;\ t \leqslant T_i,\ i = 1,\ldots,n\}$. According to the arguments of Lin et al. (2000), the distribution of the process $n^{1/2}\{\hat{\beta}(t)-\beta_0(t)\}$ can be approximated by that of the zero-mean Gaussian process $\hat{\Xi}(t)$. To approximate the distributions of $n^{1/2}\{\hat{\beta}(t)-\beta_0(t)\}$, we need to obtain a large number of realizations from $\hat{\Xi}(t)$ by repeatedly generating the normal random sample $(G_1,\ldots,G_n)$ while fixing the data $\{N_i(t), T_i, \Delta_i, X_i, Z_i; t \leqslant T_i, i = 1,\ldots,n\}$ at their observed values. Using this simulation method, we can determine an approximate $1-\alpha$ simultaneous confidence band for $\beta_0(t)$ over a time interval $[a,b]$ of interest.

4. Model Checking

In this section, we propose a formal lack-of-fit test for assessing the adequacy of model (1). Following Lin, Wei, and Ying (1993), we consider the following cumulative sums of residuals:

$$\mathcal{F}(t,x,z) = n^{-1/2}\sum_{i=1}^{n} I(X_i \leqslant x, Z_i \leqslant z)\hat{M}_i(t),$$

where $I(X_i \leqslant x, Z_i \leqslant z)$ means that each component of X_i and Z_i is no larger than the corresponding component of x and z (e.g., Lin et al., 2000). We show in the Web Appendix that the null distribution of $\mathcal{F}(t,x,z)$ can be approximated by the zero-mean Gaussian process

$$\tilde{\mathcal{F}}(t,x,z) = n^{-1/2}\sum_{i=1}^{n}\Bigg[\int_0^t \frac{\hat{R}^*(t,x,z)}{S^{(0)}(u;\hat{\alpha})}\, d\hat{M}_i^D(u) + \hat{\Upsilon}_1(t,x,z)'\hat{\Omega}^{-1} \times \int_0^\tau \{W_i - \bar{W}(u;\hat{\alpha})\}\, d\hat{M}_i^D(u) + I(X_i \leqslant x, Z_i \leqslant z)\hat{M}_i(t) - \hat{\Upsilon}_2(t,x,z)'\hat{\phi}_i(t) - \hat{\Upsilon}_3(t,x,z)'\hat{A}^{-1}\hat{\xi}_i\Bigg], \tag{11}$$

where

$$\hat{R}^*(t,x,z) = n^{-1}\sum_{i=1}^{n}\exp(\hat{\alpha}'W_i)I(X_i \leqslant x, Z_i \leqslant z)\hat{M}_i(t),$$

$$\hat{\Upsilon}_1(t,x,z) = n^{-1}\sum_{i=1}^{n} I(X_i \leqslant x, Z_i \leqslant z)M_i(t)V_i(t),$$

$$\hat{\Upsilon}_2(t,x,z) = n^{-1}\sum_{i=1}^{n}\frac{Y_i(t)}{\hat{S}(t\,|\,W_i)}I(X_i \leqslant x, Z_i \leqslant z)g \times \{\hat{\beta}(t)'X_i + \hat{\gamma}'Z_i)\}X_i,$$

and

$$\hat{\Upsilon}_3(t,x,z) = n^{-1}\sum_{i=1}^{n}\frac{Y_i(t)}{\hat{S}(t\,|\,W_i)}I(X_i \leqslant x, Z_i \leqslant z)g \times \{\hat{\beta}(t)'X_i + \hat{\gamma}'Z_i)\}Z_i.$$

As in the case of $\hat{\beta}(t)$, it is difficult to estimate the asymptotic covariance function of $\mathcal{F}(t,x,z)$ analytically. We again appeal to the resampling approach and show that the null distribution of $\mathcal{F}(t,x,z)$ can be approximated by the conditional distribution of $\hat{\mathcal{F}}(t,x,z)$, where

$$\hat{\mathcal{F}}(t,x,z) = n^{-1/2}\sum_{i=1}^{n}\Bigg[\int_0^t \frac{\hat{R}^*(t,x,z)}{S^{(0)}(u;\hat{\alpha})}\, d\hat{M}_i^D(u) + \hat{\Upsilon}_1(t,x,z)'\hat{\Omega}^{-1} \times \int_0^\tau \left\{W_i - \bar{W}(u;\hat{\alpha})\right\} d\hat{M}_i^D(u) + I(X_i \leqslant x, Z_i \leqslant z)\hat{M}_i(t) - \hat{\Upsilon}_2(t,x,z)'\hat{\phi}_i(t) - \hat{\Upsilon}_3(t,x,z)'\hat{A}^{-1}\hat{\xi}_i\Bigg]G_i.$$

Thus, one can obtain a large number of realizations from $\hat{\mathcal{F}}(t,x,z)$ by repeatedly generating the standard normal random sample $(G_1,\ldots,G_n)$ while fixing the observed data, and plot $\mathcal{F}(t,x,z)$ along with a few realizations of $\hat{\mathcal{F}}(t,x,z)$. Since the validity of approximating $\mathcal{F}(t,x,z)$ by $\hat{\mathcal{F}}(t,x,z)$ depends on the correct specification of model (1), an unusual pattern of $\mathcal{F}(t,x,z)$ compared to the realizations of $\hat{\mathcal{F}}(t,x,z)$ would suggest a lack-of-fit of model (1). Since $\mathcal{F}(t,x,z)$ is expected to fluctuate randomly around 0 under model (1),

a formal lack-of-fit test can be constructed based on the supremum statistic $\sup_{0\leqslant t\leqslant\tau,x,z}|\mathcal{F}(t,x,z)|$, with which the p-value can be obtained by comparing the observed value of $\sup_{0\leqslant t\leqslant\tau,x,z}|\mathcal{F}(t,x,z)|$ to a large number of realizations from $\sup_{0\leqslant t\leqslant\tau,x,z}|\hat{\mathcal{F}}(t,x,z)|$.

In regression analysis, it is usually of interest to test covariate effects and to test if some covariate effects are indeed time-varying. The test of $\gamma_0 = 0$ is apparently straightforward by using the Wald test. To test if a time-varying covariate effect is significant, say $\beta_{0k}(t) \equiv 0$ for the kth element of $\beta_0(t)$, one can consider the statistic

$$\mathcal{F}_k^{(1)} = \sup_{0\leqslant t\leqslant\tau}\left|\frac{\hat{\beta}_k(t)}{\hat{\sigma}_k(t)}\right|,$$

and reject the $\beta_{0k}(t) \equiv 0$ if $\mathcal{F}_k^{(1)}$ is away from zero, where $\hat{\beta}_k(t)$ is the kth elements of $\hat{\beta}(t)$ and $\hat{\sigma}_k(t)$ is an estimate of the standard error of $\hat{\beta}_k(t)$. The percentile of the observed test statistic can be computed by resampling technique as above.

To check if the coefficient, $\beta_{0k}(t)$, is significantly time-varying, we test

$$H_0 : \beta_{0k}(t) \equiv \beta_{0k}$$

for some constant β_{0k}. Let $\Psi_k(t) = \beta_{0k}(t) - \int_0^\tau \beta_{0k}(u)du/\tau$ and $\hat{\Psi}_k(t) = \hat{\beta}_k(t) - \int_0^\tau \hat{\beta}_k(u)\,du/\tau$. Then under the hypothesis H_0, we have $\Psi_k(t) \equiv 0$ and this suggests the following two test statistics. One is the Kolmogorov–Smirnov type statistic defined as

$$\mathcal{F}_k^{(2)} = \sup_{0\leqslant t\leqslant\tau}\left|n^{1/2}\hat{\Psi}_k(t)\right|$$

and the other is the Cramér-von Mises type statistic

$$\mathcal{F}_k^{(3)} = \int_0^\tau n\hat{\Psi}_k(t)^2\,dt.$$

Thus, the null hypothesis H_0 is rejected if both statistics are away from zero.

To obtain the distributions of $\mathcal{F}_k^{(2)}$ and $\mathcal{F}_k^{(3)}$, first we note that it follows from Theorem 2 that the asymptotic distribution of $n^{1/2}\{\hat{\Psi}_k(t) - \Psi_k(t)\}$ is asymptotically equivalent to the zero-mean Gaussian process

$$\hat{\Upsilon}_k(t) = n^{-1/2}\sum_{i=1}^n\left\{\hat{\phi}_{ij}(t) - \int_0^\tau \hat{\phi}_{ij}(u)du/\tau\right\},$$

where $\hat{\phi}_{ik}(t)$ is the kth element of $\hat{\phi}_i(t)$ defined in (10). Similarly, using the resampling technique, we can determine approximate critical values of the two tests.

5. Simulation Studies

We conducted simulation studies to examine the finite sample properties of the proposed estimators. In the study, we first generated a nonhomogeneous Poisson process with the following marginal model:

$$E\{N^{**}(t)\,|\,X,Z,\omega\} = \omega g\{\beta_1(t) + \beta_2(t)X + \gamma Z\}, \quad (12)$$

where X is a Bernoulli random variable with success probability 0.5, Z is a uniform random variable on $(0,1)$, and ω is an independent gamma random variable with mean 1 and variance σ^2. We considered two choices for g: an exponential link function $g_1(x) = 0.3\exp(x)$ and a Box–Cox transformation $g_2(x) = [\{1 + 0.1\exp(x)\}^2 - 1]/1.4$. In the sequel, we set $\beta_1(t) = 0.5 + \log(t), \beta_2(t) = 0.2t, \gamma = 0, 0.5$, or 1, and $\sigma^2 = 0, 0.25, 0.5$, or 1. The censoring time was taken as $C = C_1 \wedge \tau$, with C_1 generated from a uniform distribution $U(0,20)$ and $\tau = 5$ representing the largest follow-up time. For given $W = X$, the terminating event time D was generated from an exponential distribution with hazards $0.05\exp(0.6W)$, which yielded about 74% of subjects censored (i.e., $C \leqslant D$).

The recurrent event process was defined as $N^*(t) = N^{**}(t \wedge D)$. Since D is independent of $N^{**}(\cdot)$ conditional on X and Z, it can be verified that $N^*(t)$ satisfies model (1). Under the preceding settings, the average number of observed events per subject ranged from 2.7 to 6.4 for different model parameters. In all simulations, we considered two choices for the weight functions: $H_1(t) = t$, for which the integral in (5) becomes the area under the curve, and $H_2(t) = n^{-1}\sum_{i=1}^n N_i(t)$, for which the integral in (5) is a weighted sum of the integrand over all the jump points of $\{N_i(t);\ i = 1,\ldots,n\}$. The results presented below are based on 1000 replications with sample sizes $n = 100$ and 200, and final estimates were reached at convergence.

Tables 1 and 2 present the simulation results on the estimate of γ under $g_1(x)$ and $g_2(x)$, respectively. In these tables, Bias is the sample mean of the estimate minus the true value; SE is the sampling standard error of the estimate; SEE is the sampling mean of the standard error estimate of the estimate; and CP is the 95% empirical coverage probability for γ based on a normal approximation. It can be seen from Tables 1 and 2 that the proposed estimation procedures perform well for the situations considered here. It appears that the estimates are unbiased and there is a good agreement between the estimated and empirical standard errors. The empirical coverage probabilities are reasonable and the results become better when the sample size increases from 100 to 200. In addition, for these situations, the results are similar for the two weight functions $H_1(t)$ and $H_2(t)$.

In the same simulation studies as reported in Tables 1 and 2 with $H_1(t)$, we assessed the behaviors of the estimates for time-varying regression coefficients $\beta_1(t)$ and $\beta_2(t)$. Tables 3 and 4 give the simulation results on the estimates of $\beta_1(t)$ and $\beta_2(t)$ at time points $t = 1, 3$, and 5 under $g_1(t)$ and $g_2(t)$, respectively. We also computed pointwise biases and pointwise coverage probabilities for $\beta_1(t)$ and $\beta_2(t)$ with 100 grid points $0 + 0.05k$, $k = 1,\ldots,100$. The pointwise biases and pointwise coverage probabilities for some settings were depicted. Our simulation results suggest that the proposed estimators perform quite well and essentially provide unbiased estimates of the time-varying regression coefficients. The asymptotic standard errors present an excellent description of the variability for the estimates of $\beta_1(t)$ and $\beta_2(t)$, and the estimation procedures are reliable. We also considered other setups and the results were similar to those given above.

6. An Application

In this section, we apply the proposed methods to a set of recurrent event data arising from cancer clinical trial conducted by the Veterans Administration Cooperative Urological Research Group (Byar, 1980). These data have been analyzed extensively in the literature (Ghosh and Lin, 2000,

Table 1
Simulation results for the estimation of γ under $g_1(x)$

			$H_1(t)$				$H_2(t)$			
n	γ	σ^2	Bias	SE	SEE	CP	Bias	SE	SEE	CP
100	0	0.00	−0.0051	0.2410	0.2331	0.936	−0.0009	0.2431	0.2298	0.935
		0.25	0.0090	0.3314	0.3131	0.920	−0.0064	0.3434	0.3123	0.925
		0.50	−0.0276	0.4074	0.3804	0.932	−0.0019	0.4019	0.3793	0.922
		1.00	−0.0053	0.5222	0.4779	0.925	0.0349	0.5290	0.4810	0.923
	0.5	0.00	0.0059	0.2186	0.2049	0.931	0.0068	0.2113	0.2045	0.937
		0.25	−0.0073	0.3170	0.2966	0.927	−0.0073	0.3170	0.2966	0.927
		0.50	0.0040	0.3782	0.3625	0.932	−0.0058	0.3885	0.3626	0.923
		1.00	−0.0002	0.5061	0.4665	0.926	0.0189	0.5055	0.4714	0.924
	1	0.00	0.0047	0.1899	0.1821	0.937	0.0034	0.1855	0.1793	0.926
		0.25	−0.0065	0.2935	0.2827	0.932	0.0101	0.3002	0.2867	0.934
		0.50	−0.0236	0.3732	0.3538	0.932	−0.0043	0.3906	0.3563	0.902
		1.00	0.0087	0.4770	0.4576	0.933	0.0134	0.5208	0.4680	0.915
200	0	0.00	0.0069	0.1705	0.1666	0.947	0.0032	0.1664	0.1645	0.943
		0.25	0.0039	0.2332	0.2254	0.940	0.0083	0.2265	0.2268	0.940
		0.50	0.0038	0.2856	0.2722	0.934	−0.0094	0.2790	0.2713	0.937
		1.00	−0.0040	0.3545	0.3418	0.942	0.0036	0.3624	0.3492	0.941
	0.5	0.00	0.0003	0.1477	0.1461	0.941	−0.0024	0.1471	0.1452	0.943
		0.25	0.0086	0.2232	0.2132	0.932	0.0086	0.2232	0.2132	0.932
		0.50	0.0048	0.2779	0.2651	0.940	−0.0008	0.2761	0.2630	0.935
		1.00	−0.0104	0.3570	0.3372	0.934	0.0031	0.3574	0.3421	0.928
	1	0.00	−0.0048	0.1323	0.1299	0.939	0.0079	0.1301	0.1288	0.932
		0.25	0.0002	0.2134	0.2042	0.928	−0.0025	0.2170	0.2049	0.929
		0.50	−0.0039	0.2598	0.2583	0.940	−0.0105	0.2660	0.2587	0.945
		1.00	0.0054	0.3471	0.3400	0.950	0.0087	0.3645	0.3392	0.931

Table 2
Simulation results for the estimation of γ under $g_2(x)$

			$H_1(t)$				$H_2(t)$			
n	γ	σ^2	Bias	SE	SEE	CP	Bias	SE	SEE	CP
100	0	0.00	0.0052	0.1984	0.1928	0.932	−0.0100	0.1946	0.1803	0.921
		0.25	0.0137	0.2749	0.2553	0.928	0.0032	0.2631	0.2488	0.932
		0.50	−0.0137	0.3387	0.3042	0.915	−0.0029	0.3146	0.2986	0.932
		1.00	0.0044	0.4047	0.3751	0.924	0.0107	0.4236	0.3730	0.906
	0.5	0.00	0.0020	0.1704	0.1560	0.934	−0.0051	0.1481	0.1445	0.934
		0.25	−0.0093	0.2412	0.2255	0.924	0.0089	0.2401	0.2190	0.919
		0.50	−0.0033	0.3014	0.2750	0.918	0.0056	0.2936	0.2736	0.921
		1.00	−0.0029	0.3838	0.3505	0.911	0.0119	0.3989	0.3548	0.915
	1	0.00	−0.0000	0.1332	0.1242	0.925	0.0025	0.1221	0.1145	0.921
		0.25	−0.0006	0.2300	0.2068	0.923	−0.0144	0.2071	0.2013	0.931
		0.50	0.0018	0.2760	0.2590	0.919	−0.0349	0.2548	0.2547	0.939
		1.00	−0.0009	0.3662	0.3423	0.926	−0.0777	0.3184	0.3343	0.936
200	0	0.00	−0.0025	0.1412	0.1377	0.945	0.0022	0.1354	0.1285	0.938
		0.25	−0.0050	0.1960	0.1845	0.928	0.0062	0.1866	0.1792	0.935
		0.50	0.0058	0.2337	0.2193	0.922	0.0035	0.2292	0.2161	0.933
		1.00	−0.0059	0.2869	0.2760	0.939	−0.0013	0.2971	0.2735	0.922
	0.5	0.00	0.0027	0.1146	0.1107	0.940	0.0003	0.1048	0.1022	0.945
		0.25	0.0106	0.1700	0.1631	0.927	−0.0038	0.1649	0.1584	0.931
		0.50	−0.0001	0.2113	0.2011	0.936	0.0073	0.2127	0.1996	0.939
		1.00	−0.0104	0.2738	0.2562	0.924	0.0028	0.2762	0.2585	0.935
	1	0.00	−0.0064	0.0902	0.0887	0.950	0.0004	0.0858	0.0817	0.935
		0.25	−0.0115	0.1536	0.1491	0.935	−0.0058	0.1524	0.1470	0.933
		0.50	−0.0090	0.2045	0.1892	0.936	−0.0235	0.1910	0.1880	0.935
		1.00	−0.0022	0.2646	0.2496	0.940	−0.0263	0.2483	0.2513	0.942

Table 3
Simulation results for the estimations of $\beta_1(t)$ and $\beta_2(t)$ under $g_1(x)$

				$\beta_1(t)$				$\beta_2(t)$			
n	γ	σ^2	t	Bias	SE	SEE	CP	Bias	SE	SEE	CP
100	0	0	1	−0.0264	0.2557	0.2435	0.938	0.0038	0.3057	0.2918	0.937
			3	−0.0116	0.1861	0.1781	0.938	0.0092	0.1823	0.1722	0.935
			5	−0.0106	0.1720	0.1651	0.940	0.0101	0.1485	0.1405	0.935
		1	1	−0.0537	0.3842	0.3541	0.927	0.0046	0.3708	0.3593	0.944
			3	−0.0405	0.3411	0.3183	0.932	0.0098	0.3104	0.2911	0.930
			5	−0.0394	0.3457	0.3206	0.927	0.0052	0.3216	0.3006	0.927
	0.5	0	1	−0.0183	0.2271	0.2178	0.931	0.0063	0.2671	0.2534	0.950
			3	−0.0048	0.1731	0.1624	0.931	−0.0033	0.1542	0.1500	0.934
			5	−0.0053	0.1583	0.1515	0.940	−0.0061	0.1277	0.1234	0.939
		1	1	−0.0396	0.3571	0.3434	0.933	−0.0029	0.3516	0.3306	0.931
			3	−0.0411	0.3318	0.3169	0.934	0.0021	0.2838	0.2817	0.940
			5	−0.0407	0.3350	0.3202	0.935	−0.0061	0.3077	0.2945	0.926
	1	0	1	−0.0176	0.2010	0.1945	0.947	0.0003	0.2259	0.2204	0.944
			3	−0.0118	0.1550	0.1483	0.938	0.0053	0.1355	0.1309	0.937
			5	−0.0075	0.1433	0.1390	0.943	0.0010	0.1096	0.1076	0.942
		1	1	−0.0386	0.3457	0.3375	0.943	0.0007	0.3065	0.3091	0.946
			3	−0.0310	0.3259	0.3183	0.937	0.0025	0.2905	0.2739	0.934
			5	−0.0298	0.3335	0.3231	0.938	−0.0025	0.3136	0.2898	0.918
200	0	0	1	−0.0184	0.1716	0.1728	0.960	−0.0019	0.2011	0.2053	0.958
			3	−0.0082	0.1268	0.1273	0.951	0.0003	0.1203	0.1225	0.952
			5	−0.0051	0.1203	0.1181	0.949	−0.0011	0.1012	0.1007	0.950
		1	1	−0.0207	0.2609	0.2504	0.939	−0.0002	0.2546	0.2525	0.948
			3	−0.0164	0.2378	0.2276	0.953	0.0017	0.2127	0.2072	0.941
			5	−0.0185	0.2436	0.2289	0.938	0.0034	0.2225	0.2148	0.935
	0.5	0	1	−0.0026	0.1597	0.1532	0.945	−0.0071	0.1812	0.1789	0.942
			3	0.0020	0.1175	0.1149	0.939	−0.0038	0.1133	0.1068	0.934
			5	0.0029	0.1083	0.1071	0.945	−0.0014	0.0893	0.0876	0.946
		1	1	−0.0104	0.2570	0.2448	0.936	0.0027	0.2414	0.2332	0.949
			3	−0.0039	0.2396	0.2282	0.932	−0.0016	0.2029	0.2012	0.942
			5	−0.0105	0.2492	0.2316	0.928	0.0049	0.2256	0.2121	0.930
	1	0	1	−0.0084	0.1399	0.1371	0.945	0.0054	0.1604	0.1559	0.937
			3	−0.0028	0.1068	0.1048	0.950	0.0045	0.0926	0.0928	0.948
			5	−0.0024	0.1016	0.0985	0.940	0.0029	0.0762	0.0763	0.956
		1	1	−0.0200	0.2469	0.2441	0.949	−0.0013	0.2188	0.2192	0.948
			3	−0.0178	0.2372	0.2323	0.941	−0.0022	0.1984	0.1980	0.939
			5	−0.0200	0.2387	0.2364	0.946	−0.0031	0.2177	0.2131	0.939

2002; Sun and Wei, 2000; Zhang, 2002). In the study, the patients with stage I bladder cancer were randomly assigned to placebo, pyridoxine, or intravesical thiotepa and followed for recurrences of superficial bladder tumors. Following previous authors, we focus our attention on the comparison between thiotepa and placebo. There are 85 bladder cancer patients, 47 in the placebo group, and 38 in the thiotepa treatment group. At the beginning of the study, for each patient, two baseline covariates were measured—the number of initial tumors that the patients had before entering the study and the size of the largest initial tumor. Since the size of the largest initial tumor had been shown to have no effect on the recurrence rate (Sun and Wei, 2000; Zhang, 2002), here we focus on the effects of thiotepa treatment and the number of initial tumors on the mean functions of the recurrent events given survival.

For the analysis, we defined the covariates as $X_i = 0$ if the patient was in the placebo group and 1 if the patient was in the thiotepa group and Z_i as the number of initial tumors, $i = 1, \ldots, 85$. First, we assumed that the data can be described by the following marginal model:

$$E\left\{N_i^*(t) \,\middle|\, X_i, Z_i, D \geqslant t\right\} = 0.3 \exp\{\beta_1(t) + \beta_2(t)X_i + \gamma Z_i\}, \tag{13}$$

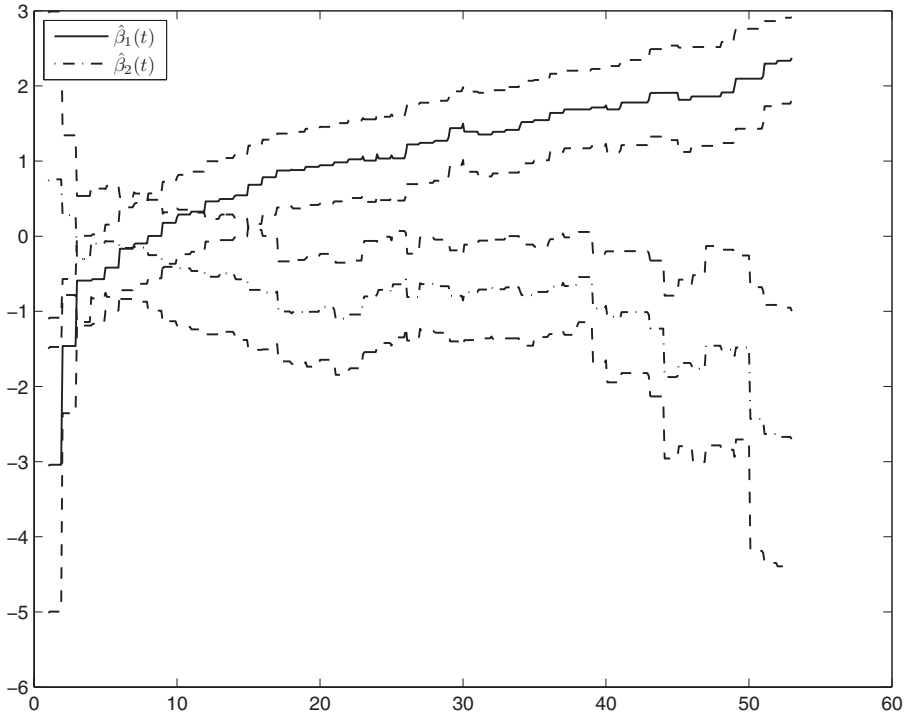
and model (2) with $W_i = (X_i, Z_i)'$. Let τ be the last observed event time being 53 months. Then, the application of the proposed method in previous sections to the data yielded $\hat{\gamma} = 0.2029$ with estimated standard error of 0.0611 for $H_1(t) = t$, and $\hat{\gamma} = 0.1679$ with estimated standard error of 0.0573 for $H_2(t) = n^{-1}\sum_{i=1}^{n} N_i(t)$, which show that the number of initial tumors has a significant positive effect on the mean functions of the recurrent events given survival. Table 5 gives the estimates of $\beta_1(t)$ and $\beta_2(t)$ at some time points. Figure 1 displays the estimates of $\beta_1(t)$ and $\beta_2(t)$ with the pointwise 95% confidence band based on 530 grid points $0.1k$, $k = 1, \ldots, 530$. This means that the treatment effect seems to change with time.

Table 4
Simulation results for the estimations of $\beta_1(t)$ and $\beta_2(t)$ under $g_2(x)$

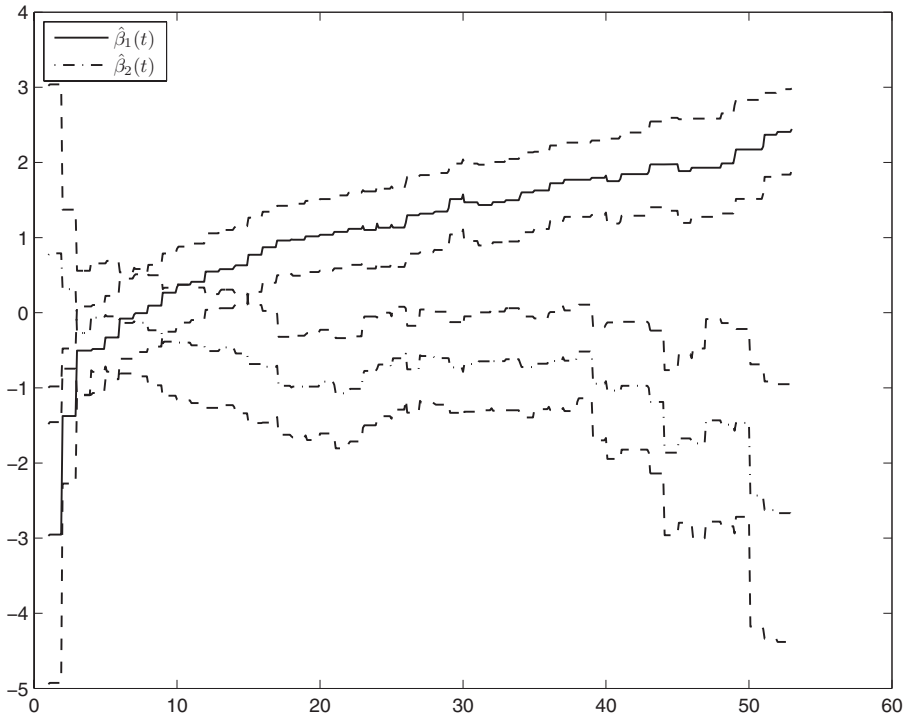
				$\beta_1(t)$				$\beta_2(t)$			
n	γ	σ^2	t	Bias	SE	SEE	CP	Bias	SE	SEE	CP
100	0	0	1	−0.0559	0.3162	0.2987	0.948	0.0171	0.4012	0.3824	0.956
			3	−0.0240	0.1838	0.1772	0.943	0.0119	0.1845	0.1776	0.945
			5	−0.0186	0.1510	0.1475	0.945	0.0129	0.1272	0.1223	0.949
		1	1	−0.0602	0.3903	0.3644	0.942	−0.0035	0.4425	0.4256	0.945
			3	−0.0402	0.2924	0.2748	0.937	0.0178	0.2632	0.2563	0.948
			5	−0.0405	0.2739	0.2598	0.937	0.0165	0.2394	0.2250	0.930
	0.5	0	1	−0.0340	0.2678	0.2539	0.938	0.0141	0.3395	0.3199	0.937
			3	−0.0096	0.1550	0.1495	0.944	0.0040	0.1426	0.1434	0.945
			5	−0.0098	0.1316	0.1254	0.938	0.0075	0.1032	0.0972	0.930
		1	1	−0.0514	0.3582	0.3337	0.937	−0.0126	0.4077	0.3738	0.935
			3	−0.0339	0.2816	0.2608	0.933	0.0023	0.2481	0.2319	0.933
			5	−0.0313	0.2652	0.2509	0.939	0.0002	0.2213	0.2079	0.939
	1	0	1	−0.0179	0.2233	0.2105	0.944	−0.0034	0.2746	0.2625	0.945
			3	−0.0051	0.1278	0.1235	0.949	0.0018	0.1178	0.1116	0.936
			5	−0.0041	0.1099	0.1044	0.939	0.0045	0.0769	0.0742	0.928
		1	1	−0.0407	0.3253	0.3104	0.935	0.0065	0.3349	0.3206	0.949
			3	−0.0243	0.2731	0.2592	0.932	−0.0035	0.2215	0.2117	0.934
			5	−0.0239	0.2639	0.2536	0.926	−0.0045	0.2129	0.1962	0.922
200	0	0	1	−0.0233	0.2097	0.2079	0.956	0.0019	0.2645	0.2652	0.953
			3	−0.0054	0.1269	0.1250	0.950	0.0049	0.1302	0.1251	0.943
			5	−0.0040	0.1073	0.1048	0.944	0.0030	0.0891	0.0868	0.940
		1	1	−0.0331	0.2630	0.2596	0.947	0.0154	0.2957	0.2997	0.967
			3	−0.0164	0.2042	0.1980	0.954	0.0046	0.1795	0.1844	0.954
			5	−0.0122	0.1934	0.1875	0.936	0.0040	0.1628	0.1618	0.946
	0	0	1	−0.0291	0.1818	0.1796	0.949	0.0143	0.2206	0.2250	0.962
			3	−0.0086	0.1083	0.1056	0.955	0.0056	0.1038	0.1014	0.945
			5	−0.0053	0.0928	0.0892	0.939	0.0036	0.0722	0.0692	0.941
		1	1	−0.0069	0.2483	0.2356	0.937	−0.0036	0.2629	0.2604	0.944
			3	−0.0053	0.1954	0.1886	0.933	0.0025	0.1678	0.1648	0.943
			5	−0.0091	0.1869	0.1819	0.943	0.0085	0.1506	0.1486	0.946
	1	0	1	−0.0122	0.1507	0.1485	0.945	0.0059	0.1903	0.1837	0.947
			3	−0.0008	0.0880	0.0877	0.953	0.0025	0.0793	0.0792	0.959
			5	0.0016	0.0726	0.0745	0.957	0.0014	0.0528	0.0527	0.961
		1	1	−0.0203	0.2303	0.2228	0.941	0.0017	0.2318	0.2290	0.948
			3	−0.0147	0.1917	0.1876	0.949	−0.0009	0.1490	0.1517	0.952
			5	−0.0155	0.1901	0.1842	0.955	−0.0015	0.1448	0.1420	0.946

Table 5
Regression analysis for the bladder cancer data

	$H_1(t)$				$H_2(t)$			
t	$\hat{\beta}_1(t)$	SE($\hat{\beta}_1(t)$)	$\hat{\beta}_2(t)$	SE($\hat{\beta}_2(t)$)	$\hat{\beta}_1(t)$	SE($\hat{\beta}_1(t)$)	$\hat{\beta}_2(t)$	SE($\hat{\beta}_2(t)$)
5	−0.4185	0.2922	−0.0627	0.3520	−0.3302	0.2830	−0.0291	0.3495
10	0.2506	0.2728	−0.4081	0.3970	0.3411	0.2597	−0.3857	0.3930
15	0.6843	0.2661	−0.7147	0.4070	0.7706	0.2518	−0.6835	0.4006
20	0.9437	0.2609	−0.9406	0.3594	1.0363	0.2431	−0.9172	0.3519
25	1.0760	0.2779	−0.7570	0.3576	1.1724	0.2594	−0.7304	0.3476
30	1.4985	0.2471	−0.8551	0.3084	1.5737	0.2376	−0.7877	0.3064
35	1.5451	0.2723	−0.7417	0.3068	1.6276	0.2624	−0.6794	0.3079
40	1.7398	0.2600	−0.9261	0.3704	1.8240	0.2507	−0.9109	0.3854
45	1.9089	0.3210	−1.8659	0.5536	1.9757	0.3156	−1.8538	0.5608
50	2.0953	0.3390	−1.4803	0.6244	2.1727	0.3360	−1.4664	0.6376



(a)



(b)

Figure 1. Estimates of time-varying coefficients and their pointwise 95% confidence bands for the bladder cancer data. (a) Plots under $H_1(t) = t$; (b) Plots under $H_2(t) = n^{-1}\sum_{i=1}^{n} N_i(t)$. The solid lines are the estimates of $\beta_1(t)$, the dash-dotted lines are the estimates of $\beta_2(t)$, and the dashed lines are their pointwise 95% confidence bands.

To test whether the treatment effect is constant or varying over time, that is, $H_0 : \beta_2(t)$ is a constant, we computed the Kolmogorov-Smirnov type statistic $\mathcal{F}_2^{(2)}$ and the Cramér-von Mises type statistic $\mathcal{F}_2^{(3)}$ by the simulation technique described in Section 4, based on 5000 realizations. The p-values of the test statistics $\mathcal{F}_2^{(2)}$ and $\mathcal{F}_2^{(3)}$ are found to be 0.0300 (0.0258) and 0.0256 (0.0252), respectively, with the weight function $H_1(t)(H_2(t))$. These small p-values indicate that the treatment effect is varying over time in model (13). We further test the significance of the time-varying treatment effect by calculating the test statistic $\mathcal{F}_2^{(1)}$ defined in Section 4. The p-values of $\mathcal{F}_2^{(1)}$ are 0.0132 and 0.0092 under $H_1(t)$ and $H_2(t)$, respectively. This confirms that the thiotepa treatment has a significant time-varying effect in reducing the recurrence of bladder tumor among survivors at a given time.

Finally, we apply the model checking techniques introduced in Section 4 to assess the adequacy of models (1) and (2) for these data. Since the two covariates are time-invariant, we used the following supremum test statistic $\sup_{1<t\leqslant\tau,x,z} |\mathcal{F}(t,x,z)|$. We calculated the statistic $\mathcal{F}(t,x,z)$ and obtained $\sup_{1<t\leqslant\tau,x,z} |\mathcal{F}(t,x,z)| = 1.3131$ and 1.5314 with p-values of 0.3184 and 0.1706 under $H_1(t)$ and $H_2(t)$, respectively, based on 5000 realizations of the corresponding statistic $\sup_{1<t\leqslant\tau,x,z} |\hat{F}(t,x,z)|$. This result indicates that model (13) fits the data adequately.

7. Concluding Remarks

In this article, we proposed a family of semiparametric transformation models with time-varying coefficients for recurrent event data in the presence of a terminal event, which includes the semiparametric transformation models studied by Lin et al. (2001) as special cases when there is no dependent terminal event. The new model is more versatile and flexible as it can summarize effects more clearly and is more useful as a diagnostic tool for time-varying coefficients in the mean functions of the recurrent events given survival. An estimation procedure was proposed for the parametric as well as the nonparametric components of the model, and asymptotic properties of the proposed estimators were established. A lack-of-fit test was provided for assessing the adequacy of the model, and some tests were also presented for investigating whether or not covariate effects vary with time. Simulation results showed that the proposed methods work well for the situations considered. The methodology was illustrated by the analysis of the bladder cancer data from a clinic trial.

Note that for recurrent event data in the absence of a terminal event, $E\{N^*(t) \mid X, Z\} = \int_0^t E\{dN^*(u) \mid X, Z\}$, that is, the rate function is the derivative of the mean function. Thus, a marginal mean model is associated with a marginal rate model, and vice versa. But for recurrent event data in the presence of a terminal event, the conditional mean function $E\{N^*(t) \mid X, Z, D \geqslant t\}$ is usually not equal to $\int_0^t E\{dN^*(u) \mid X, Z, D \geqslant u\}$ and the latter does not lead to anything interpretable unless the recurrent events are independent of the terminal event (e.g., Cook and Lawless, 1997). Hence a conditional mean model is different from a conditional rate model in this situation. The conditional rate model can be used to yield the marginal mean model since $E\{N^*(t) \mid X, Z\} = \int_0^t P\{D_i \geqslant u \mid X, Z\} E\{dN^*(u) \mid X, Z, D_i \geqslant u\}$, while the conditional mean model is more intuitive than the conditional rate model, especially to practitioners, since the mean number of events is a more interpretable quantity than any other quantities in the context of recurrent event data. However, the proposed estimation procedure for model (1) cannot be extended in a straightforward manner to deal with a conditional rate model with time-varying coefficients.

For the case of time-dependent covariates, let $\boldsymbol{X}(t) = \{X(s); s \in [0,t]\}$ and $\boldsymbol{Z}(t) = \{Z(s); s \in [0,t]\}$. An obvious extension of model (1) would be

$$E\{N^*(t) \mid \boldsymbol{X}(t), \boldsymbol{Z}(t), D \geqslant t\} = g\{\beta_0(t)'X(t) + \gamma_0' Z(t)\}. \quad (14)$$

All time-dependent covariates are assumed to be external as defined by Kalbfleisch and Prentice (2002, p. 196). The proposed estimation procedure can be extended in a straightforward manner to deal with time-dependent covariates.

Since estimating functions (4) and (5) were given in a somewhat ad hoc fashion using the generalized estimating equation and the technique of inverse probability weighting, it would be worthwhile to further investigate the efficiency of the proposed estimators. If $N_i^*(\cdot)$ is a Poisson process, then it might be possible to estimate $\beta_0(t)$ and γ_0 more efficiently by the nonparametric maximum likelihood approach, and the resulting inference procedure would be much more complicated.

The proposed estimation procedure requires modeling the survival distribution, and we have used the Cox proportional hazards model for the survival time. Other competing models, such as the additive hazards model, the accelerated failure time model, and the linear transformation model may be used as well. It would be worthwhile to investigate the potential bias due to misspecification for each of these models.

In practice, the choice of an appropriate link function g may be based on prior data or the desiring interpretation of the regression parameters (e.g., Lin et al., 2001). Notice that the magnitudes of the parameter estimates are quite different for various choices of g. This is because the parameters have different interpretations for different g. If some covariate has a nonzero and time-independent effect for one choice of g, then it would appear that its effect would be time-dependent for any other link function. Further research is needed to provide a method for selecting or comparing different link functions.

Another interesting issue is the effect of the weight $H(t)$. Ideally, we would choose $H(t)$ to minimize the variances of $\hat{\beta}(t)$ and $\hat{\gamma}$. However, it does not appear to be possible to derive an optimal weight without specification of the dependence structure on the increments of $N^*(t)$, and the selection of weight functions is usually a complicated problem (Lin et al., 2001). Development of a simple but more efficient inference procedure merits future research.

8. Supplementary Materials

The Web Appendix referenced in Sections 3 and 4 is available under the Paper Information link at the *Biometrics* website `http://www.biometrics.tibs.org`.

Acknowledgements

The authors are grateful to the editor, Dr David Zucker, the associate editor, and the referee for their many helpful

comments and suggestions that greatly improved this article. This research is partly supported by the AMSS-PolyU Joint Research Institute. The first author's research was partly supported by the Research Grant Council of Hong Kong. The third author's research was partly supported by the National Natural Science Foundation of China Grants (10731010, 10971015, and 10721101), the National Basic Research Program of China (973 Program) (2007CB814902), and Key Laboratory of RCSDS, CAS (2008DP173182).

References

Andersen, P. K. and Gill, R. D. (1982). Cox's regression model for counting processes: A large sample study. *Annals of Statistics* **10,** 1100–1120.

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*, M. Pavane-Macaluso, P. H. Smith, and F. Edsmyr (eds), 363–370. New York: Plenum.

Chang, S. H. and Wang, M. C. (1999). Conditional regression analysis for recurrence time data. *Journal of the American Statistical Association* **94,** 1221–1230.

Cook, R. J. and Lawless, J. F. (1997). Marginal analysis of recurrent events and a terminating event. *Statistics in Medicine* **16,** 911–924.

Cook, R. J. and Lawless, J. F. (2007). *The Statistical Analysis of Recurrent Events*. New York: Springer.

Ghosh, D. (2004). Accelerated rates regression models for recurrent failure time data. *Lifetime Data Analysis* **10,** 247–261.

Ghosh, D. and Lin, D. Y. (2000). Nonparametric analysis of recurrent events and death. *Biometrics* **56,** 554–562.

Ghosh, D. and Lin, D. Y. (2002). Marginal regression models for recurrent and terminal events. *Statistics Sinica* **12,** 663–688.

Ghosh, D. and Lin, D. Y. (2003). Semiparametric analysis of recurrent events data in the presence of dependent censoring. *Biometrics* **59,** 877–885.

Huang, C. Y. and Wang, M. C. (2004). Joint modeling and estimation for recurrent event processes and failure time data. *Journal of the American Statistical Association* **99,** 1153–1165.

Kalbfleisch, J. D. and Prentice, R. L. (2002). *The Statistical Analysis of Failure Time Data*, 2nd edition. New York: Wiley.

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

Lee, E. W., Wei, L. J., and Amato, D. A. (1992). Cox-type regression analysis for large numbers of small groups of correlated failure time observations. In *Survival Analysis: State of the Art*, J. P. Klein and P. K. Goel (eds), 237–247. Dordrecht: Kluwer Academic Publishers.

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

Lin, D. Y., Wei, L. J., Yang, I., and 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, D. Y., Wei, L. J., and Ying, Z. (1993). Checking the Cox model with cumulative sums of martingale-based residuals. *Biometrika* **80,** 557–572.

Lin, D. Y., Wei, L. J., and Ying, Z. (1998). Accelerated failure time models for counting processes. *Biometrika* **85,** 605–618.

Lin, D. Y., Wei, L. J., and Ying, Z. (2001). Semiparametric transformation models for point processes. *Journal of the American Statistical Association* **96,** 620–628.

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

Murphy, S. A. (1994). Consistency in a proportional hazards model incorporating a random effect. *Annals of Statistics* **22,** 712–731.

Murphy, S. A. (1995). Asymptotic theory for the frailty model. *Annals of Statistics* **23,** 182–198.

Nielsen, G. G., Gill, R. D., Andersen, P. K., and Sorensen, T. I. A. (1992). A counting process approach to maximum likelihood estimation in frailty models. *Scandinavian Journal of Statistics* **19,** 25–43.

Pan, Q. and Schaubel, D. E. (2009). Flexible estimation of differences in treatment-specific recurrent event means in the presence of a terminating event. *Biometrics* **65,** 753–761.

Pepe, M. S. and Cai, J. (1993). Some graphical displays and marginal regression analyses for recurrent failure times and time-dependent covariates. *Journal of the American Statistical Association* **88,** 811–820.

Prentice, R. L., Williams, B. J., and Peterson, A. V. (1981). On the regression analysis of multivariate failure time data. *Biometrika* **68,** 373–379.

Schaubel, D. E. and Cai, J. (2005). Analysis of clustered recurrent event data with application to hospitalization rates among renal failure patients. *Biostatistics* **6,** 404–419.

Schaubel, D. E., Zeng, D., and Cai, J. (2006). A semiparametric additive rates model for recurrent event data. *Lifetime Data Analysis* **12,** 389–406.

Sun, J. and 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.

Wei, L. J., Lin, D. Y., and Weissfeld, L. (1989). Regression analysis of multivariate incomplete failure time data by modeling marginal distributions. *Journal of the American Statistical Association* **84,** 1065–1073.

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

Zeng, D. and Lin, D. Y. (2006). Efficient estimation of semiparametric transformation models for counting processes. *Biometrika* **93,** 627–640.

Zeng, D. and Lin, D. Y. (2007). Semiparametric transformation models with random effects for recurrent events. *Journal of the American Statistical Association* **102,** 167–180.

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

Received December 2009. *Revised April* 2010.
Accepted April 2010.