The Canadian Journal of Statistics
Vol. 39, No. 4, 2011, Pages 578–590
La revue canadienne de statistique

A class of mixed models for recurrent event data

Liuquan SUN[1], Xingqiu ZHAO[2]* and Jie ZHOU[1]

[1] *Institute of Applied Mathematics, Academy of Mathematics and Systems Science, Chinese Academy of Sciences, Beijing 100190, China*
[2] *Department of Applied Mathematics, The Hong Kong Polytechnic University, Hong Kong*

Key words and phrases: Counting process; marginal rate model; mixed model; partial-score function; proportional and Convergent Effects.

MSC 2010: Primary 62N01; secondary 62N02

Abstract: In this article, we propose a class of mixed models for recurrent event data. The new models include the proportional rates model and Box–Cox transformation rates models as special cases, and allow the effects of covariates on the rate functions of counting processes to be proportional or convergent. For inference on the model parameters, estimating equation approaches are developed. The asymptotic properties of the resulting estimators are established and the finite sample performance of the proposed procedure is evaluated through simulation studies. A real example with data taken from a clinic study on chronic granulomatous disease (CGD) is also illustrated for the use of the proposed methodology. *The Canadian Journal of Statistics* 39: 578–590; 2011 © 2011 Statistical Society of Canada

Résumé: Dans cet article, nous proposons une classe de modèles mixtes pour des donnes d'événements récurrents. Les nouveaux modèles incluent le modèle de taux proportionels et ceux des taux avec transformation de Box-Cox comme cas spéciaux. Ils permettent aussi aux effets des covariables sur les fonctions de taux des processus de comptage d'être proportionnel ou convergent. Une approche utilisant les équations d'estimation est développée pour faire l'inférence sur les paramètres des modèles. Les propriétés asymptotiques des estimateurs résultants sont obtenues et la performance, pour de petits échantillons, de la procédure proposée est évaluée par des études de simulation. La méthodologie proposée est illustrée à l'aide d'un vrai exemple avec des données Tirées d'un essai clinique sur la granulomatose chronique familiale (CGD). *La revue canadienne de statistique* 39: 578–590; 2011 © 2011 Société statistique du Canada

1. INTRODUCTION

For the analysis of recurrent event data, the most popular model is the proportional intensity model proposed by Anderson & Gill (1982). Under this model, the underlying counting process arising from recurrent events is a Poisson process. This may not be true in practice. To relax such Poisson assumption, several authors have studied proportional means and rates models; see Pepe & Cai (1993), Lawless & Nadeau (1995), Lin et al. (2000), and Ghosh & Lin (2002) among others. However, in the proportional means and rates models, the covariates have a fixed multiplicative effect on the mean and rate functions. In many applications, it is unreasonable to assume that the effects of covariates measured at the beginning of a study remain fixed over time. For example, Lin, Wei, & Ying (2001) examined data on recurrent pulmonary exacerbations from a cystic fibrosis clinical trial and found that the two mean functions for the numbers of events in the rhDNase and placebo groups converge with time rather than being proportional over time.

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

To allow nonproportional means and rates, several semiparametric regression methods have been studied. For example, Lin, Wei, & Ying (1998) considered the accelerated failure time models for counting processes, Lin, Wei, & Ying (2001) and Zeng & Lin (2006) studied a class of semiparametric transformation models for point processes through a specified link function, and Ghosh (2004) presented an accelerated rates regression model for recurrent event data in which the effect of covariates is to transform the time scale for a baseline rate function. Cook & Lawless (2007) and Aalen, Borgan, & Gjessing (2008) provided comprehensive reviews of methods for recurrent events data. Recently, Schaubel, Zeng, & Cai (2006) and Zeng & Cai (2010) provided an additive rates model, in which the covariate effects are assumed to be additive to the unspecified baseline rate, Zeng & Lin (2007) considered a class of semiparametric transformation models with random effects for recurrent events, Sun & Su (2008) studied a very general mean model which includes the accelerated and proportional mean models as special cases, Liu et al. (2010) considered an additive–multiplicative rates model, wherein some covariate effects are additive while others are multiplicative, and Sun, Tong, & Zhou (2011) proposed a family of Box–Cox transformation models, which takes Box–Cox transformation for both the conditional mean function and the baseline mean function of the point process. All the above models cannot allow for both proportional and converging covariate effects on the rate function of recurrent event data.

In this article, we propose a class of mixed models to display a variety of patterns for the effects of covariates on the rate function in the spirit of the work of Barker & Henderson (2004), which introduced a mixed model for classical survival data. The proposed models include the proportional rates model and a family of Box–Cox transformation models as subclasses, and allow for both proportional and converging effects.

Specifically, let $N^*(t)$ be the number of events that have occurred by time t and let Z be a vector of covariates. Define $\mathrm{d}\mu(t|Z) = E\{\mathrm{d}N^*(t)|Z\}$. The proportional rates model is given by

$$\mathrm{d}\mu(t|Z) = \mathrm{e}^{\beta' Z}\, \mathrm{d}\mu_0(t),$$

and a class of Box–Cox transformation models take the form

$$\mathrm{d}\mu(t|Z) = \frac{\mathrm{e}^{\beta' Z}\, \mathrm{e}^{\gamma\mu_0(t)}}{1 + \mathrm{e}^{\beta' Z}(\mathrm{e}^{\gamma\mu_0(t)} - 1)}\, \mathrm{d}\mu_0(t)$$

where $\mathrm{d}\mu_0(t) = E\{\mathrm{d}N^*(t)|Z = 0\}$, which is referred to as an unspecified baseline rate function, β is a vector of regression parameters, and $\gamma \geq 0$ is an unknown scalar parameter (Sun, Tong, & Zhou, 2011). Note that the proportional rates model only allows for proportional covariate effects and the Box–Cox transformation model only allows for converging covariate effects on the rate function. In practice, however, some covariates may have proportional effects, and the other covariates may have converging effects. To this end, let $Z = (Z_1', Z_2')'$, where Z_1 and Z_2 are $p \times 1$ and $q \times 1$ vectors of covariates, respectively. Assume that covariates Z_1 have proportional effects, and Z_2 have converging effects. By combining the proportional rates model and a class of Box–Cox transformation models, we obtain a class of mixed models, which take the form

$$\mathrm{d}\mu(t|Z) = \frac{\mathrm{e}^{\beta_1' Z_1 + \beta_2' Z_2}\, \mathrm{e}^{\gamma\mu_0(t)}}{1 + \mathrm{e}^{\beta_2' Z_2}(\mathrm{e}^{\gamma\mu_0(t)} - 1)}\, \mathrm{d}\mu_0(t), \tag{1}$$

where β_1 and β_2 are vectors of regression parameters. Clearly, when $\gamma = 0$, (1) reduces to the proportional rates model with covariates Z. If $\beta_1 = 0$, the model reduces to a family of Box–Cox transformation models with covariates Z_2. If $\beta_2 = 0$, the model reduces to the proportional rates model with covariate Z_1. When $N^*(t)$ is a simple counting process (i.e., only take a value of 0 or 1), this family of models reduces to the mixed proportional and converging hazards models (Barker & Henderson, 2004).

For the analysis, we introduce the following baseline function

$$R(t) = \frac{1}{\gamma}\left(\mathrm{e}^{\gamma\mu_0(t)} - 1\right).$$

Note that $R(t)$ is positive, and increases with time. As $\gamma \to 0$, $R(t) \to \mu_0(t)$. Substituting $R(t)$ into (1), we get

$$\mathrm{d}\mu(t|Z) = \frac{\mathrm{e}^{\beta_1' Z_1 + \beta_2' Z_2}}{1 + \gamma\,\mathrm{e}^{\beta_2' Z_2} R(t)}\,\mathrm{d}R(t). \tag{2}$$

To demonstrate some properties of the mixed model, we consider rate ratios. Fixing the Z_2 covariates and taking a ratio of rates with two different values of Z_1, we have

$$\frac{\mathrm{d}\mu(t|Z_1, Z_2)}{\mathrm{d}\mu(t|Z_1^*, Z_2)} = \mathrm{e}^{\beta_1'(Z_1 - Z_1^*)},$$

which is proportional in Z_1. However, fixing Z_1 and taking rate ratios for two values of Z_2 gives

$$\frac{\mathrm{d}\mu(t|Z_1, Z_2)}{\mathrm{d}\mu(t|Z_1, Z_2^*)} = \frac{\mathrm{e}^{\beta_2'(Z_2 - Z_2^*)} + \gamma\,\mathrm{e}^{\beta_2' Z_2} R(t)}{1 + \gamma\,\mathrm{e}^{\beta_2' Z_2} R(t)}.$$

This ratio is $\mathrm{e}^{\beta_2'(Z_2 - Z_2^*)}$ at $t = 0$ and then converges to one as t increases. The mixed model therefore allows for a proportional effect in Z_1 and a converging effect in Z_2. Clearly, $\mathrm{d}\mu(t|Z)/\mathrm{d}\mu_0(t) \to \mathrm{e}^{\beta_1' Z_1}$ as $t \to \infty$, and the parameter γ controls the rate of convergence. If γ is close to zero, the rates converge to baseline levels extremely slowly; when γ is extremely large, the rate would converge almost immediately. This situation could occur when a covariate has a large effect on recurrent event times at the beginning of a study and then has little or no effect shortly afterwards. To ensure the identifiability of γ, it is necessary to assume that β_2 is not identically zero.

The remainder of the paper is organised as follows. Section 2 presents inference procedures for the regression parameters and the scalar parameter based on recurrent event data, and the asymptotic properties of the proposed estimates are established. Section 3 reports some results from simulation studies conducted for evaluating the proposed methods. In Section 4, we apply the methodology to the CGD data and Section 5 concludes with some remarks.

2. ESTIMATION PROCEDURES

In practice, $N^*(\cdot)$ may not be fully observed since the subject is often followed for a limited period of time. Specifically, $N(t) = N^*(t \wedge C)$ is observed instead of $N^*(t)$, where C is the follow-up or censoring time, and $a \wedge b = \min(a, b)$. Following Lin, Wei, & Ying (2001), we assume that C is independent of $N^*(\cdot)$ conditional on Z. For a random sample of n subjects, the observed data consist of $\{C_i, N_i(t), Z_i;\ t \le C_i\}$ $(i = 1, \ldots, n)$, where $Z_i = (Z_{1i}', Z_{2i}')'$. Here, we focus on estimation of $\theta = (\beta_1', \beta_2', \gamma)'$ in model (1). Let $\theta_0 = (\beta_{10}', \beta_{20}', \gamma_0)'$ be the true value of θ. Define

$$M_{iR}(t) = \int_0^t \left\{\mathrm{e}^{-\beta_{10}' Z_{1i} - \beta_{20}' Z_{2i}} + \gamma_0\,\mathrm{e}^{-\beta_{10}' Z_{1i}} R_0(u)\right\} \mathrm{d}N_i(u) - \int_0^t Y_i(u)\,\mathrm{d}R_0(u),$$

where $Y_i(t) = I(C_i \geq t)$, and $R_0(t) = \gamma_0^{-1}(\mathrm{e}^{\gamma_0 \mu_0(t)} - 1)$. Note that under model (1), $M_{iR}(t)$'s are zero-mean stochastic processes. Then, for given θ. a natural estimator for $R_0(t)$ is the solution to

$$\sum_{i=1}^{n} \left[\left\{ \mathrm{e}^{-\beta_1' Z_{1i} - \beta_2' Z_{2i}} + \gamma\, \mathrm{e}^{-\beta_1' Z_{1i}} R(t) \right\} \mathrm{d}N_i(t) - Y_i(t)\, \mathrm{d}R(t) \right] = 0, \quad 0 \leq t \leq \tau,$$

where τ is a prespecified constant such that $P(C_i \geq \tau) > 0$. Denote this estimator by $\hat{R}(t;\theta)$, which can be obtained from the following integral equation

$$\hat{R}(t;\theta) = \int_0^t \gamma \hat{R}(u;\theta)\, \mathrm{d}\hat{\Lambda}(u;\beta_1, 0) + \hat{\Lambda}(t;\beta_1,\beta_2), \tag{3}$$

where

$$\hat{\Lambda}(t;\beta_1,\beta_2) = \int_0^t \frac{\sum_{i=1}^n \mathrm{e}^{-\beta_1' Z_{1i} - \beta_2' Z_{2i}}\, \mathrm{d}N_i(u)}{\sum_{i=1}^n Y_i(u)}.$$

Thus it follows from (3) that $\hat{R}(t;\theta)$ has the following explicit expression (e.g., Yang, 1992)

$$\hat{R}(t;\theta) = \frac{1}{\hat{P}(t;\beta_1,\gamma)} \int_0^t \hat{P}(u-;\beta_1,\gamma)\, \mathrm{d}\hat{\Lambda}(u;\beta_1,\beta_2),$$

where $\hat{P}(t;\beta_1,\gamma) = \prod_{u \leq t}\{1 - \gamma \Delta \hat{\Lambda}(u;\beta_1,0)\}$ is the product-integral of $\gamma\hat{\Lambda}(u;\beta_1,0)$ over $[0,t]$ and $\Delta\hat{\Lambda}(u;\beta_1,0)$ is the jump of $\hat{\Lambda}(u;\beta_1,0)$ at u.

Define

$$\hat{\Omega}_k(t;,\beta_1,\beta_2) = \int_0^t \frac{\sum_{i=1}^n Z_{ki}\, \mathrm{e}^{-\beta_1' Z_{1i} - \beta_2' Z_{2i}}\, \mathrm{d}N_i(u)}{\sum_{i=1}^n Y_i(u)}, \quad k = 1, 2,$$

$$\hat{\Phi}_1(t;\theta) = -\frac{1}{\hat{P}(t;\beta_1,\gamma)} \int_0^t \hat{P}(u-;\beta_1,\gamma)\big[\mathrm{d}\hat{\Omega}_1(u;\beta_1,\beta_2) + \gamma\{\hat{\Omega}_1(t;\beta_1,0)$$
$$-\hat{\Omega}_1(u-;\beta_1,0)\}\, \mathrm{d}\hat{\Lambda}(u;\beta_1,\beta_2)\big],$$

$$\hat{\Phi}_2(t;\theta) = -\frac{1}{\hat{P}(t;\beta_1,\gamma)} \int_0^t \hat{P}(u-;\beta_1,\gamma)\, \mathrm{d}\hat{\Omega}_2(u;\beta_1,\beta_2),$$

$$\hat{\Phi}_3(t;\theta) = \frac{1}{\hat{P}(t;\beta_1,\gamma)} \int_0^t \hat{P}(u-;\beta_1,\gamma)\big[\hat{\Lambda}(t;\beta_1,0) - \hat{\Lambda}(u-;\beta_1,0)\big]\, \mathrm{d}\hat{\Lambda}(u;\beta_1,\beta_2),$$

and $\hat{\Phi}(t;\theta) = (\hat{\Phi}_1'(t;\theta), \hat{\Phi}_2'(t;\theta), \hat{\Phi}_3(t;\theta))'$. Using the partial-score function (Lin et al., 2000) with $R(t)$ replaced by $\hat{R}(t;\theta)$, we propose to estimate θ by the solution to the following score equation

$$U(\theta) = \sum_{i=1}^{n} \int_0^{\tau} \left\{ X_i(t;\theta) - \bar{X}(t;\theta) \right\} \mathrm{d}N_i(t) = 0, \tag{4}$$

where

$$X_i(t;\theta) = \frac{\left((\mathrm{e}^{-\beta_2' Z_{2i}} + \gamma \hat{R}(t;\theta)) Z_{1i}', \mathrm{e}^{-\beta_2' Z_{2i}} Z_{2i}', -\hat{R}(t;\theta) \right)' - \gamma \hat{\Phi}(t;\theta)}{\{\mathrm{e}^{-\beta_2' Z_{2i}} + \gamma \hat{R}(t;\theta)\}},$$

and

$$\bar{X}(t;\theta) = \frac{\sum_{i=1}^{n} Y_i(t) X_i(t;\theta) \left\{ \mathrm{e}^{-\beta_1' Z_{1i} - \beta_2' Z_{2i}} + \gamma\, \mathrm{e}^{-\beta_1' Z_{1i}} \hat{R}(t;\theta) \right\}^{-1}}{\sum_{i=1}^{n} Y_i(t) \left\{ \mathrm{e}^{-\beta_1' Z_{1i} - \beta_2' Z_{2i}} + \gamma\, \mathrm{e}^{-\beta_1' Z_{1i}} \hat{R}(t;\theta) \right\}^{-1}}.$$

Denote the solution to $U(\theta) = 0$ by $\hat{\theta} = (\hat{\beta}_1', \hat{\beta}_2, \hat{\gamma})'$. Then $\hat{\theta}$ can be obtained easily through Newton–Raphson method. As in the simulation studies below, the Newton–Raphson method usually converges quickly, and it took about one second for one run with $n = 500$ in Matlab.

To find out the asymptotic distribution of $\hat{\theta}$, define $\hat{R}(t) = \hat{R}(t;\hat{\theta})$, $\hat{\pi}(t) = n^{-1} \sum_{i=1}^{n} Y_i(t)$, $\hat{D}_i(t) = \mathrm{e}^{Z_{1i}'\hat{\beta}_1} \{\mathrm{e}^{-Z_{2i}'\hat{\beta}_2} + \hat{\gamma}\hat{R}(t)\}^{-1}$, $\hat{D}_i^*(t) = \{\mathrm{e}^{-Z_{2i}'\hat{\beta}_2} + \hat{\gamma}\hat{R}(t)\}^{-1}$,

$$\bar{D}^*(t) = \sum_{i=1}^{n} Y_i(t)\hat{D}_i(t)\hat{D}_i^*(t) \Big/ \sum_{i=1}^{n} Y_i(t)\hat{D}_i(t),$$

$$\hat{M}_i(t) = N_i(t) - \int_0^t Y_i(u)\hat{D}_i(u)\,\mathrm{d}\hat{R}(u),$$

$$\hat{\xi}(t) = \frac{\hat{\gamma}\hat{P}(t;\hat{\beta}_1,\hat{\gamma})}{n\hat{\pi}(t)} \sum_{i=1}^{n} \int_t^{\tau} \frac{1}{\hat{P}(u;\hat{\beta}_1,\hat{\gamma})} Y_i(u)\hat{D}_i(u)\{\hat{D}_i^*(u) - \bar{D}^*(u)\} X_i(u;\hat{\theta})\,\mathrm{d}\hat{R}(u),$$

and

$$\hat{\eta}_i = \int_0^{\tau} \left\{ X_i(t;\hat{\theta}) - \bar{X}(t;\hat{\theta}) + \hat{D}_i(t)^{-1}\hat{\xi}(t) \right\} \mathrm{d}\hat{M}_i(t).$$

Then under some regularity conditions, following the argument given in the Appendix of Sun, Tong, & Zhou (2011), it can be shown that $\hat{\theta}$ is strongly consistent and $n^{1/2}(\hat{\theta} - \theta_0)$ is asymptotically normal with mean zero and covariance matrix that can be consistently estimated by $\hat{A}^{-1}\hat{\Sigma}\hat{A}^{-1}$, where $\hat{\Sigma} = n^{-1}\sum_{i=1}^{n} \hat{\eta}_i^{\otimes 2}$, and

$$\hat{A} = n^{-1} \sum_{i=1}^{n} \int_0^{\tau} Y_i(t)\hat{D}_i(t) \left\{ X_i(t;\hat{\theta}) - \bar{X}(t;\hat{\theta}) \right\}^{\otimes 2} \mathrm{d}\hat{R}(t).$$

Here $v^{\otimes 2} = vv'$ for a vector v.

3. SIMULATION STUDY

We conducted a simulation study to examine the finite sample properties of the proposed estimators. Let Z_{1i} and Z_{2i} be independent Bernoulli random variables with success probability 0.5. We set $\mu_0(t) = t$ (a homogeneous process) or $\mu_0(t) = 0.4(t^2 + 2t)$ (a nonhomogeneous process). For given Z_{1i} and Z_{2i}, we generated recurrent event times from the Poisson process with the intensity function (1). The censoring time C_i was taken as $\min(C_i^*, \tau)$, where C_i^* follows a uniform distribution on $(1, 6)$ and $\tau = 3$ representing the largest follow-up time. For each simulation study, we considered $\beta_2 = 1$, and $\beta_1 = 0, 0.5$ representing no effect and a positive effect of Z_{1i} on the recurrent event times, respectively. Set $\gamma = 0.05$, 0.1, or 0.5, which control the convergence rate of the effect of Z_2. The average number of events per subject ranged from 2.9 to 5.9 for different settings. The results presented below are based on 1,000 replications with sample sizes $n = 100$, 200, 300, and 500.

Tables 1 and 2 show the simulation results on estimation of $\beta = (\beta_1, \beta_2)$ and γ for $\mu_0(t) = t$, and Tables 3 and 4 present the simulation results for $\mu_0(t) = 0.4(t^2 + 2t)$. The tables include the

TABLE 1: Simulation results for estimation of β_1 and β_2 with $\mu_0(t) = t$.

			β_1				β_2			
β_1	n	γ	Bias	SSE	ESE	CP	Bias	SSE	ESE	CP
0	100	0.05	0.0030	0.0947	0.0941	0.956	0.0319	0.2177	0.2170	0.951
		0.1	0.0048	0.1006	0.0968	0.938	0.0351	0.2454	0.2321	0.941
		0.5	0.0025	0.1102	0.1078	0.939	0.0761	0.3889	0.3509	0.941
	200	0.05	−0.0032	0.0678	0.0666	0.943	0.0269	0.1523	0.1525	0.947
		0.1	0.0002	0.0698	0.0687	0.943	0.0102	0.1632	0.1623	0.950
		0.5	−0.0001	0.0800	0.0767	0.942	0.0317	0.2428	0.2396	0.943
	300	0.05	−0.0014	0.0564	0.0546	0.942	0.0094	0.1235	0.1232	0.954
		0.1	0.0000	0.0556	0.0559	0.958	0.0138	0.1335	0.1325	0.949
		0.5	−0.0016	0.0659	0.0623	0.947	0.0239	0.1920	0.1917	0.953
	500	0.05	−0.0004	0.0439	0.0422	0.941	−0.0026	0.0998	0.0949	0.939
		0.1	−0.0010	0.0440	0.0434	0.956	0.0081	0.1044	0.1026	0.951
		0.5	0.0003	0.0484	0.0485	0.952	0.0128	0.1421	0.1477	0.952
0.5	100	0.05	0.0031	0.0891	0.0837	0.933	0.0219	0.1919	0.1857	0.935
		0.1	−0.0051	0.0908	0.0865	0.933	0.0333	0.2094	0.2005	0.936
		0.5	0.0057	0.0980	0.0963	0.943	0.0764	0.3085	0.3000	0.943
	200	0.05	−0.0013	0.0616	0.0597	0.943	0.0061	0.1344	0.1311	0.949
		0.1	0.0010	0.0599	0.0612	0.951	0.0054	0.1397	0.1406	0.956
		0.5	−0.0002	0.0698	0.0687	0.951	0.0311	0.2186	0.2070	0.933
	300	0.05	0.0006	0.0491	0.0488	0.952	0.0129	0.1065	0.1075	0.951
		0.1	0.0007	0.0502	0.0500	0.947	−0.0021	0.1120	0.1140	0.952
		0.5	0.0018	0.0570	0.0559	0.957	0.0160	0.1654	0.1657	0.950
	500	0.05	0.0016	0.0371	0.0378	0.954	0.0062	0.0826	0.0829	0.958
		0.1	0.0005	0.0389	0.0389	0.955	0.0006	0.0874	0.0886	0.957
		0.5	0.0010	0.0443	0.0434	0.952	0.0127	0.1258	0.1286	0.954

TABLE 2: Simulation results for estimation of γ with $\mu_0(t) = t$.

		$\beta_1 = 0$				$\beta_1 = 0.5$			
n	γ	Bias	SSE	ESE	CP	Bias	SSE	ESE	CP
100	0.05	0.0079	0.0927	0.0895	0.948	0.0016	0.0765	0.0761	0.950
	0.1	0.0119	0.1154	0.1077	0.940	0.0106	0.0959	0.0914	0.947
	0.5	0.0817	0.4739	0.3288	0.942	0.0598	0.2982	0.2704	0.944
200	0.05	0.0092	0.0604	0.0620	0.963	0.0013	0.0551	0.0539	0.946
	0.1	0.0038	0.0786	0.0747	0.939	0.0019	0.0634	0.0647	0.957
	0.5	0.0370	0.2360	0.2145	0.946	0.0264	0.2011	0.1809	0.936
300	0.05	0.0021	0.0510	0.0505	0.942	0.0034	0.0436	0.0440	0.954
	0.1	0.0043	0.0643	0.0603	0.938	−0.0004	0.0524	0.0526	0.945
	0.5	0.0130	0.1672	0.1652	0.946	0.0101	0.1492	0.1432	0.942
500	0.05	−0.0017	0.0408	0.0390	0.936	0.0015	0.0336	0.0339	0.947
	0.1	0.0028	0.0467	0.0468	0.952	−0.0011	0.0395	0.0404	0.951
	0.5	0.0056	0.1205	0.1264	0.957	0.0105	0.1118	0.1105	0.948

estimated biases (Bias) given by the sample means minus the true values, the sampling standard errors (SSE), the sampling means of the estimated standard errors estimates (ESE), and the 95% empirical coverage probabilities (CP). The results indicate that the estimates seem to be unbiased and the proposed variance estimation procedure provides reasonable estimates. Also the results on the empirical coverage probabilities indicate that the normal approximations are appropriate. Note that both the estimated bias and standard error of $\hat{\beta}_1$ seem not affected by the value of γ. This is reasonable because Z_1 has the same effect for different values of γ. But both the estimated bias and standard error of $\hat{\beta}_2$ slightly increase with γ. This is probably because Z_2 has a converging effect, and the rate depends on γ. For larger values of γ, the effect of Z_2 converges more rapidly, and hence, the data contain less information on β_2.

To investigate the distributional behaviour of the proposed estimators $\hat{\beta}_1$, $\hat{\beta}_2$, and γ in a finite sample situation, we provide some QQ-plots in Figure 1 with $n = 300$, $\beta_1 = 0.5$, and $\mu_0(t) = t$, which suggest that the normal approximation is reasonable.

To compare the proposed estimation procedure with the maximum partial likelihood estimates (MPLE) proposed by Lin et al. (2000) for the proportional rates model, we conducted additional simulation studies for the situation where $\gamma = 0$. Table 5 presents the results on estimation of $\beta = (\beta_1, \beta_2)$ for $\gamma = 0$ and $\mu_0(t) = t$ under the same setups as those in Table 1. It can be seen from Table 5 that the estimates for β_1 are almost the same for the two methods, and our estimated standard errors for β_2 are somewhat larger than those of the MPLE. Both methods provide reasonable estimates for β, and the proposed estimator $\hat{\gamma}$ performs well for $\gamma = 0$. We also generated data with $\mu_0(t) = 0.4(t^2 + 2t)$, and obtained similar results.

Although $\gamma = 0$ looks to be a boundary point, the mixed model (1) is still a valid marginal rate model for small negative γ provided the denominator of (1) stays positive for $0 \le t \le \tau$. We further investigated the size and power of the Wald test for the hypothesis $\gamma = 0$ based on the asymptotic distribution of $\hat{\gamma}$ presented in Section 2 at the significance level of 5%, and the estimated sizes and powers are summarised in Table 6. The results suggest that the Wald test procedure seems to have the right size and reasonable power. As expected, the power increases as the sample size and the value of γ increase.

TABLE 3: Simulation results for estimation of β_1 and β_2 with $\mu_0(t) = 0.4(t^2 + 2t)$.

			β_1				β_2			
β_1	n	γ	Bias	SSE	ESE	CP	Bias	SSE	ESE	CP
0	100	0.05	−0.0013	0.0994	0.0965	0.939	0.0296	0.2195	0.2181	0.953
		0.1	0.0023	0.0997	0.0988	0.943	0.0475	0.2421	0.2341	0.946
		0.5	0.0014	0.1165	0.1108	0.929	0.0634	0.3715	0.3511	0.940
	200	0.05	0.0034	0.0705	0.0684	0.943	0.0128	0.1533	0.1530	0.947
		0.1	−0.0028	0.0699	0.0703	0.953	0.0149	0.1634	0.1638	0.942
		0.5	0.0056	0.0787	0.0786	0.946	0.0465	0.2452	0.2410	0.950
	300	0.05	−0.0021	0.0561	0.0559	0.956	0.0023	0.1240	0.1241	0.946
		0.1	−0.0001	0.0576	0.0573	0.942	0.0020	0.1341	0.1330	0.951
		0.5	−0.0000	0.0652	0.0641	0.940	0.0068	0.2014	0.1922	0.936
	500	0.05	−0.0004	0.0436	0.0432	0.949	0.0008	0.0958	0.0961	0.941
		0.1	0.0005	0.0452	0.0445	0.940	0.0060	0.1025	0.1030	0.955
		0.5	−0.0021	0.0507	0.0497	0.944	0.0152	0.1467	0.1490	0.953
0.5	100	0.05	0.0030	0.0906	0.0866	0.937	0.0061	0.1937	0.1863	0.939
		0.1	0.0020	0.0948	0.0883	0.930	0.0184	0.2028	0.2019	0.945
		0.5	0.0001	0.0982	0.0993	0.955	0.0564	0.3316	0.2985	0.933
	200	0.05	−0.0014	0.0645	0.0613	0.944	0.0040	0.1335	0.1323	0.941
		0.1	0.0023	0.0641	0.0631	0.939	0.0077	0.1420	0.1418	0.953
		0.5	0.0014	0.0727	0.0701	0.937	0.0365	0.2171	0.2084	0.938
	300	0.05	0.0031	0.0507	0.0500	0.948	0.0059	0.1051	0.1078	0.963
		0.1	−0.0005	0.0537	0.0515	0.945	0.0092	0.1144	0.1157	0.946
		0.5	0.0007	0.0571	0.0574	0.948	0.0175	0.1632	0.1675	0.951
	500	0.05	0.0011	0.0384	0.0388	0.951	0.0040	0.0853	0.0837	0.943
		0.1	0.0001	0.0413	0.0399	0.946	0.0061	0.0889	0.0897	0.954
		0.5	0.0010	0.0431	0.0445	0.952	0.0084	0.1243	0.1286	0.950

4. APPLICATION

In this section we apply the methodology proposed in the previous sections to a set of recurrent event data arising from a double-blinded clinical trial of chronic granulomatous disease (CGD) patients, where the CGD is a group of inherited rare disorders of the immune function characterised by recurrent pyogenic infections which usually occur early in life and may lead to death in childhood. The trial involving two treatment groups, placebo or gamma interferon, was conducted by the International CGD Cooperative Study Group and consists of 128 eligible patients recruited between October 1988 and March 1989. The goal of the study is to investigate the ability of gamma interferon to reduce the rate of serious infections requiring hospitalisation. The data set includes the dates of randomisation and each serious infection during the follow-up period for each patient. In total, 30 of the 65 patients in the placebo group and 14 of 63 in the gamma interferon group had experienced at least one serious infection. The data are given in Appendix D of Fleming & Harrington (1991) and were analysed by Lin et al. (2000) and Sun & Su (2008) among others.

TABLE 4: Simulation results for estimation of γ with $\mu_0(t) = 0.4(t^2 + 2t)$.

		$\beta_1 = 0$				$\beta_1 = 0.5$			
n	γ	Bias	SSE	ESE	CP	Bias	SSE	ESE	CP
100	0.05	0.0123	0.0942	0.0911	0.953	0.0002	0.0811	0.0782	0.938
	0.1	0.0135	0.1106	0.1082	0.954	0.0049	0.0935	0.0937	0.956
	0.5	0.0692	0.4112	0.3356	0.938	0.0454	0.3431	0.2738	0.933
200	0.05	0.0026	0.0646	0.0634	0.950	0.0011	0.0553	0.0548	0.955
	0.1	0.0037	0.0776	0.0756	0.943	0.0022	0.0646	0.0657	0.955
	0.5	0.0416	0.2487	0.2187	0.945	0.0319	0.2086	0.1864	0.945
300	0.05	0.0011	0.0516	0.0514	0.950	0.0018	0.0437	0.0446	0.952
	0.1	−0.0012	0.0619	0.0615	0.951	0.0032	0.0527	0.0531	0.952
	0.5	0.0090	0.1876	0.1713	0.927	0.0151	0.1527	0.1476	0.947
500	0.05	−0.0000	0.0397	0.0396	0.946	0.0010	0.0342	0.0345	0.957
	0.1	0.0021	0.0467	0.0474	0.958	0.0029	0.0403	0.0412	0.955
	0.5	0.0129	0.1301	0.1307	0.948	0.0056	0.1120	0.1126	0.952

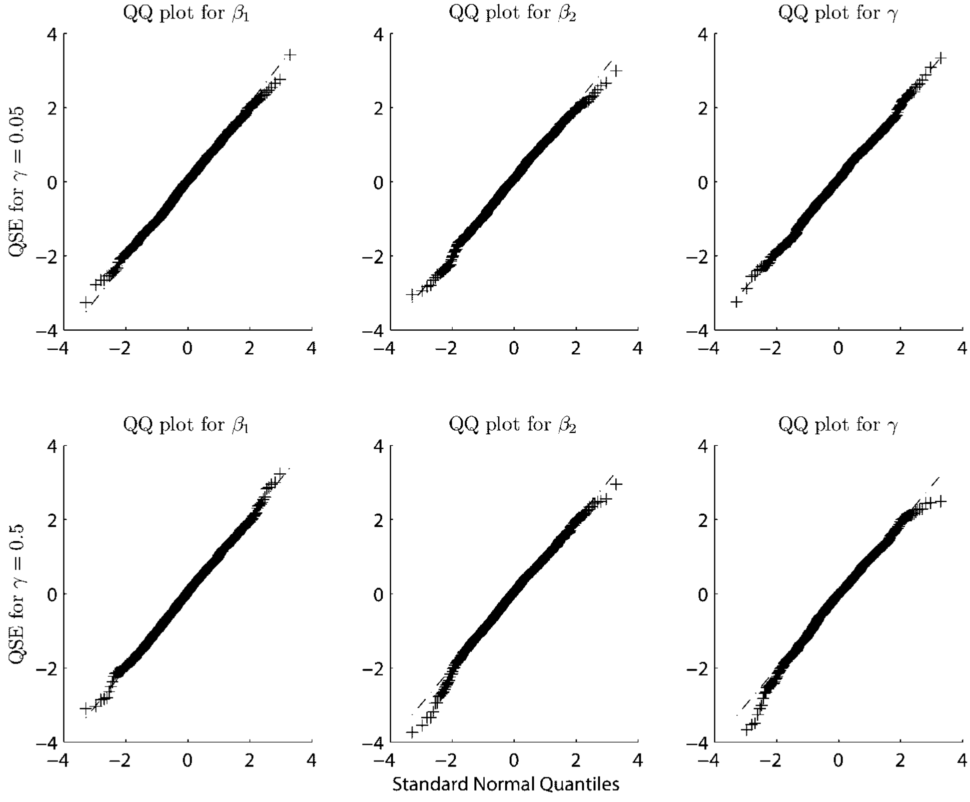

FIGURE 1: QQ-plots with $n = 300$, $\beta_1 = 0.5$, and $\mu_0(t) = t$. QSE stands for quantile of standardised estimates.

TABLE 5: Comparison of the proposed method with the MPLE when $\gamma = 0$ and $\mu_0(t) = t$.

			Our method				MPLE			
β_1	n	Parameter	Bias	SSE	ESE	CP	Bias	SSE	ESE	CP
0	100	β_1	0.0002	0.0903	0.0908	0.952	0.0002	0.0903	0.0908	0.952
		β_2	0.0208	0.1968	0.1975	0.941	0.0018	0.1044	0.1017	0.937
		γ	0.0039	0.0716	0.0724	0.954	–	–	–	–
	200	β_1	−0.0013	0.0649	0.0642	0.956	−0.0013	0.0648	0.0642	0.956
		β_2	0.0186	0.1411	0.1402	0.957	0.0032	0.0734	0.0726	0.948
		γ	0.0043	0.0513	0.0507	0.955	–	–	–	–
	300	β_1	0.0002	0.0509	0.0524	0.962	0.0002	0.0509	0.0524	0.962
		β_2	0.0086	0.1143	0.1135	0.954	0.0003	0.0605	0.0591	0.950
		γ	0.0021	0.0420	0.0412	0.947	–	–	–	–
	500	β_1	0.0017	0.0403	0.0407	0.946	0.0018	0.0403	0.0407	0.946
		β_2	0.0018	0.0887	0.0879	0.955	−0.0007	0.0461	0.0459	0.944
		γ	0.0002	0.0325	0.0320	0.940	–	–	–	–
0.5	100	β_1	0.0016	0.0820	0.0812	0.939	0.0017	0.0821	0.0811	0.939
		β_2	0.0153	0.1842	0.1714	0.924	0.0002	0.0922	0.0885	0.931
		γ	0.0025	0.0661	0.0629	0.942	–	–	–	–
	200	β_1	0.0016	0.0585	0.0575	0.947	0.0015	0.0585	0.0575	0.947
		β_2	0.0031	0.1215	0.1203	0.938	−0.0017	0.0641	0.0627	0.947
		γ	0.0005	0.0435	0.0438	0.953	–	–	–	–
	300	β_1	0.0006	0.0483	0.0469	0.943	0.0006	0.0484	0.0469	0.942
		β_2	0.0039	0.0991	0.0984	0.947	−0.0009	0.0542	0.0513	0.943
		γ	0.0010	0.0353	0.0358	0.961	–	–	–	–
	500	β_1	0.0006	0.0366	0.0364	0.950	0.0006	0.0366	0.0364	0.951
		β_2	0.0010	0.0778	0.0763	0.938	0.0000	0.0402	0.0398	0.948
		γ	−0.0002	0.0279	0.0276	0.949	–	–	–	–

For the analysis, we consider fitting model (1) with $Z = (Z_1, Z_2)'$ for the data, where Z_1 denotes the patient's age in year at enrolment, and Z_2 is the treatment indicator, taking value 1 for the patients who received gamma interferon. This model would be useful for predicting the experience of infection for patients with specific age and treatment assignments. Let τ be the largest follow-up time. The application of the proposed method gives estimates of the covariate coefficients as $\hat{\beta}_1 = -0.0303$ and $\hat{\beta}_2 = -1.4560$ with the estimated standard errors of 0.0143 and 0.5429, respectively. Both covariate effects are significant at 5% level, and the results are consistent with those obtained by Fleming & Harrington (1991) and Lin et al. (2000).

To give a more concrete interpretation, we also obtained the estimate of the convergence rate parameter as $\hat{\gamma} = 0.5245$ with the estimated standard error 0.6646, which implies a high degree of uncertainty, and the scale parameter γ is not significantly different from 0. This result suggests that the proportional rates model is appropriate for this set of data. Sun & Su (2008) also analysed the same CGD data using a class of accelerated means regression models. Their results showed that the covariates have no effect of time-scale change, and the proportional effect is highly significant

TABLE 6: Estimated sizes and powers of the Wald test for testing $\gamma = 0$.

		$\mu_0(t) = t$				$\mu_0(t) = 0.4(t^2 + 2t)$			
β_1	n	$\gamma = 0$	$\gamma = 0.1$	$\gamma = 0.3$	$\gamma = 0.5$	$\gamma = 0$	$\gamma = 0.1$	$\gamma = 0.3$	$\gamma = 0.5$
0	100	0.055	0.165	0.384	0.381	0.047	0.156	0.345	0.362
	200	0.041	0.272	0.712	0.795	0.044	0.248	0.666	0.792
	300	0.037	0.412	0.859	0.943	0.043	0.359	0.861	0.924
	500	0.052	0.603	0.980	0.998	0.058	0.587	0.970	0.995
0.5	100	0.054	0.198	0.505	0.596	0.053	0.171	0.477	0.521
	200	0.052	0.349	0.825	0.904	0.066	0.320	0.815	0.891
	300	0.044	0.480	0.952	0.983	0.046	0.516	0.944	0.974
	500	0.041	0.708	0.996	1.000	0.059	0.717	0.995	1.000

on the mean function of recurrent infections. These findings are consistent with those obtained by our proposed method.

To assess the overall fit of model (1) for the CGD data, following Lin et al. (2000), we consider the following cumulative sum of residuals:

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

where the event $I(Z_i \le z)$ means that each component of Z_i is bounded above by the corresponding component of z. Generate n independent standard normal random variables $G_1, \ldots, G_n$, independent of the data. It can be shown that the distribution of the process $\mathcal{F}(t, z)$ can be approximated by that of the zero-mean Gaussian process

$$\hat{\mathcal{F}}(t, z) = n^{-1/2} \sum_{i=1}^{n} \hat{\Psi}_i(t, z)G_i,$$

where

$$\hat{\Psi}_i(t, z) = \int_0^t \left\{ I(Z_i \le z) - \hat{D}_i(u)^{-1} \frac{\hat{S}(u, z)}{\hat{\pi}(u)} \right\} \mathrm{d}\hat{M}_i(u)$$
$$- \hat{B}(t, z)' \hat{A}^{-1} \int_0^\tau \left\{ X_i(u; \hat{\theta}) - \bar{X}(u; \hat{\theta}) + \hat{D}_i(u)^{-1} \hat{\mu}(u) \right\} \mathrm{d}\hat{M}_i(u),$$

$$\hat{S}(u, z) = n^{-1} \sum_{i=1}^{n} \left\{ I(Z_i \le z) \left[Y_i(u)\hat{D}_i(u) - \hat{\gamma}\hat{P}(u; \hat{\beta}_1, \hat{\gamma}) \right.\right.$$
$$\left.\left. \times \int_u^t \frac{1}{\hat{P}(v; \hat{\beta}_1, \hat{\gamma})} Y_i(v)\hat{D}_i(v)\{\hat{D}_i(v)\, \mathrm{d}\hat{R}(v) - \mathrm{d}\hat{\Lambda}(v; \hat{\beta}_1, 0)\} \right] \right\},$$

and

$$\hat{B}(t,z) = n^{-1}\sum_{i=1}^{n}\left\{\int_0^t I(Z_i \le z)Y_i(u)\hat{D}_i(u)\left[X_i(u;\hat{\theta})\,\mathrm{d}\hat{R}(u) + \mathrm{d}\hat{\Phi}(u;\hat{\theta})\right]\right\}$$

in notation of Section 2.

For the CGD data, we obtained $\sup_{0\le t\le\tau,z} |\mathcal{F}(t,z)| = 0.7040$ with a p-value of 0.375 based on 1,000 realisations of the statistic $\sup_{0\le t\le\tau,z} |\tilde{\mathcal{F}}(t,z)|$, which can be obtained by repeatedly generating the standard normal random sample $(G_1,\ldots,G_n)$ while fixing the data $\{C_i, N_i(\cdot), Z_i\}$ $(i = 1,\ldots,n)$ at their observed values. This result indicates that model (1) fits the data adequately.

5. CONCLUDING REMARKS

In this article we have studied a class of mixed models for recurrent event data. A key advantage of the proposed modeling approach is that the dependence structure of recurrent event times is left unspecified and it allows for both the proportional and convergent covariate effects on the rate function of the recurrent event process. Using the partial-score function, an estimation procedure was proposed for the model parameters, and the asymptotic properties of the resulting estimators were established. The numerical studies showed that the proposed methods work well for practical situations.

An important application of the proposed approach is checking the assumption that effects of covariates on recurrent event times are proportional or convergent. If the estimate of γ is close to zero, then it is reasonable to proceed on Z_2 with a proportional effect. Under such situation, the proposed method and the maximum partial likelihood method provide reasonable estimates for β. However, if the estimated γ is significantly different from zero, then it will be desirable to use the proposed mixed models instead of the proportional rates models.

In the mixed model, there is a problem with the choice of Z_1 and Z_2. The backward selection method presented in Barker & Henderson (2004) can be applied by replacing the likelihood with the pseudo-partial likelihood. If covariates are of small dimension, we can identify proportional and converging covariate effects on recurrent event times by fitting different models.

Note that our proposed method relies on the assumption that the recurrent event and censoring processes are independent given covariates. In some applications, however, this noninformative censoring assumption might be violated, especially if censoring is induced by informative dropouts and/or failure events. One possible way to adjust the method for such dependent censoring is to model the censoring mechanism as in Liu, Wolfe, & Huang (2004). It would be interesting to extend the proposed inference procedures to this situation.

ACKNOWLEDGEMENTS

The authors thank Editor-in-Chief, Professor Jiahua Chen, the Associate Editor, and two referees for their insightful comments and suggestions that greatly improved the article. The first author's research was partly supported by the National Natural Science Foundation of China Grants (No. 10731010, 10971015, and 11021161), the National Basic Research Program of China (973 Program) (No. 2007CB814902), and Key Laboratory of RCSDS, CAS (No. 2008DP173182). The second author's research was partly supported by the Research Grant Council of Hong Kong (PolyU504011).

BIBLIOGRAPHY

Aalen, O. O., Borgan, Ø., & Gjessing, H. K. (2008). *Survival and Event History Analysis: A Process Point of View*, Springer, New York.

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

Barker, P. & Henderson, R. (2004). Modelling coverging hazards in survival analysis. *Lifetime Data Analysis,* 10, 263–281.

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

Fleming, T. R. & Harrington, D. P. (1991). *Counting Process and Survival Analysis*, Wiley, New York.

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

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

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

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

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

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

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

Liu, Y., Wu, Y., Cai, J., & Zhou, H. (2010). Additive–multiplicative rates model for recurrent events. *Lifetime Data Analysis,* 16, 353–373.

Pepe, M. S. & 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.

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

Sun, L. & Su, B. (2008). A class of accelerated means regression models for recurrent event data. *Lifetime Data Analysis,* 14, 357–375.

Sun, L., Tong, X., & Zhou, X. (2011). A class of Box–Cox transformation models for recurrent event data. *Lifetime Data Analysis,* 17, 280–301.

Yang, S. (1992). Some inequalities about the Kaplan–Meier estimator. *Annals of Statistics,* 20, 535–544.

Zeng, D. & Cai, J. (2010). Semiparametric additive rate model for recurrent events with informative terminal event. *Biometrika,* 97, 699–712.

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

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

Received 17 March 2011
Accepted 4 September 2011