Lifetime Data Anal (2008) 14:357–375
DOI 10.1007/s10985-008-9087-z

A class of accelerated means regression models for recurrent event data

Liuquan Sun · Bin Su

Received: 30 August 2007 / Accepted: 19 May 2008 / Published online: 1 June 2008

Abstract In this article, we propose a general class of accelerated means regression models for recurrent event data. The class includes the proportional means model, the accelerated failure time model and the accelerated rates model as special cases. The new model offers great flexibility in formulating the effects of covariates on the mean functions of counting processes while leaving the stochastic structure completely unspecified. For the inference on the model parameters, estimating equation approaches are developed and both large and final sample properties of the proposed estimators are established. In addition, some graphical and numerical procedures are presented for model checking. An illustration with multiple-infection data from a clinic study on chronic granulomatous disease is also provided.

Keywords Counting process · Marginal model · Model checking · Semiparametric model · Recurrent events

1 Introduction

In many research settings, the event of interest can be experienced more than once per subject. Such outcomes have been termed recurrent events, which are commonly encountered in longitudinal follow-up studies. Medical examples of recurrent events are multiple infection episodes and tumor recurrences. Other examples include repeated breakdowns of a certain machinery in reliability testing and repeated purchases

L. Sun (✉) · B. Su
Institute of Applied Mathematics, Academy of Mathematics and Systems Science, Chinese Academy of Sciences, Beijing 100080, People's Republic of China
e-mail: slq@amt.ac.cn

B. Su
e-mail: subin@nus.edu.sg

of a certain product in marketing research. The structure of recurrent events is that of naturally ordered multivariate failure time data since different events within an individual are correlated. Correspondingly, the analysis of recurrent event data has recently been the subject of much methodological research, and the investigators are often interested in assessing the effects of covariates on certain features of the recurrent events process.

For recurrent event data, there are several estimating procedures proposed in the survival analysis literature, including conditional models (Andersen and Gill 1982; Prentice et al. 1981) and marginal models (Wei et al. 1989). Following the traditional development of survival analysis, these methods are based on modeling the intensity and hazard functions. Because the mean number of events is more interpretable quantity than the hazard in the context of recurrent event data, some authors have proposed to model the mean and rate functions (Pepe and Cai 1993; Lawless and Nadeau 1995; Lin et al. 2000, 2001). For example, Lawless and Nadeau (1995) have proposed a class of marginal means models, and Lin et al. (2000) studied the proportional means and rates models for counting processes.

One class of models which have been developed in many contexts is time-transformation models, in which all subjects have similar trajectories and the effect of covariates is to alter the time scale of the trajectories. For example, Lin et al. (1998) provided the accelerated failure time model to formulate the effects of covariates on the mean function of the counting process for recurrent events. Ghosh (2004) presented the accelerated rates model for counting processes in which the effect of covariates is to transform the time scale for a baseline rate function. In this paper, we propose a more general class of accelerated means regression models for recurrent event data, which includes the proportional means model, the accelerated failure time model and the accelerated rates model as special cases. The class are related to a class of semi-parametric hazards regression models for univariate survival data studied by Chen and Jewell (2001). In the new model, a covariate's effect is identified as having two separate components, namely a time-scale change and a relative ratio for the mean function of the counting process.

The remainder of the paper is organized as follows. In Sect. 2, we present a semi-parametric formulation of the general model, and proposes an estimating procedure for the model parameters. The asymptotic properties of the proposed estimators are established. In Sect. 3, we develop a technique for checking the adequacy of the general model. Section 4 reports some results from simulation studies conducted for evaluating the proposed methods. In Sect. 5, the methodology is applied to a data set from a clinic study on chronic granulomatous disease, followed by concluding remarks in Sect. 6. The details of the proofs are relegated to Appendix.

2 Model and estimation procedure

Suppose that a total of n subjects are observed over time. Let $N_i^*(t)$ be the number of events that occur over the interval $[0, t]$ for subject i, and Z_i be a p-vector of covariates. In most applications, the subject is followed for a limited period of time so $N_i^*(\cdot)$ is not fully observed. Let C_i denote the follow-up or censoring time. Assume that C_i is independent of $N_i^*(\cdot)$ conditional on Z_i. Define $N_i(t) = N_i^*(t \wedge C_i)$ and $Y_i(t) = I(C_i \geq t)$, where $a \wedge b = \min(a, b)$, and $I(\cdot)$ is the indicator function. The observable data consist of $\{N_i(\cdot), Y_i(\cdot), Z_i\}$ $(i = 1, \ldots, n)$.

The proposed accelerated means regression model takes the form

$$E\{N_i^*(t)|Z_i\} = \mu_0(te^{\beta_{10}' Z_i})g(\beta_{20}' Z_i), \tag{1}$$

where $\beta_{10}(t)$ and β_{20} are p-vectors of unknown regression parameters, and $\mu_0(t)$ is an unspecified baseline mean function. The link function $g(\cdot)$ is pre-specified and twice continuously differentiable with $g(\cdot) \geq 0$. Clearly, model (1) becomes the proportional means model when $\beta_{10} = 0$ and $g(x) = e^x$. The choice of $g(x) = 1$ yields the accelerated failure time model for counting processes. When $g(x) = e^x$ and $\beta_{20} = -\beta_{10}$, model (1) reduces to the accelerated rates regression model for recurrent events. If $N^*(t)$ is a simple counting process (i.e., can only take a value of 0 or 1), model (1) is equivalent to a general hazards regression model studied by Chen and Jewell (2001) for univariate survival data.

Flexibility of the general model may lead to concern of identifiability. Following the arguments of Chen and Jewell (2001), we address the issue of identifiability in the following proposition.

Proposition 1 *Under model (1), if there exist a sequence of constants* $\{c_k\}_{k=-\infty}^{+\infty}$ *and a large enough* $t_0 > 0$ *such that* $\mu_0(t) = \sum_{k=-\infty}^{+\infty} c_k t^k$, *for any* $t \in [0, t_0]$, *and* $g(\cdot)$ *is a strictly monotone function, then* β_{10} *and* β_{20} *are identifiable if and only if there exist* $k_1, k_2 \in \{0, \pm 1, \pm 2, \ldots\}$ *such that* $k_1 \neq k_2$ *and* $c_{k_1} c_{k_2} \neq 0$.

Proposition 1 implies that the model is not identifiable if and only if the baseline mean function is in the form of $c_k t^k$ for some k, in which all three subclasses of special models and the general model (1) coincide. We assume that β_{10} and β_{20} are identifiable throughout the paper.

The two parameters β_{10} and β_{20} can be interpreted as measuring two different effects the covariate may have on the mean function of the recurrent event process. For example, suppose that Z_i is binary, and $Z_i = 1$ stands for the treatment group and $Z_i = 0$ for the control group, as in a randomized clinical trial. The first parameter β_{10} identifies the acceleration or deceleration of the mean function process in the treatment group, which is called *time-scale effect*, while β_{20} characterizes the relative ratio after adjusting for the different mean function process in the treatment and the control groups, which is called *relative-ratio effect*. Therefore, model (1) implies that the treatment can alter both the magnitude of the mean function and the frequency of recurrences simultaneously. Correctly identifying and estimating these two components may better describe a given recurrent event data, although the main value of the general model (1) may be in quantifying and highlighting differences between the three subclasses of models that are commonly used individually.

Let $\tilde{N}_i(t; \beta_1) = N_i(te^{-\beta_1' Z_i})$, and $Y_i(t; \beta_1) = I(C_i \geq te^{-\beta_1' Z_i})$. Define

$$M_i(t; \beta) = \tilde{N}_i(t; \beta_1) - \int_0^t Y_i(s; \beta_1) g(\beta_2' Z_i) d\mu_0(s),$$

where $\beta = (\beta_1', \beta_2')'$. Under model (1), $M_i(t; \beta_0)$ are zero-mean stochastic processes, where $\beta_0 = (\beta_{10}', \beta_{20}')'$. Thus, for given β, a reasonable estimator for $\mu_0(t)$ is the solution to

$$\sum_{i=1}^{n}\left[\tilde{N}_i(t; \beta_1) - \int_0^t Y_i(s; \beta_1) g(\beta_2' Z_i) d\mu(s)\right] = 0, \quad 0 \le t \le \tau,$$

where τ is a prespecified constant such that $P(C_i \ge \tau e^{-\beta_{10} Z_i}) > 0$. Denote this estimator by $\hat{\mu}_0(t; \beta)$, which can be expressed as

$$\hat{\mu}_0(t; \beta) = \sum_{i=1}^{n} \int_0^t \frac{d\tilde{N}_i(s; \beta_1)}{\sum_{j=1}^{n} Y_j(s; \beta_1) g(\beta_2' Z_j)}. \tag{2}$$

To estimate β_0, using the generalized estimating equation methods (Liang and Zeger 1986) and replacing $\mu_0(t)$ with its estimator obtained above, we specify the following two estimating functions for β_{10} and β_{20}:

$$U_1(\beta) = \sum_{i=1}^{n} \int_0^\tau \left\{Z_i - \bar{Z}(t; \beta)\right\} d\tilde{N}_i(t; \beta_1), \tag{3}$$

and

$$U_2(\beta) = \sum_{0}^{\tau} \int_0^\tau \left\{W(t, Z_i; \beta) - \bar{W}(t, Z_i; \beta)\right\} d\tilde{N}_i(t; \beta_1), \tag{4}$$

where $W(t, Z_i; \beta)$ is a known p-dimensional weight function of t, Z_i and β, not in the span of the functions 1 and Z_i,

$$\bar{Z}(t; \beta) = \frac{\sum_{i=1}^{n} Y_i(t; \beta_1) g(\beta_2' Z_i) Z_i}{\sum_{i=1}^{n} Y_i(t; \beta_1) g(\beta_2' Z_i)},$$

and

$$\bar{W}(t; \beta) = \frac{\sum_{i=1}^{n} Y_i(t; \beta_1) g(\beta_2' Z_i) W(t, Z_i; \beta)}{\sum_{i=1}^{n} Y_i(t; \beta_1) g(\beta_2' Z_i)}.$$

Let $Z_i^*(t; \beta) = (Z_i', W(t, Z_i; \beta)')'$, $\bar{Z}^*(t; \beta) = (\bar{Z}(t; \beta)', \bar{W}(t; \beta)')'$, and $U(\beta) = (U_1(\beta)', \ U_2(\beta)')'$. Then (3) and (4) can be written as

$$U(\beta) = \sum_{i=1}^{n} \int_0^\tau \left\{Z_i^*(t; \beta) - \bar{Z}^*(t; \beta)\right\} d\tilde{N}_i(t; \beta_1). \tag{5}$$

Since $U(\beta)$ is a discrete function of β_1, we define the estimator $\hat{\beta} = (\hat{\beta}_1', \hat{\beta}_2')'$ as a zero-crossing of $U(\beta)$ or as a minimiser of $\|U(\beta)\|$ (Lin et al. 1998; Chen and Jewell 2001; Ghosh 2004), where $\|v\| = (v'v)^{1/2}$ for a vector v. Various methods of solving this equation exist, such as direct grid search, the bisection method or the technique of simulated annealing (SA). When there are only a small number of covariates, direct grid search and the bisection method are recommended. For high-dimensional covariate vectors, the SA method may be more efficient, which is a generic probabilistic meta-algorithm for the global optimization problem, namely locating a good approximation to the global optimum of a given function in a large search space. Each step of the SA algorithm replaces the current solution by a random "nearby" solution, chosen with a probability that depends on the difference between the corresponding function values and on a global parameter, that is gradually decreased during the process (see, for example, Lin and Geyer 1992). When $\hat{\beta}$ is available, the baseline mean function $\mu_0(t)$ can be estimated by the Nelson–Aalen-type estimator $\hat{\mu}_0(t) \equiv \hat{\mu}_0(t; \hat{\beta})$ given by (2).

In general, to establish the asymptotic properties of $\hat{\beta}$, we need first to establish the asymptotic properties of $U(\beta_0)$. Under certain regularity conditions, we show in Theorem 1 of the Appendix that $n^{-1/2}U(\beta_0)$ is asymptotically normal with mean zero and covariance matrix that can be consistently estimated by $\hat{\Sigma}$, where

$$\hat{\Sigma} = n^{-1} \sum_{i=1}^{n} D_i(\hat{\beta}) D_i(\hat{\beta})', \tag{6}$$

$$D_i(\beta) = \int_0^{\tau} \left\{ Z_i^*(t; \beta) - \bar{Z}^*(t; \beta) \right\} d\hat{M}_i(t; \beta),$$

and

$$\hat{M}_i(t; \beta) = \tilde{N}_i(t; \beta_1) - \int_0^t Y_i(t; \beta_1) g(\beta_2' Z_i) d\hat{\mu}_0(t; \beta).$$

In the following, define

$$S^{(0)}(t; \beta) = n^{-1} \sum_{i=1}^{n} Y_i(t; \beta_1) g(\beta_2' Z_i),$$

$$S^{(1)}(t; \beta) = n^{-1} \sum_{i=1}^{n} Y_i(t; \beta_1) \dot{g}(\beta_2' Z_i) Z_i,$$

$$S_z^{(2)}(t; \beta) = n^{-1} \sum_{i=1}^{n} Y_i(t; \beta_1) g(\beta_2' Z_i) Z_i^{\otimes 2},$$

$$S_z^{(3)}(t; \beta) = n^{-1} \sum_{i=1}^{n} Y_i(t; \beta_1) \dot{g}(\beta_2' Z_i) Z_i^{\otimes 2},$$

$$S_w^{(2)}(t;\beta)=n^{-1}\sum_{i=1}^{n}Y_i(t;\beta_1)g(\beta_2'Z_i)W(t,Z_i;\beta)Z_i',$$

and

$$S_w^{(3)}(t;\beta)=n^{-1}\sum_{i=1}^{n}Y_i(t;\beta_1)\dot{g}(\beta_2'Z_i)W(t,Z_i;\beta)Z_i',$$

where $\dot{g}=dg(t)/dt$ and $v^{\otimes 2}=vv'$ for a vector v.

We also prove in Theorem 2 of the Appendix that $n^{1/2}(\hat{\beta}-\beta_0)$ is asymptotically normal with mean zero and covariance matrix that can be consistently estimated by $\hat{A}^{-1}\hat{\Sigma}(\hat{A}^{-1})'$, where

$$\begin{aligned}
\hat{A}&=\begin{pmatrix}\hat{A}_{11} & \hat{A}_{12}\\ \hat{A}_{21} & \hat{A}_{22}\end{pmatrix},\\
\hat{A}_{11}&=\int_0^{\tau}\left[S_z^{(2)}(t;\hat{\beta})-\bar{Z}(t;\hat{\beta})^{\otimes 2}S^{(0)}(t;\hat{\beta})\right]d\{\hat{\lambda}_0(t)t\},\\
\hat{A}_{12}&=\int_0^{\tau}\left[S_z^{(3)}(t;\hat{\beta})-\bar{Z}(t;\hat{\beta})S^{(1)}(t;\hat{\beta})'\right]d\hat{\mu}_0(t),\\
\hat{A}_{21}&=\int_0^{\tau}\left[S_w^{(2)}(t;\hat{\beta})-\bar{W}(t;\hat{\beta})\bar{Z}(t;\hat{\beta})S^{(0)}(t;\hat{\beta})'\right]d\{\hat{\lambda}_0(t)t\},\\
\hat{A}_{22}&=\int_0^{\tau}\left[S_w^{(3)}(t;\hat{\beta})-\bar{W}(t;\hat{\beta})S^{(1)}(t;\hat{\beta})'\right]d\hat{\mu}_0(t),
\end{aligned}\tag{7}$$

and $\hat{\lambda}_0(t)$ is a density-type estimator of $\lambda_0(t)=d\mu_0(t)/dt$, that is,

$$\hat{\lambda}_0(t)=h^{-1}\int K\left(\frac{u-t}{h}\right)d\hat{\mu}_0(u),$$

where h is the bandwidth and $K(\cdot)$ is a kernel function with a compact support (e.g., Ramlau-Hansen 1983).

Based on the above result, we see that the estimator $\hat{A}$ requires an estimate for the derivative of $\mu_0(t)$. Because such density-type estimator $\hat{\lambda}_0(t)$ tends to be numerically unstable, the resulting variance estimator for $\hat{\beta}$ will also be unreliable. In order to obtain a stable variance estimate of $\hat{\beta}$, we adapt a resampling technique due to Parzen et al. (1994). Specifically, let $\hat{\beta}^*$ be the solution to

$$U(\beta)=\sum_{i=1}^{n}D_i(\hat{\beta})G_i,\tag{8}$$

where $\{G_1, \ldots, G_n\}$ are independent standard normal variables. By the arguments of Parzen et al. (1994) and Lin et al. (1998), the asymptotic distribution of $n^{1/2}(\hat{\beta} - \beta_0)$ can be approximated by the conditional distribution of $n^{1/2}(\hat{\beta}^* - \hat{\beta})$ given the data $\{N_i(\cdot), Y_i(\cdot), Z_i\}(i = 1, \ldots, n)$. To approximate the distribution of $\hat{\beta}$, we produce a large number of realisations of $\hat{\beta}^*$ by repeatedly generating the random samples $(G_1, \ldots, G_n)$ while fixing the data $\{N_i(\cdot), Y_i(\cdot), Z_i\}(i = 1, \ldots, n)$ at their observed values. The covariance matrix of $\hat{\beta}$ can then be approximated by the empirical covariance matrix of $\hat{\beta}^*$. Hence, confidence intervals for β_0 can be constructed using the empirical distribution of $\hat{\beta}^*$.

Let $V(t) = n^{1/2}\{\hat{\mu}_0(t) - \mu_0(t)\}$. We show in Theorem 3 of the Appendix that $V(t)$ converges weakly to a zero-mean Gaussian process whose covariance function at (s, t) can be consistently estimated by $\hat{\Gamma}(s, t) = n^{-1} \sum_{i=1}^{n} \hat{\Psi}_i(s)\hat{\Psi}_i(t)$, where

$$\begin{aligned} \hat{\Psi}_i(t) &= \int_0^t \frac{d\hat{M}_i(u; \hat{\beta})}{S^{(0)}(u; \hat{\beta})} - \hat{H}(t)'\hat{A}^{-1} \int_0^{\tau} \left\{ Z_i^*(u; \hat{\beta}) - \bar{Z}^*(u; \hat{\beta}) \right\} d\hat{M}_i(u; \hat{\beta}), \\ \hat{H}_1(t) &= \int_0^t \bar{Z}(u; \hat{\beta}) d\{\hat{\lambda}_0(u)u\}, \qquad (9) \\ \hat{H}_2(t) &= \int_0^t \frac{S^{(1)}(u; \beta)}{S^{(0)}(u; \beta)} d\hat{\mu}_0(u), \end{aligned}$$

and $\hat{H}(t) = (\hat{H}_1(t)', \hat{H}_2(t)')'$.

As in the case of $\hat{\beta}$, it is difficult to estimate the asymptotic covariance function of $V(t)$ analytically. Following Lin et al. (1998), we can show that the asymptotic distribution of $V(t)$ can be approximated by the conditional distribution of $\hat{V}(t)$, where

$$\hat{V}(t) = n^{1/2} \left\{ \hat{\mu}_0(t; \hat{\beta}) - \hat{\mu}_0(t; \hat{\beta}^*) \right\} + n^{-1/2} \sum_{i=1}^{n} \int_0^t \frac{d\hat{M}_i(u; \hat{\beta})}{S^{(0)}(u; \hat{\beta})} G_i,$$

where $\hat{\beta}^*$ is the solution to (8). Thus, we may use the simulated distribution of $\hat{V}(t)$ to make inference about $\mu_0(t)$ along the lines of Lin et al. (1998).

Remark Note that in order to approximate the distribution $\hat{\beta}$, we adapt a simple resampling method (Parzen et al. 1994), which is essentially a parametric bootstrap procedure. Of course, a nature approach following the work of Efron (1979) is simply to resample the data with replacement and create bootstrap analogues of $\hat{\beta}$, which is the (naive) bootstrap method. However, implementation of this type scheme is computationally slow and intensive. This is essentially due to the high dimension of the data and the use of the resampling mechanism to rebuild bootstrap replicates at each stage. In addition, there is no analytical proof that the naive bootstrap method is valid for recurrent event data. Theoretical justification of the existing resampling methods is generally nontrivial and has to be made on a case by case basis.

3 Goodness-of-fit tests

As with any regression model, it is important to develop goodness-of-fit methods for assessing the adequacy of model (1). Following Lin et al. (2000), we consider the following cumulative sums of residual:

$$F(t, z; \hat{\beta}) = n^{-1/2} \sum_{i=1}^{n} I(Z_i \leq z) \hat{M}_i(t; \hat{\beta}), \tag{10}$$

where the event $I(Z_i \leq z)$ means that each of the components of Z_i is no larger than the corresponding component of z. We show in Theorem 4 of the Appendix that the null distribution of $F(t, z; \hat{\beta})$ can be approximated by the zero-mean Gaussian process

$$\begin{aligned}\tilde{F}(t, z) = n^{-1/2} \sum_{i=1}^{n} \int_0^t \left\{ I(Z_i \leq z) - \frac{S(u, z; \hat{\beta})}{S^{(0)}(u; \hat{\beta})} \right\} d\hat{M}_i(u; \hat{\beta}) \\ -\hat{B}(t, z)' \hat{A}^{-1} n^{-1/2} \sum_{i=1}^{n} \int_0^\tau \left\{ Z_i^*(u; \hat{\beta}) - \bar{Z}^*(u; \hat{\beta}) \right\} d\hat{M}_i(u; \hat{\beta}),\end{aligned} \tag{11}$$

where

$$S(u, z; \beta) = n^{-1} \sum_{i=1}^{n} Y_i(u; \beta_1) g(\beta_2' Z_i) I(Z_i \leq z),$$

$$\hat{B}_1(t, z) = n^{-1} \sum_{i=1}^{n} \int_0^t Y_i(u; \hat{\beta}_1) g(\hat{\beta}_2' Z_i) I(Z_i \leq z) \left\{ Z_i - \bar{Z}(u; \hat{\beta}) \right\} d\{\hat{\lambda}_0(u) u\},$$

$$\hat{B}_2(t, z) = n^{-1} \sum_{i=1}^{n} \int_0^t Y_i(u; \hat{\beta}_1) \dot{g}(\hat{\beta}_2' Z_i) Z_i \left\{ I(Z_i \leq z) - \frac{S(u, z; \hat{\beta})}{S^{(0)}(u; \hat{\beta})} \right\} d\hat{\mu}_0(u),$$

and $\hat{B}(t, z) = (\hat{B}_1(t, z)', \hat{B}_2(t, z)')'$.

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

$$\hat{F}(t,z) = \left\{ F(t,z;\hat{\beta}) - F(t,z;\hat{\beta}^*) \right\} + n^{-1/2} \sum_{i=1}^{n} \int_0^t \left\{ I(Z_i \le z) - \frac{S(u,z;\hat{\beta})}{S_0(u;\hat{\beta})} \right\} d\hat{M}_i(u;\hat{\beta}) G_i.$$

Thus, to approximate the distribution of $F(t,z;\hat{\beta})$, one can obtain a large number of realizations from $\hat{F}(t,z)$, by repeatedly generating the standard normal random sample $(G_1,\ldots,G_n)$ while fixing the data $\{N_i(\cdot), Y_i(\cdot), Z_i\}$ $(i = 1,\ldots,n)$ at their observed values. To assess the overall fit of model (1), one can plot a few realizations from $\hat{F}(t,z)$ along with the observed $F(t,z;\hat{\beta})$ and see if they can be regarded as arising from the same population. More formally, we can apply the supremum test statistic $\sup_{0\le t\le\tau,z} |F(t,z;\hat{\beta})|$, with which the p-value can be obtained by comparing the observed value of $\sup_{0\le t\le\tau,z} |F(t,z;\hat{\beta})|$ to a large number of realizations from $\sup_{0\le t\le\tau,z} |\hat{F}(t,z)|$.

4 Simulation studies

Simulation studies were conducted to examine the finite sample properties of the proposed estimators. In the study, gap time between the recurrences were generated from the random-effect intensity model

$$\lambda(t|Z_i,\eta) = \eta\lambda_0(te^{\beta_1 Z_i})e^{\beta_1 Z_i} g(\beta_2 Z_i), \tag{12}$$

where $g(x) = \exp(x)$, $\lambda_0(t) = t^2$, η is a gamma random variable with mean 1 and variance σ^2, and Z_i is a Bernoulli random variable with success probability 0.5. The follow-up time was generated from the uniform distribution $U(0, 3.5)$, which yielded an average of approximately two observed events per subject. Simple integration implies that model (12) is of the form (1). When $\sigma^2 = 0$, recurrent events within an subject are independent. On the other hand, nonzero values of σ^2 induce correlation between recurrences.

For each simulation study, we considered $\sigma^2 = 0, 0.25, 0.5$ and 1. The weight function $W(t, Z_i;\beta)$ was taken as the Gehan weight function

$$W(t, Z_i;\beta) = n^{-1} \sum_j Y_j(t;\beta_1) g(\beta_2 Z_j) Z_i.$$

For each simulation setting, 1,000 simulation samples were considered, and 1,000 resamplings were generated for each simulation sample. We took τ as the largest recurrence time so that all data were used in the analysis. The estimates of β_0 were obtained by the bisection method with the accuracy of ± 0.0001.

Table 1 presents the simulation results on estimation of $\beta_0 = (\beta_{10}, \beta_{20}) = (-0.5, 0.5)$ with the sample sizes $n = 50$ and 100. The table includes the biases (Bias) given by the sample means of the point estimates $\hat{\beta}$ minus the true value, the sampling means (SEE) of the estimated standard errors of $\hat{\beta}$, the sampling standard errors (SSE) of

Table 1 Monte carlo simulation results

Num	σ^2	$\beta_{10} = -0.5$				$\beta_{20} = 0.5$			
		Bias	SSE	SEE	CP	Bias	SSE	SEE	CP
50	0.00	−0.002	0.068	0.068	0.946	−0.003	0.080	0.079	0.951
50	0.25	0.000	0.085	0.084	0.945	0.002	0.096	0.094	0.947
50	0.50	−0.004	0.098	0.095	0.942	−0.009	0.107	0.105	0.945
50	1.00	0.005	0.117	0.113	0.937	−0.013	0.129	0.124	0.938
100	0.00	0.001	0.044	0.043	0.949	0.008	0.049	0.048	0.952
100	0.25	−0.007	0.067	0.064	0.947	0.009	0.078	0.075	0.947
100	0.50	0.008	0.070	0.068	0.944	−0.010	0.082	0.081	0.946
100	1.00	−0.008	0.076	0.074	0.938	−0.007	0.089	0.085	0.940

$\hat{\beta}$, and the 95% empirical coverage probabilities (CP) for β_0 based on the empirical distribution of β^*. It can be seen from Table 1 that the proposed estimation procedures performed well for the situations considered here. Specifically, the proposed estimators are practically unbiased, and both the variance estimation and coverage probabilities seem reasonable.

5 An example of application

We now apply the proposed method to the multiple-infection data taken from the CGD study presented by Fleming and Harrington (1991) and Lin et al. (2000). CGD is a heterogeneous group of uncommon inherited disorders of the immune function characterized by recurrent pyogenic infections. In order to assess the efficacy of gamma interferon in reducing the rate of infections, a double-blinded clinical trial was conducted in which patients were randomized to either placebo or gamma interferon group. A total of 128 patients were enrolled into the study. By the end of the trial, 30 of the 65 patients in the placebo group and 14 of 63 in the gamma interferon group had experienced at least one infection. The full data set appears in the Appendix D.2 of Fleming and Harrington (1991). Lin et al. (2000) also analyzed the data using compared multiple endpoint Cox models to analyze these data. Although the data set contains several covariates, but for the illustration purpose, we focus on the effects of treatment and the patients' age on the rate of infections (Lin et al. 2000; Ghosh 2004).

For the analysis, we defined Z_{i1} as the treatment indicator, which take the value 1 if the subject received gamma interferon and Z_{i2} to be the patients's age at enrollment. Let τ be the largest observed infection time. The bisection method with the accuracy of ± 0.00001 was employed to find the estimates of the regression parameters, and 1,000 resamplings were used to approximate the distribution $\hat{\beta}$. Firstly, we only used the treatment indicator Z_{i1} in model (1) with $g(x) = \exp(x)$, and the results are shown in model A of Table 2. In model A, the treatment covariate has no effect of time-scale change, but has the proportional effect on the mean function of recurrent infections, indicating that gamma interferon is effective in reducing the mean of infection. Next, we added the age covariate into the model, and the results are presented in model B of Table 2. This also suggests that the covariates have no effect of time-scale change,

Table 2 Estimation of the effect for the CGD study

	Estimation	Model A	Model B	
		Treatment only	Treatment	Age
$\hat{\beta}_{10}$	Coefficient	−0.0005	0.0021	−0.0049
	Standard error	0.0016	0.0025	0.0102
	p-Value	0.756	0.401	0.632
$\hat{\beta}_{20}$	Coefficient	−1.080	−1.1044	−0.0204
	Standard error	0.3068	0.3478	0.0077
	p-Value	0.004	0.002	0.008

and the proportional effect is highly significant on the mean function of recurrent infections. In addition, The above results are close to that of Lin et al. (2000) using the proportional means model.

Using the results of model B in the Table 2, Fig. 1 displays the estimates of the cumulative frequencies of recurrent infections for 18-year-old patients who received gamma interferon versus who did not receive gamma interferon. The estimates are shown from day 4 to day 373, which are respectively the smallest and largest infection

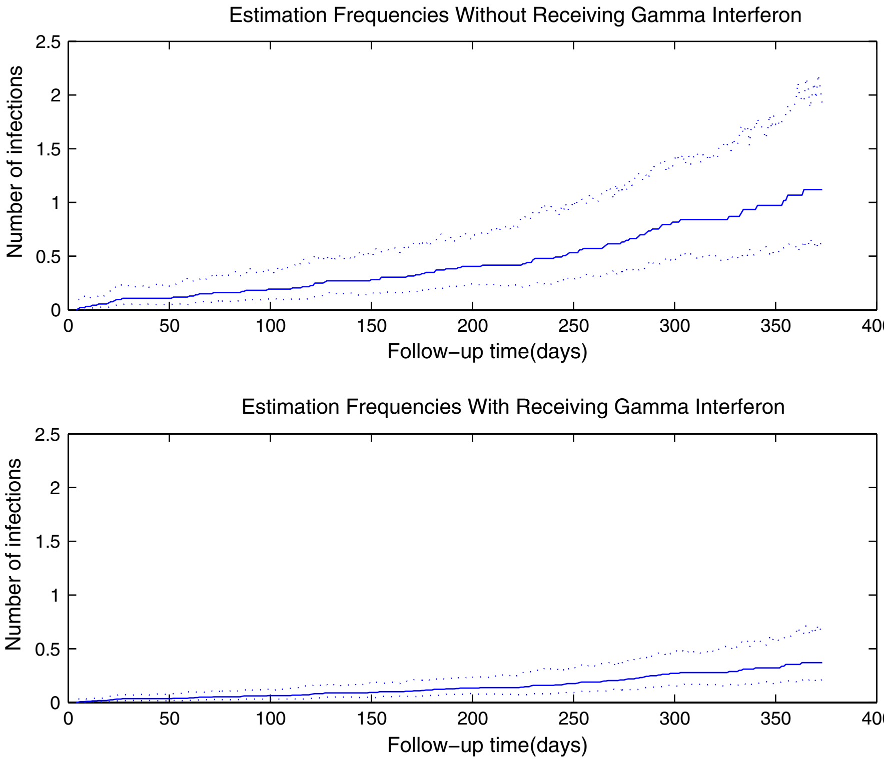

Fig. 1 Estimation of the cumulative frequencies of infections for 18-year-old CGD patients

times observed in the data set. The 95% simultaneous confidence bands are based on the 1,000 simulated realizations of $\hat{V}(\cdot)$. Again, the treated patients tend to have fewer infection episodes over time.

Now we consider the application of the model checking procedures given in Sect. 3 to the data. We only show the analysis for model B with both covariates in the model. The p-value using the supremum test statistic $\sup_{0 \le t \le \tau, z} |F(t, z; \hat{\beta})|$ based on 1,000 realizations of $\sup_{0 \le t \le \tau, z} |\hat{F}(t, z)|$ is 0.659, which means that the model is appropriate for the CGD data set.

6 Concluding remarks

In this article we have studied a general class of accelerated means regression models for recurrent event data, which are flexible and include some commonly used models as special cases. An estimation procedure was proposed for the model parameters, and asymptotic properties of the estimators were derived. The methodology was applied to data from a clinic study on chronic granulomatous disease, and the simulation results show that the proposed methods work well for the situations considered.

One potential use of the general model (1) is to test if the proportional means model, the accelerated failure time model and the accelerated rates model can be identified from the data. For example, we can check the proportional means assumption by testing $\beta_{10} = 0$ with $g(x) = e^x$, the accelerated failure time assumption by testing $\beta_{20} = 0$, and the accelerated rates assumption by testing $\beta_{20} = -\beta_{10}$ with $g(x) = e^x$. These can be done by the Wald or Score-type statistics, and we are currently conducting more simulations to further address these issues both theoretically and numerically.

Our proposed method assumes that the censoring and event processes are independent conditional on covariates. However, in many applications, this noninformative censoring assumption might not hold, especially when censoring could be caused by informative dropouts or failure events. It would be of interest to extend procedures for a general class of accelerated means regression models to handle these problems.

Another limitation of the approach given here is that the covariates Z are time-invariant. In some applications, it might be of interest to use time-dependent covariates in the model. An obvious extension of model (1) in this case would be

$$E\{N_i^*(t)|Z_i(t)\} = \mu_0(te^{\beta_{10}' Z_i(t)})g(\beta_{20}' Z_i(t)).$$

However, the proposed estimation procedure cannot be extended in a straightforward manner to deal time-dependent covariates.

Since estimating functions (3) and (4) were given in a somewhat ad hoc fashion using the generalized estimating equation methods, it would be worthwhile to further investigate the efficiencies of the proposed estimators. If $N_i^*(\cdot)$ is a Poisson process, then it might be possible to estimate β_0 and $\mu_0(\cdot)$ more efficiently by the nonparametric maximum likelihood approach, and the resulting inference procedure would be much more complicated.

Note that in (4), there is a weight function that needs to be specified. We used Gehan weight function in the simulation studies, and there are many potential candidates for the weight function $W(t, Z_i; \beta)$. For instance, we may choose $W(t, Z_i; \beta) = tZ_i$ and $W(t, Z_i; \beta) = \{t/(1+t)\}Z_i$ as in Chen and Jewell (2000). Ideally, we would choose $W(t, Z_i; \beta)$ to minimize the variance of $\hat{\beta}$. However, it does not appear possible to derive an optimal weight without specification of dependence structure on the increments of $N^*(t)$, and the selection of weight functions is usually a complicated problem (Lin et al. 2001). Thus, as for further research, it would be useful to investigate how to find a weight function that gives the optimal efficiency of the proposed estimate of regression parameters if it exists.

Acknowledgments The authors would like to thank the Editor and two reviewers for their constructive and insightful comments and suggestions that greatly improved the paper. This research was partly supported by the National Natural Science Foundation of China Grants (No. 10571169 and 10731010) and the National Basic Research Program of China (973 Program) (No. 2007CB814902).

Appendix: Asymptotic properties of $U(\beta_0)$, $\hat{\beta}$, $\hat{\mu}_0(t)$ and $F(t, z; \hat{\beta})$

Let $s^{(0)}(t)$, $s^{(1)}(t)$, $s_z^{(2)}(t)$, $s_z^{(3)}(t)$, $s_w^{(2)}(t)$, $s_w^{(3)}(t)$, $\bar{z}(t)$ and $\bar{w}(t)$ be the limits of $S^{(0)}(t; \beta_0)$, $S^{(1)}(t; \beta_0)$, $S_z^{(2)}(t; \beta_0)$, $S_z^{(3)}(t; \beta_0)$, $S_w^{(2)}(t; \beta_0)$, $S_w^{(3)}(t; \beta_0)$, $\bar{Z}(t; \beta_0)$ and $\bar{W}(t; \beta_0)$, respectively.

We will use the same notation defined in the previous sections and assume that the following regularity conditions hold:

(C1) (N_i^*, C_i, Z_i) are independent and identically distributed for $i = 1, \ldots, n$.
(C2) $P(Y_i(\tau; \beta_{10}) = 1) > 0$.
(C3) $N_i(t)$, Z_i and $W(t, Z_i; \beta_0)$ are bounded on $[0, \tau]$ for $i = 1, \ldots, n$.
(C4) $g(\cdot)$ is twice continuously differentiable with $g(\cdot) \geq 0$, and $g(\beta_{20}' Z_i)$ is locally bounded away.
(C5) $C_i e^{\beta_{10} Z_i}$ has a bounded density and $\mu_0(t)$ has a bounded second derivative.
(C6) A is nonsingular, where

$$A = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix},$$

$$A_{11} = \int_0^\tau \left[s_z^{(2)}(t) - \bar{z}(t)^{\otimes 2} s^{(0)}(t) \right] d\{\lambda_0(t)t\},$$

$$A_{12} = \int_0^\tau \left[s_z^{(3)}(t) - \bar{z}(t) s^{(1)}(t)' \right] d\mu_0(t),$$

$$A_{21} = \int_0^\tau \left[s_w^{(2)}(t) - \bar{w}(t)\bar{z}(t) s^{(0)}(t)' \right] d\{\lambda_0(t)t\},$$

$$A_{22} = \int_0^\tau \left[s_w^{(3)}(t) - \bar{w}(t) s^{(1)}(t)' \right] d\mu_0(t).$$

Theorem 1 *Under conditions (C1)–(C4), $n^{-1/2}U(\beta_0)$ is asymptotically normal with mean zero and covariance matrix $\Sigma = E\{d_i d_i'\}$, where*

$$d_i = \int_0^\tau \left\{ Z_i^*(t;\beta_0) - \bar{z}^*(t) \right\} dM_i(t;\beta_0),$$

and $\bar{z}^*(t) = (\bar{z}(t)', \bar{w}(t)')'$.

Proof It can be checked that

$$U(\beta_0) = \sum_{i=1}^n \int_0^\tau \left\{ Z_i^*(t;\beta_0) - \bar{Z}^*(t;\beta_0) \right\} dM_i(t;\beta_0).$$

By following the argument used in Theorem 1 of Lin et al. (1998), it is easily to get that

$$n^{-1/2}U(\beta_0) = n^{-1/2} \sum_{i=1}^n d_i + o_p(1).$$

Utilizing the multivariate central limit theorem, $n^{-1/2}U(\beta_0)$ converges in distribution to a normal random variable with mean zero and variance matrix $\Sigma = E\{d_i d_i'\}$, which can be consistently estimated by $\hat{\Sigma}$ defined in (6).

Let $\mathcal{U}(\beta)$ be the limit of $n^{-1}U(\beta)$, and $\mathcal{N}$ be a compact neighborhood of β_0 on which $\|U(\beta)\|$ is minimised to obtain $\hat{\beta}$. □

Theorem 2 *Assume that conditions (C1)–(C6) hold and $\mathcal{U}(\beta) \neq 0$ for all $\beta \in \mathcal{N}$ but $\beta \neq \beta_0$. Then $\hat{\beta}$ is strongly consistent and $n^{1/2}(\hat{\beta} - \beta_0)$ converges in distribution to zero-mean normal with covariance matrix $A^{-1}\Sigma(A^{-1})'$.*

Proof Write

$$U_1(\beta) - U_1(\beta_0) = \{U(\beta_1, \beta_2) - U_1(\beta_1, \beta_{20})\} + \{U(\beta_1, \beta_{20}) - U_1(\beta_{10}, \beta_{20})\}.$$

For any sequence $\varepsilon_n \to 0$, using a Taylor series expansion and the uniform strong law of large numbers (Pollard 1990), we have that for $\|\beta - \beta_0\| \leq \varepsilon_n$,

$$U_1(\beta_1, \beta_2) - U_1(\beta_1, \beta_{20}) = -A_{12} n(\beta_2 - \beta_{20}) + o(n\|\beta_2 - \beta_{20}\|)$$

almost surely. Note that

$$\begin{aligned}
U_1(\beta_1, \beta_{20}) - U_1(\beta_{10}, \beta_{20}) = \Bigg[& \sum_{i=1}^{n} \int_0^{\tau} \left\{ Z_i(t; \beta_1, \beta_{20}) - \bar{Z}(t; \beta_1, \beta_{20}) \right\} \left\{ \tilde{N}_i(t; \beta_1) \right. \\
& \left. - Y_i(t; \beta_1) g(\beta_{20}' Z_i) d\mu_0(t e^{(\beta_{10}-\beta_1)' Z_i}) \right\} \\
& - \sum_{i=1}^{n} \int_0^{\tau} \left\{ Z_i(t; \beta_{10}, \beta_{20}) - \bar{Z}(t; \beta_{10}, \beta_{20}) \right\} \\
& \times \left\{ \tilde{N}_i(t; \beta_{10}) - Y_i(t; \beta_{10}) g(\beta_{20}' Z_i) d\mu_0(t) \right\} \Bigg] \\
& + \sum_{i=1}^{n} \int_0^{\tau} \left\{ Z_i(t; \beta_1, \beta_{20}) - \bar{Z}(t; \beta_1, \beta_{20}) \right\} \\
& \times Y_i(t; \beta_1) g(\beta_{20}' Z_i) d\{\mu_0(t e^{(\beta_{10}-\beta_1)' Z_i}) - \mu_0(t)\}.
\end{aligned} \tag{A.1}$$

Applying the technique of Ying (1993) and Lin et al. (1998), we can show that the first term on the right-hand side of (A.1) is of order $o(n^{1/2})$. It follows from a Taylor series expansion that

$$\mu_0(t e^{(\beta_{10}-\beta_1)' Z_i}) - \mu_0(t) = \{\lambda_0(t) + o(1)\} t (\beta_{10} - \beta_1)' Z_i.$$

Therefore, the second term on the right-hand side of (A.1) is

$$\begin{aligned}
& \sum_{i=1}^{n} \int_0^{\tau} \left\{ Z_i(t; \beta_1, \beta_{20}) - \bar{Z}(t; \beta_1, \beta_{20}) \right\} Y_i(t; \beta_1) g(\beta_{20}' Z_i) Z_i' d\{t\lambda_0(t)\} (\beta_{10} - \beta_1) \\
& \quad + o(n \|\beta_1 - \beta_{10}\|) \\
& = -A_{11} n (\beta_1 - \beta_{10}) + o(n \|\beta_1 - \beta_{10}\|)
\end{aligned}$$

almost surely. It then follows that for any sequence $\varepsilon_n \to 0$,

$$\sup_{\|\beta-\beta_0\| \le \varepsilon_n} \left\{ \|U_1(\beta) - U_1(\beta_0) + (A_{11}, A_{12}) n (\beta - \beta_0)\| / \left(n^{1/2} + n\|\beta - \beta_0\|\right) \right\} = o(1)$$

almost surely. Similarly, we get that for any sequence $\varepsilon_n \to 0$,

$$\sup_{\|\beta-\beta_0\| \le \varepsilon_n} \left\{ \|U_2(\beta) - U_2(\beta_0) + (A_{21}, A_{22}) n (\beta - \beta_0)\| / \left(n^{1/2} + n\|\beta - \beta_0\|\right) \right\} = o(1)$$

almost surely. Hence for any sequence $\varepsilon_n \to 0$,

$$\sup_{\|\beta-\beta_0\|\le\varepsilon_n} \left\{\|U(\beta) - U(\beta_0) + An(\beta-\beta_0)\| / \left(n^{1/2} + n\|\beta-\beta_0\|\right)\right\} = o(1) \quad \text{(A.2)}$$

almost surely. It is easy to show that $\mathcal{U}(\beta_0) = 0$. Note that $n^{-1}U(\beta) \to \mathcal{U}(\beta)$ uniformly in $\mathcal{N}$ and $\mathcal{U}(\beta) \neq 0$ for all $\beta \neq \beta_0$. Then following the argument used in Theorem 2 of Lin et al. (1998), we can get that $\hat{\beta}$ is strongly consistent under the regularity conditions (C1)–(C5). In addition, by the definition of $\hat{\beta}$ and condition (C6), it follows from (A.2) that $n^{1/2}(\hat{\beta} - \beta_0)$ is asymptotically normal with mean zero and covariance matrix $A^{-1}\Sigma(A^{-1})'$, which can be consistently estimated by $\hat{A}^{-1}\hat{\Sigma}(\hat{A}^{-1})'$. □

Theorem 3 *Under conditions (C1)–(C6), $V(t)$ converges weakly to a zero-mean Gaussian process with covariance function $\Gamma(s,t) = E\{\Psi_i(s)\Psi_i(t)\}$ at (s,t), where*

$$\Psi_i(t) = \int_0^t \frac{dM_i(u;\beta_0)}{s^{(0)}(u)} - h(t)'A^{-1}\int_0^\tau \left\{Z_i^*(u;\beta_0) - \bar{z}^*(u)\right\} dM_i(u;\beta_0),$$

$$h_1(t) = \int_0^t \bar{z}(u)d\{\lambda_0(u)u\}, \qquad h_2(t) = \int_0^t \frac{s^{(1)}(u)}{s^{(0)}(u)} d\mu_0(u),$$

and $h(t) = (h_1(t)', h_2(t)')'$.

Proof To derive the asymptotic normality of $V(t)$, first note that

$$\begin{aligned}\hat{\mu}_0(t) - \mu_0(t) = &\{\hat{\mu}_0(t;\hat{\beta}_1,\hat{\beta}_2) - \hat{\mu}_0(t;\hat{\beta}_1,\beta_{20})\} \\ &+ \{\hat{\mu}_0(t;\hat{\beta}_1,\beta_{20}) - \hat{\mu}_0(t;\beta_0)\} \\ &+ \{\hat{\mu}_0(t;\beta_0) - \mu_0(t)\}.\end{aligned}$$

Using a Taylor series expansion and the uniform strong law of large numbers (Pollard 1990), we obtain that uniformly in $t \in [0,\tau]$,

$$\{\hat{\mu}_0(t;\hat{\beta}_1,\hat{\beta}_2) - \hat{\mu}_0(t;\hat{\beta}_1,\beta_{20})\} = -h_2(t)'(\hat{\beta}_2 - \beta_{20}) + o_p(n^{-1/2}).$$

By following the proof of (A.2), it is seen that

$$\{\hat{\mu}_0(t;\hat{\beta}_1,\beta_{20}) - \hat{\mu}_0(t;\beta_{10},\beta_{20})\} = -h_1(t)'(\hat{\beta}_1 - \beta_{10}) + o_p(n^{-1/2})$$

uniformly in $t \in [0,\tau]$. It can be checked that

$$\hat{\mu}_0(t;\beta_0) - \mu_0(t) = n^{-1}\sum_{i=1}^n \int_0^t \frac{dM_i(u;\beta_0)}{s^{(0)}(u)} + o_p(n^{-1/2})$$

uniformly in $t \in [0, \tau]$. Thus, it follows from Theorem 1 that

$$V(t) = n^{-1/2} \sum_{i=1}^{n} \Psi_i(t) + o_p(1) \tag{A.3}$$

uniformly in $t \in [0, \tau]$. Because $\Psi_i(t)$ are independent zero-mean random variables for each t, the multivariate central limit theorem implies that $n^{-1/2} \sum_{i=1}^{n} \Psi_i(t)$ converges in finite dimensional distributions to a zero-mean Gaussian process. Using the modern empirical theory as in Lin et al. (2000), we can show that $n^{-1/2} \sum_{i=1}^{n} \Psi_i(t)$ is tight. Thus, $V(t)$ converges weakly to a zero-mean Gaussian process with covariance function $\Gamma(s, t)$ at (s, t). By the arguments of Lin et al. (2000), the covariance function $\Gamma(s, t)$ can be consistently estimated by $\hat{\Gamma}(s, t)$ defined in (9). □

Theorem 4 *Under conditions (C1)–(C6), the null distribution of $F(t, z; \hat{\beta})$ converges weakly to a zero-mean Gaussian process with covariance function $E\{\Phi_i(t, z)\Phi_i(t^\dagger, z^\dagger)\}$ at (t, z) and $(t^\dagger, z^\dagger)$, where*

$$\Phi_i(t, z) = \int_0^t \left\{ I(Z_i \le z) - \frac{s(u, z)}{s^{(0)}(u)} \right\} dM_i(u; \beta_0)$$
$$- b(t, z)' A^{-1} \int_0^\tau \left\{ Z_i^*(u; \beta_0) - \bar{z}^*(u) \right\} dM_i(u; \beta_0),$$
$$s(u, z) = E\{Y_i(u; \beta_{10}) g(\beta_{20}' Z_i) I(Z_i \le z)\},$$
$$b_1(t, z) = E\left\{ \int_0^t Y_i(u; \beta_{10}) g(\beta_0' Z_i) I(Z_i \le z) \{Z_i - \bar{z}(u)\} \, d\{\lambda_0(u) u\} \right\},$$
$$b_2(t, z) = E\left\{ \int_0^t Y_i(u; \beta_{10}) \dot{g}(\beta_{20}' Z_i) Z_i \left\{ I(Z_i \le z) - \frac{s(u, z)}{s^{(0)}(u)} \right\} d\mu_0(u) \right\},$$

and $b(t, z) = (b_1(t, z)', b_2(t, z)')'$.

Proof Write

$$F(t, z; \hat{\beta}) = \Bigg[n^{-1/2} \sum_{i=1}^{n} \int_0^t I(Z_i \le z) \Big\{ d\tilde{N}_i(u; \hat{\beta}_1)$$
$$- Y_i(u; \hat{\beta}_1) g(\beta_{20}' Z_i) d\mu_0(u e^{(\beta_{10} - \hat{\beta}_1)' Z_i}) \Big\} - n^{-1/2} \sum_{i=1}^{n}$$
$$\times \int_0^\tau I(Z_i \le z) \left\{ d\tilde{N}_i(u; \beta_{10}) - Y_i(u; \beta_{10}) g(\beta_{20}' Z_i) d\mu_0(u) \right\} \Bigg]$$

$$
\begin{aligned}
& - n^{-1/2} \sum_{i=1}^{n} \int_0^t I(Z_i \le z) Y_i(u; \hat{\beta}_1) g(\hat{\beta}_2' Z_i) d\left[\hat{\mu}_0(u, \hat{\beta}) - \mu_0(u)\right] \\
& - n^{-1/2} \sum_{i=1}^{n} \int_0^t I(Z_i \le z) Y_i(u; \hat{\beta}_1) \Big[g(\hat{\beta}_2' Z_i) d\mu_0(u) \\
& - g(\beta_{20}' Z_i) d\mu_0(u e^{(\beta_{10} - \hat{\beta}_1)' Z_i}) \Big] \\
& + n^{-1/2} \sum_{i=1}^{n} \int_0^t I(Z_i \le z) dM_i(u, \beta_0). \qquad \text{(A.4)}
\end{aligned}
$$

Applying the technique of Ying (1993) and Lin et al. (1998), we can show that the first term on the right-hand side of (A.4) is of order $o(1)$ uniformly in t and z. Similarly to (A.3), the second term on the right-hand side of (A.4) is equivalent to

$$
-n^{-1/2} \sum_{i=1}^{n} \int_0^t \frac{s(u, z)}{s^{(0)}(u)} dM_i(u; \beta_0) + \tilde{b}(t, z)' n^{1/2} (\hat{\beta} - \beta_0) + o_p(1),
$$

where

$$
\begin{aligned}
\tilde{b}_1(t, z) &= E\left\{ \int_0^t Y_i(u; \beta_{10}) g(\beta_0' Z_i) I(Z_i \le z) \bar{z}(u) d\{\lambda_0(u) u\} \right\}, \\
\tilde{b}_2(t, z) &= E\left\{ \int_0^t Y_i(u; \beta_{10}) g(\beta_{20}' Z_i) I(Z_i \le z) \frac{s^{(1)}(u)}{s^{(0)}(u)} d\mu_0(u) \right\},
\end{aligned}
$$

and $\tilde{b}(t, z) = (\tilde{b}_1(t, z)', \tilde{b}_2(t, z)')'$. It follows from a Taylor series expansion that the third term on the right-hand side of (A.4) equals

$$
-b^*(t, z)' n^{1/2} (\hat{\beta} - \beta_0) + o_p(1),
$$

where

$$
\begin{aligned}
b_1^*(t, z) &= E\left\{ \int_0^t Y_i(u; \beta_{10}) g(\beta_0' Z_i) I(Z_i \le z) Z_i d\{\lambda_0(u) u\} \right\}, \\
b_2^*(t, z) &= E\left\{ \int_0^t Y_i(u; \beta_{10}) \dot{g}(\beta_{20}' Z_i) I(Z_i \le z) Z_i d\mu_0(u) \right\},
\end{aligned}
$$

and $b^*(t, z) = (b_1^*(t, z)', b_2^*(t, z)')'$. Therefore, it follows that uniformly in t and z,

$$F(t, z; \hat{\beta}) = n^{-1/2} \sum_{i=1}^{n} \Phi_i(t, z) + o_p(1),$$

which is a sum of i.i.d. zero-mean terms for fixed t and z. By the multivariate central limit theorem, $F(t, z; \hat{\beta})$ converges in finite dimensional distributions to a zero-mean Gaussian process. Using the modern empirical theory as in Lin et al. (2000), we can show that $n^{-1/2} \sum_{i=1}^{n} \Phi_i(t, z)$ is tight. Thus, $F(t, z; \hat{\beta})$ converges weakly to a zero-mean Gaussian process with covariance function $E\{\Phi_i(t, z)\Phi_i(t^\dagger, z^\dagger)\}$ at (t, z) and $(t^\dagger, z^\dagger)$. By the arguments of Lin et al. (2000), this Gaussian process can be approximated by the zero-mean Gaussian process $\tilde{F}(t, z)$ given by (11). □

References

Andersen PK, Gill RD (1982) Cox's regression model for counting processes: a large sample study. Ann Statist 10:1100–1120

Chen YQ, Jewell NP (2001) On a general class of semiparametric hazards regression models. Biometrika 88:687–702

Efron B (1979) Bootstrap methods: another look at the jackknife. Ann Statist 7:1–26

Fleming TR, Harrington DP (1991) Counting processes and survival analysis. Wiley, New York

Ghosh D (2004) Accelerated rates regression models for recurrent failure data. Lifetime Data Anal 10:247–261

Lawless JF, Nadeau C (1995) Some simple robust methods for the analysis of recurrent events. Technometrics 37:158–168

Liang KY, Zeger SL (1986) Longitudinal data analysis using generalized linear models. Biometrika 73:13–22

Lin DY, Geyer CJ (1992) Computational methods for semiparametric linear regression with censored dada. J Comp Graph Statist 1:77–92

Lin DY, Wei LJ, Ying Z (1998) Accelerated failure time models for counting processes. Biometrika 85:609–618

Lin DY, Wei LJ, Yang I, Ying Z (2000) Semiparametric regression for the mean and rate function of recurrent events. J R Statist Soc B 69:711–730

Lin DY, Wei LJ, Ying Z (2001) Semiparametric transformation models for point processes. J Amer Statist Assoc 96:620–628

Parzen MI, Wei LJ, Ying Z (1994) A resampling method based on pivotal estimating functions. Biometrika 81:341–350

Pepe MS, Cai J (1993) Some graphical displays and marginal regression analyses for recurrent failure times and time-dependent covariates. J Am Statist Assoc 88:811–820

Prentice RL, Williams BJ, Peterson AV (1981) On the regression analysis of multivariate failure time data. Biometrika 68:373–379

Pollard D (1990) Empirical processes: theory and applications. Institute of Mathematical Statistics, Hayward, California

Ramlau-Hansen H (1983) Smoothing counting process intensities by means of kernel functions. Ann Statist 11:453–466

Wei LJ, Lin DY, Weissfeld L (1989) Regression analysis of multivariate incomplete failure time data by modeling marginal distributions. J Am Statist Assoc 84:1065–1073

Ying Z (1993) A large sample study of rank estimation for censored regression data. Ann Statist 21:76–99