



# Accelerated Rates Regression Models for Recurrent Failure Time Data

DEBASHIS GHOSH

*Department of Biostatistics, University of Michigan, 1420 Washington Heights, Ann Arbor,  
MI 48109-2029, USA*

*Received September 14, 2002; Revised October 27, 2003; Accepted November 4, 2003*

**Abstract.** In this article, we formulate a semiparametric model for counting processes in which the effect of covariates is to transform the time scale for a baseline rate function. We assume an arbitrary dependence structure for the counting process and propose a class of estimating equations for the regression parameters. Asymptotic results for these estimators are derived. In addition, goodness of fit methods for assessing the adequacy of the accelerated rates model are proposed. The finite-sample behavior of the proposed methods is examined in simulation studies, and data from a chronic granulomatous disease study are used to illustrate the methodology.

**Keywords:** counting process, multiple events, Poisson process, survival data

## 1. Introduction

In many medical and scientific studies, subjects can experience recurrent or repeated events during follow-up. Examples of recurrent events include transient ischemic attacks in subjects with epilepsy, repeated opportunistic infections in HIV-infected subjects, and warranty claims in actuarial studies. Other areas in which recurrent events occur include sociology and reliability engineering. In these studies, the investigators are often interested in assessing the effects of covariates on certain features of the recurrent events process.

The majority of the work in the analysis of recurrent failure time data has focused on the counting process model of Andersen and Gill (1982). In their proposal, a nonhomogeneous Poisson process structure is assumed for the recurrent events; dependence between recurrent events is modelled using time-varying covariates. The Andersen–Gill model in the original paper specifies the effects of time-varying covariates on the intensity function of the process.

In the case of external covariates (Kalbfleisch and Prentice, 1980, Chapter 5.2.1), the Andersen–Gill model has been generalized to a proportional rates regression model by other authors (Pepe and Cai, 1993; Lawless et al., 1997; Lin et al., 2000). While there is a conditioning on history in the intensity-based model, there exists no such conditioning for the rate model. In the proportional rates models, the effect of covariates is multiplicative on the rate function, which often is of interest in many scientific settings (Wang et al., 2001). The rate function of a counting process is the first derivative of the mean function, which has been the quantity most often

modelled in regression settings for recurrent events data (Lawless and Nadeau, 1995; Lin et al., 1998; Lin et al., 2001). However, there are instances in which the assumptions here may not be valid or that the rate function is directly of interest. This motivates the need to consider alternative regression models for the analysis of recurrent failure time data.

One class of models which have been developed in many contexts is time-transformation models. In this framework, all subjects have similar trajectories, and the effect of covariates is to alter the time scale of the trajectories. Although these models have been studied for univariate survival data (Cox and Oakes, 1984, Chapter 5; Chen and Wang, 2000) and mean functions of recurrent failure time data (Lin et al., 1998), no such work has been developed for rate modelling in the context of recurrent failure time data. In this paper, we propose a model for arbitrary counting processes in which the effect of covariates is to change the time scale for a baseline rate function. This function will be left unspecified so that the model will be semiparametric; this model will be referred to as the accelerated rates model. It is related to the accelerated hazards model for univariate survival data studied by Chen and Wang (2000); we later discuss the connection between the two. We will assume an arbitrary dependence structure for the recurrent events process and develop procedures based on estimating equations for the regression parameters in the model using estimating equations. Asymptotic properties of the resulting estimators are derived; the proofs of the results are deferred to the Appendix. In Section 3, we also develop goodness of fit statistics for checking the adequacy of the accelerated rates model. In Section 4, we examine the finite-sample properties of the proposed methodology through simulation studies and application to data from a chronic granulomatous disease study. We conclude with some discussion in Section 5.

## 2. Model and inference procedures

Let  $N^*(t)$  denote the number of recurrences that occur in the absence of censoring in  $[0, t]$  and let  $\mathbf{Z}$  denote a  $p$ -vector of external (Kalbfleisch and Prentice, 1980, Chapter 5.2.1) covariates. Define  $d\mu(t|\mathbf{Z}) \equiv E\{dN^*(t)|\mathbf{Z}\}$  and  $d\mu_0(t) = E\{dN^*(t)|\mathbf{Z} = \mathbf{0}\}$ . The model we propose for the analysis of recurrent events is given by

$$d\mu(t|\mathbf{Z}) = d\mu_0(e^{\beta_0^T \mathbf{Z}} t), \quad (1)$$

where  $\beta_0$  is a vector of unknown regression coefficients, and  $a^T$  is the transpose of the vector  $a$ . We refer to (1) as the accelerated rates model for counting processes. The effect of covariates in this model is to expand or contract the time scale of the baseline rate function, depending on the sign of  $\beta_0$ . Typically, we leave  $d\mu_0(\cdot)$  unspecified so that (1) is a semiparametric model. Because  $N^*(t)$  is increasing in  $t$ , we have that  $d\mu_0 \geq 0$ . If  $N^*(t)$  is a simple counting process (i.e., can only take a value of 0 or 1), then (1) reduces to a version of the accelerated hazards model studied by Chen and Wang (2000) for independent survival data. It is also important for  $\mu_0(t)$  to not be

constant, or else  $\beta_0$  will be nonidentifiable. This is one of the regularity conditions assumed in the theoretical derivations in the Appendix.

In order to highlight the difference between this model with the AFT model for counting processes of Lin et al. (1998), it is necessary to consider the model for the mean function of the point process induced by (1). Simple integration yields the following model:

$$\mu(t|Z) = \mu_0(e^{\beta_0 Z} t) e^{-\beta_0 Z},$$

where  $\mu(t|Z) = E\{N^*(t)|Z\}$ . Inspection shows that this model differs from the AFT model of Lin et al. (1998) by a factor of  $e^{-\beta_0 Z}$ . Thus, formulation of covariate effects through model (1) induces a model for the mean function of the point process that is more complicated than the AFT model.

Let  $C$  denote the censoring time. Because of censoring, it is impossible to observe  $N^*(\cdot)$ . Rather, we observe  $\{N_i(\cdot), C_i, \mathbf{Z}_i\}, i = 1, \dots, n$ , a random sample from  $\{N(\cdot), C, \mathbf{Z}\}$ , where  $N(t) = N^*(t \wedge C)$  and  $a \wedge b$  is the minimum of  $a$  and  $b$ . If  $T_{ik}$  is the time to the  $k$ th event for the  $i$ th subject ( $i = 1, \dots, n; k = 1, 2, \dots$ ), then it is easy to see that

$$N_i^*(t) = \sum_{k=1}^{\infty} I(T_{ik} \leq t),$$

and

$$N_i(t) = \sum_{k=1}^{\infty} I(T_{ik} \leq t \wedge C_i).$$

Let us define

$$N_i^*(t; \beta) = \sum_{k=1}^{\infty} I(T_{ik} e^{\beta^T \mathbf{Z}_i} \leq t),$$

$Y_i(t; \beta) = I(C_i e^{\beta^T \mathbf{Z}_i} \geq t)$ , and  $\tilde{N}_i(t; \beta) = N_i^*(t \wedge C_i e^{\beta^T \mathbf{Z}_i}; \beta)$ , ( $i = 1, \dots, n$ ). It then follows that  $E\{N_i^*(t; \beta_0)\} = E\{N_i^*(t e^{-\beta_0^T \mathbf{Z}_i})\}$ ; in view of this fact and model (1), we have that

$$E\{N_i^*(t; \beta_0)\} = \mu_0(t) e^{-\beta_0^T \mathbf{Z}_i}. \quad (2)$$

By (2), for  $i = 1, \dots, n$ ,

$$\begin{aligned} M_i(t; \beta) &= \int_0^t Y_i(s; \beta) d\{N_i^*(s; \beta) - \mu_0(s) e^{-\beta^T \mathbf{Z}_i}\} \\ &= \tilde{N}_i(t; \beta) - \int_0^t Y_i(s; \beta) e^{-\beta^T \mathbf{Z}_i} d\mu_0(s) \end{aligned} \quad (3)$$

are mean-zero processes when  $\beta = \beta_0$ . Since  $e^{\beta^T \mathbf{Z}_i}$  ( $i = 1, \dots, n$ ) is constant with respect to time, we can multiply both sides of (3) by it and still preserve the zero mean structure of the processes. This motivates the following class of estimating functions for  $\beta_0$

$$\mathbf{U}(\beta) = \sum_{i=1}^n \int_0^\infty W(t; \beta) \{ \mathbf{Z}_i - \bar{\mathbf{Z}}(t; \beta) \} e^{\beta^T \mathbf{Z}_i} d\tilde{N}_i(t; \beta), \quad (4)$$

where  $W(t; \beta)$  is a bounded weight function that converges in probability to a function  $w(t)$  and  $\bar{\mathbf{Z}}(t; \beta) = \sum_{j=1}^n Y_j(t; \beta) \mathbf{Z}_j / \sum_{j=1}^n Y_j(t; \beta)$ . Since  $\mathbf{U}(\beta)$  is a discrete function of  $\beta$ , we define  $\hat{\beta}$  to be the zero-crossing from setting (4) equal to zero. Various methods of solving this equation exist, such as bisection (Press et al., 1992) or simulated annealing (Lin and Geyer, 1992). Given  $\hat{\beta}$ , an Aalen–Breslow type estimator of  $\mu_0(\cdot)$  is given by  $\hat{\mu}_0(t; \hat{\beta})$ , where

$$\hat{\mu}_0(t; \beta) = \sum_{i=1}^n \int_0^t \frac{e^{\beta^T \mathbf{Z}_i} d\tilde{N}_i(s; \beta)}{\sum_{j=1}^n Y_j(s; \beta)}.$$

The baseline rate function  $d\mu_0(\cdot)$  in (1) can then be computed using nonparametric regression techniques in a manner similar to estimation of the hazard function (Ramlau-Hansen, 1983). We briefly describe this kernel estimation procedure in the appendix.

We note that in (4), there is a weight function that needs to be specified. Ideally, we would choose  $W(t; \beta)$  to minimize the variance of  $\mathbf{U}(\beta)$ . However, it does not appear possible to derive an optimal weight without specification of the dependence structure on the increments of  $N^*(t)$ . Therefore, we suggest two possible choices for  $W$ . The first option is to take  $W(t; \beta) \equiv 1$ , which we term the log-rank weight function. The second choice of a weight function is  $W(t; \beta) \equiv n^{-1} \sum_{i=1}^n Y_i(t; \beta)$ , which is called the Gehan weight function. The performance of these weight functions is explored later in the simulation studies and in the numerical application.

For  $i = 1, \dots, n$ , let  $\hat{M}_i(t; \beta) = \tilde{N}_i(t; \beta) - \int_0^t Y_i(s; \beta) e^{-\beta^T \mathbf{Z}_i} d\hat{\mu}_0(s; \beta)$ . Assuming certain regularity conditions, we prove in Appendix B.1 that  $n^{-1/2} \mathbf{U}(\hat{\beta}_0)$  converges in distribution to a normal distribution with mean zero vector and a covariance matrix that can be consistently estimated by  $\mathbf{V}(\hat{\beta})$ , where

$$\mathbf{V}(\beta) = n^{-1} \sum_{i=1}^n \mathbf{Q}_i(\beta) \mathbf{Q}_i^T(\beta),$$

and  $\mathbf{Q}_i(\beta) = \int_0^\infty W(t; \beta) \{ \mathbf{Z}_i - \bar{\mathbf{Z}}(t; \beta) \} e^{\beta^T \mathbf{Z}_i} d\hat{M}_i(t; \beta)$  ( $i = 1, \dots, n$ ). It is also shown there that  $n^{1/2}(\hat{\beta} - \beta_0)$  converges in distribution to a normal random vector with mean zero and variance matrix  $\mathbf{A}^{-1} \mathbf{V} \mathbf{A}^{-1}$ , where  $\mathbf{V} = \lim_{n \rightarrow \infty} \mathbf{V}(\hat{\beta})$  and

$$\mathbf{A} = \int_0^\infty w(t) E[Y_1(t; \beta_0) \{ \mathbf{Z}_1 - \bar{\mathbf{z}}(t) \}^{\otimes 2} e^{2\beta_0^T \mathbf{Z}_1}] d\dot{\mu}_0(t),$$

with  $\bar{\mathbf{z}}(t) = \lim_{n \rightarrow \infty} n^{-1} \bar{\mathbf{Z}}(t; \beta_0)$ ,  $\dot{\mu}_0(t) = d\mu_0(t)/dt$  and  $a^{\otimes 2} = aa^T$ . In Appendix B.2, we prove that  $n^{1/2} \{ \hat{\mu}_0(t; \hat{\beta}) - \mu_0(t) \}$  converges weakly to a mean-zero Gaussian process with covariance function

$$\begin{aligned} \zeta(t,s) = E & \left( \left[ \int_0^t \frac{e^{\beta_0^T \mathbf{Z}_1} dM_1(u; \beta_0)}{E\{Y_1(u; \beta_0)\}} + c^T(t) \mathbf{A}^{-1} \int_0^\infty w(u) \{\mathbf{Z}_1 - \bar{\mathbf{z}}(u)\} e^{\beta_0^T \mathbf{Z}_1} dM_1(u; \beta_0) \right] \right. \\ & \left. \times \left[ \int_0^s \frac{e^{\beta_0^T \mathbf{Z}_1} dM_1(u; \beta_0)}{E\{Y_1(u; \beta_0)\}} + c^T(s) \mathbf{A}^{-1} \int_0^\infty w(u) \{\mathbf{Z}_1 - \bar{\mathbf{z}}(u)\} e^{\beta_0^T \mathbf{Z}_1} dM_1(u; \beta_0) \right] \right), \end{aligned}$$

where  $c(t) = - \int_0^t \bar{\mathbf{z}}^*(u) d\dot{\mu}_0(u)$  and

$$\bar{\mathbf{z}}^*(u) = \lim_{n \rightarrow \infty} n^{-1} \sum_{i=1}^n \frac{e^{\beta_0^T \mathbf{Z}_i} \mathbf{Z}_i Y_i(u; \beta_0)}{\sum_{j=1}^n Y_j(u; \beta_0)}.$$

Based on this result, we see that estimating  $A$  requires an estimate for the derivative of  $d\mu_0(\cdot)$ . Because such density-type estimation tends to be numerically unstable, the resulting variance estimator will also be unreliable. In order to perform variance estimation for  $\hat{\beta}$ , we use a device from Parzen et al. (1994). We fix  $\{N_i(\cdot), C_i, \mathbf{Z}_i\}$  ( $i = 1, \dots, n$ ) and solve the following set of equations for  $\beta^*$

$$\mathbf{U}(\beta^*) = - \sum_{i=1}^n \mathbf{Q}_i(\hat{\beta}) G_i, \tag{5}$$

where  $\{G_1, \dots, G_n\}$  are  $n$  iid  $N(0, 1)$  random variables. Note that in (5), the only stochastic components are the  $\{G_i\}$ . We generate  $\{G_1, \dots, G_n\}$  many times and solve for  $\hat{\beta}^*$  each time. In Appendix B.3, we prove that the distribution of  $n^{1/2}(\hat{\beta}^* - \hat{\beta})$ , conditional on  $\{N_i(\cdot), C_i, \mathbf{Z}_i, i = 1, \dots, n\}$  is asymptotically equivalent to the unconditional distribution of  $n^{1/2}(\hat{\beta} - \beta_0)$ . Hence, confidence intervals for  $\beta_0$  can be constructed using the empirical distribution of  $\hat{\beta}^*$ .

### 3. Goodness of Fit Methods

As with any regression model, it is important to develop model checking methods for assessing the adequacy of (1). We propose the use of the following cumulative sums of residuals

$$\mathbf{U}_o(t; \beta) = \sum_{i=1}^n \mathbf{Z}_i e^{\beta^T \mathbf{Z}_i} \hat{M}_i(t; \beta). \tag{6}$$

If the accelerated rates model (1) holds, then these processes will fluctuate randomly about zero. Similar types of processes were considered by Wei (1984) and Lin et al. (1993) in other contexts. In Appendix B.4, we prove that the null distribution of  $n^{-1/2} \mathbf{U}_o(t; \hat{\beta})$  converges weakly to a Gaussian process and is asymptotically equivalent to that of  $n^{-1/2} \mathbf{U}_o^*(t)$ , where

$$\mathbf{U}_o^*(t) = \sum_{i=1}^n \int_0^t \left\{ \mathbf{Z}_i - \bar{\mathbf{z}}(u; \hat{\beta}) \right\} e^{\hat{\beta}^T \mathbf{Z}_i} d\hat{M}_i(u; \hat{\beta}) G_i + \{ \mathbf{U}_o(t; \hat{\beta}^*) - \mathbf{U}_o(t; \hat{\beta}) \},$$

and  $\hat{\beta}^*$  is defined as in (5).

Using this result, we can approximate the null distribution of  $n^{-1/2}\mathbf{U}_o(t; \widehat{\beta})$  by simulation. We plot both  $n^{-1/2}\mathbf{U}_o(t; \widehat{\beta})$  and a set of  $K$  realizations from  $n^{-1/2}\mathbf{U}_o^*(t)$ , the simulated realizations from the null distribution. In the examples considered here, we take  $K = 20$ . If the path of the process for the observed data is more extreme than those of the simulated processes, it suggests that the model fit (1) is not adequate. More formal tests can be constructed based on the Kolmogorov–Smirnov type statistic  $\sup_t |n^{-1/2}\mathbf{U}_o(t; \widehat{\beta})|$ . As in Lin et al. (1993), we can compare  $\sup_t |n^{-1/2}\mathbf{U}_o(t; \widehat{\beta})|$  to  $\sup_t |n^{-1/2}\mathbf{U}_o^*(t)|$  for  $K$  simulated realizations. We can then compute a p-value using the simulated processes. Our experience has shown that using  $K = 1000$  seems to work well.

## 4. Numerical Results

### 4.1. Simulation Studies

To assess the finite-sample properties of the proposed methods, extensive numerical studies were conducted. We considered sample sizes  $n = 50$  and  $100$ . Our primary interest was in assessing the robustness of the estimators with respect to the correlation between the recurrent event times. Gap times between recurrences were generated from the following model

$$\lambda(t|Z, \psi) = \psi \lambda_0(te^{\beta_0 Z}),$$

where  $\psi$  is a gamma random variable with mean 1 and variance  $\sigma^2$ ,  $Z$  is a 0/1 treatment indicator (0 for placebo, 1 for treatment), and  $\lambda_0(x) = x^2$ . Because  $\lambda(t|Z, \psi) = E\{dN^*(t)|Z, \psi\}$ , by the law of iterated expectation,  $E\{dN^*(t)|Z\} = E[E\{dN^*(t)|Z, \psi\}] = \lambda_0(te^{\beta_0 Z})$ , which is of the form (1). When  $\sigma^2 = 0$ , recurrent events within an individual are independent; nonzero values of  $\sigma^2$ , on the other hand, induce correlation between recurrences.

For each simulation study, we considered  $\sigma^2 = 0, 1.0, 2.0$  and  $4.0$ . We compared the estimators of  $\beta_0$  using the log-rank and Gehan weight functions. For each simulation setting, 1000 simulation samples were considered, and 1000 resamplings were generated for each simulation sample.

For the results reported here, we took  $\beta_0 = -0.5$ . Thus, the effect of treatment is a rightward shift in the rate function for subjects on treatment by a factor of  $\exp(0.5) = 1.6$  relative to the rate function for the subjects on placebo. Independent censoring was generated using both a uniform  $(0, 5)$  and a uniform  $(0, 15)$  random variable. Across the settings described, an average of 2.3 recurrences per individual were observed for the first scenario and 4.4 recurrences per individual for the second. For coverage probabilities, we will consider those based on the 95% Wald confidence interval, as well as those based on the empirical distribution of  $\beta^*$ . These results are summarized in Table 1.

Based on Table 1, we find that the estimators tend to have reasonable small-sample behavior. There is virtually no bias in the estimator of  $\beta_0$  for the sample sizes and

Table 1. Summary of simulation results for  $\beta_0$  in model (1) with  $\beta_0 = -0.5$ .

n	$\sigma^2$	$\tau$	Log-rank weight					Gehan weight				
			Bias	SE	SEE	CP	CP*	Bias	SE	SEE	CP	CP*
50	0	5	-0.01	0.36	0.3	0.91	0.96	0.01	0.35	0.31	0.92	0.97
		15	0.00	0.25	0.21	0.92	0.95	-0.01	0.24	0.21	0.91	0.96
	1	5	-0.02	0.46	0.42	0.92	0.97	-0.01	0.47	0.43	0.93	0.97
		15	-0.02	0.35	0.34	0.93	0.95	-0.01	0.34	0.33	0.93	0.95
	4	5	-0.01	0.5	0.45	0.92	0.96	0.00	0.49	0.45	0.92	0.97
		15	0.01	0.38	0.35	0.93	0.96	0.01	0.38	0.34	0.93	0.96
100	0	5	-0.01	0.25	0.24	0.94	0.95	0.01	0.24	0.22	0.94	0.94
		15	-0.01	0.18	0.17	0.95	0.94	0.00	0.17	0.16	0.95	0.94
	1	5	0.00	0.33	0.31	0.94	0.95	-0.01	0.33	0.32	0.95	0.95
		15	0.01	0.25	0.24	0.95	0.94	-0.01	0.25	0.24	0.94	0.95
	4	5	0.00	0.35	0.33	0.94	0.95	0.00	0.35	0.34	0.95	0.95
		15	0.01	0.27	0.26	0.95	0.95	0.01	0.28	0.26	0.94	0.94

Note: Bias is the mean of the estimators of  $\beta_0$  minus  $\beta_0$ ; SE: the standard error of the estimators of  $\beta_0$ ; SEE: the mean of the standard error estimate; CP: the coverage probability of the Wald 95% confidence interval, and CP\*: the coverage probability of the 95% confidence interval based on the empirical distribution of  $\beta^*$ . Independent censoring was generated using a uniform  $(0, \tau)$  random variable.

correlations considered. The Wald confidence intervals tend to be slightly anticonservative in smaller samples, but this behavior diminishes for larger sample sizes. The confidence intervals based on the empirical distribution of  $\beta^*$  tend to behave better for the sample sizes considered.

In the simulations, the choice of weight functions did not seem to have much effect on the variance of  $\beta_0$ . This trend held true for the other simulation scenarios we considered. More work needs to be done in investigating the efficiency of the proposed procedures.

#### 4.2. Chronic Granulomatous Data

We now give an example of the proposed methods to data from a chronic granulomatous disease (CGD) study previously analyzed by Therneau and Hamilton (1997). The study was a randomized clinical trial in which patients with CGD were randomized to either placebo or interferon-gamma (rIFN-g). The goal of the study was to assess the efficacy of rIFN-g in reducing the frequency of serious infections in CGD patients. There were 128 patients randomized in the study; 65 patients received rIFN-g, and 63 received placebo. There were 20 infections in the rIFN-g arm, while there were 55 infections among patients in the placebo group. In Table 2, we summarize the recurrent events data by treatment group.

We now analyze the recurrent events data using treatment as the only covariate in the model (1). The results of the two-sample comparison using the log-rank and Gehan weight functions are given in Table 3. We find that the effect of treatment on the rate of recurrent infection is significant. Based on the log-rank weight function,

Table 2. Summary of recurrent infection experiences in CGD study, stratified by treatment arm.

	No. of recurrent infections				Total
	0	1	2	$\geq 3$	
Treatment	0	1	2	$\geq 3$	
Placebo	35	19	4	7	55
rIFN-g	49	9	4	1	20

Table 3. Summary of regression analyses of recurrent infection data for subjects with chronic granulomatous disease.

Weight function	Covariate	Coefficient	SE	95% Confidence Intervals	
				Resampling	Wald
Gehan	Treatment only	-0.79	0.36	(-1.99, -0.21)	(-1.49, -0.08)
	Treatment	-0.80	0.28	(-1.82, 0.30)	(-1.34, -0.25)
	Age	-0.024	0.001	(-0.026, -0.021)	(-0.026, -0.021)
Log-rank	Treatment only	-0.64	0.33	(-1.91, -0.17)	(-1.28, 0.01)
	Treatment	-0.64	0.26	(-1.63, -0.26)	(-1.14, -0.13)
	Age	-0.018	0.001	(-0.022, -0.017)	(-0.019, -0.016)

Note: treatment is coded for placebo and 1 for rIFN-g. SE: standard error computed from empirical distribution of  $\beta^*$ ; resampling: 95% confidence interval is the 95% confidence interval based on the empirical distribution of  $\beta^*$  (10,000 resamplings).

this effect is estimated to be  $\exp(-0.64) = 0.53$ , while the Gehan weight function yields an estimated treatment effect of  $\exp(-0.79) = 0.45$ . The interpretation of this estimate is that use of rIFN-g is associated with a contraction in the time scale for rate of infection by a factor of  $\exp(0.64) = 1.9$  or  $\exp(0.79) = 2.2$ , depending on the weight function used. By comparison, using the transformation model of Lin et al. (2001) yields a treatment effect of  $-0.271$ , and the proportional means model gives a treatment effect of  $-1.10$  (Therneau and Hamilton, 1997). All of the estimated treatment effects are significant. Both of these results suggest that treatment is leading a decrease in mean occurrences of infections.

The analysis is now repeated using both treatment and age in the model. These results are also given in Table 3. While age is a statistically significant predictor of the rate of recurrences, it does not appear to be much of a confounder of the treatment effect. We also tested for an interaction between treatment and age, but it was found to be nonsignificant ( $P > 0.2$ ).

The model checking techniques are now applied in order to assess the adequacy of (1). We only show the analyses for the model with both covariates in the model. Similar results hold for the two-sample comparison. The graphs of the cumulative sum processes are given in Figures 1 and 2. Based on visual inspection, it does not appear that there are any violations of the model with respect to treatment and age. The  $p$ -values using Kolmogorov-Smirnov statistics for treatment and age, based on 10000 resamplings, are 0.475 and 0.543, respectively.



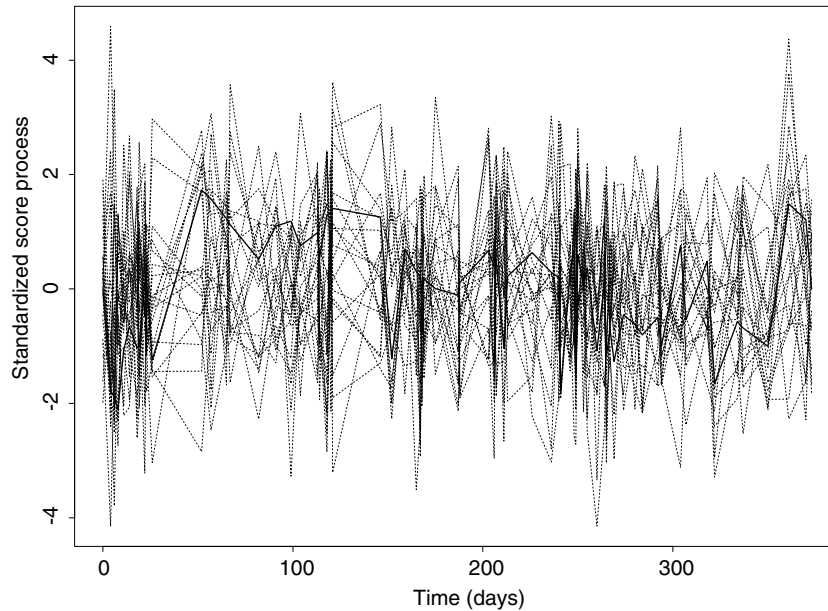


Figure 1. Standardized score process for assessing goodness of fit of the treatment covariate in regression model (1) with treatment and age as covariates. The solid line represents the observed process, while the dotted line represents 20 simulated realizations from the null distribution.

## 5. Remarks

We have assumed throughout the paper that the covariates  $Z$  are time-invariant. In some applications, it might be of interest to use time-varying covariates in the model. An obvious extension of (1) in this case would be

$$d\mu\{t|\mathbf{Z}(t)\} = d\mu_0\{e^{n_0^T \mathbf{Z}(t)}\}, \quad (7)$$

where  $d\mu\{t|\mathbf{Z}(t)\} = E\{dN^*(t)|\mathbf{Z}(s): 0 \leq s \leq t\}$ . However, the estimation procedure described in the paper cannot be extended in a straightforward manner to incorporate time-varying covariates. Thus, it would be of interest to develop estimation procedures for model (7).

It should be noted that while much of the estimation is related to that suggested by Lin et al. (1998), our model is fundamentally different from theirs, as explained in Section 2. We are positing a regression model for the derivative of the quantity suggested by Lin et al. (1998).

Another assumption that has been made here is that the only censoring present is independent, i.e., due to random loss to followup or end of the study. However, it is quite plausible that in medical studies involving serious recurrent events, there may be dependent censoring due to patient withdrawal or death. It would be of interest to extend procedures for the accelerated rates model to handle these problems.

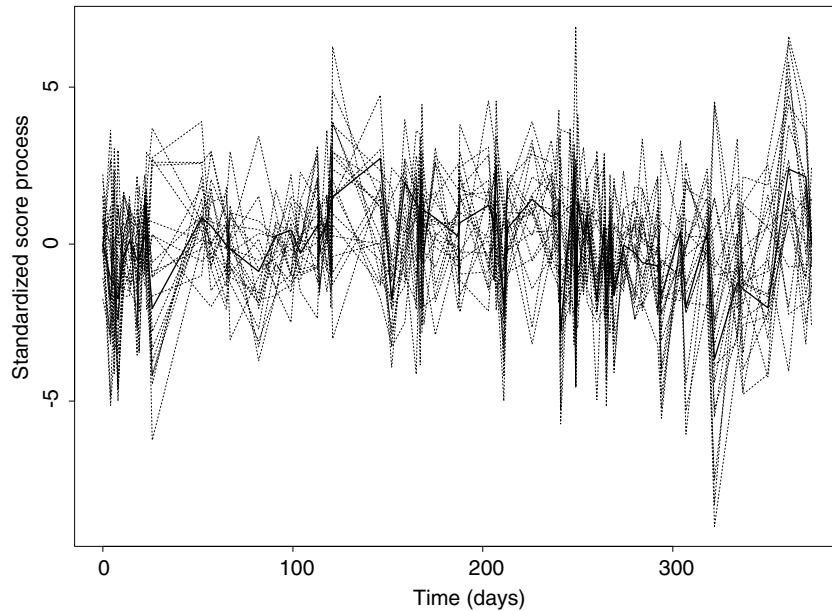


Figure 2. Standardized score process for assessing goodness of fit of the age covariate in regression model (1) with treatment and age as covariates. The solid line represents the observed process, while the dotted line represents 20 simulated realizations from the null distribution.

The resampling method for constructing confidence intervals and goodness of fit method can be computationally intensive for simulation purposes, but for a given analysis, the computation requirements are relatively moderate. Recently, Jin et al. (2003) have developed some fast computing algorithms for the AFT model with univariate survival data. It should be possible to adapt their algorithms to our model as well.

The issue of efficiency has only been touched on briefly. While the form of the estimating equations is similar to those for the univariate accelerated failure time model (Prentice, 1978; Wei et al., 1990), there has been no formal study done here examining potential efficiency loss. In addition, it is of interest to determine situations in which use of the Gehan weight function would lead to greater efficiency relative to the log-rank weight function, and vice versa. It would also be useful to derive other types of weight functions.

Finally, the times of the recurrences were assumed to be known exactly if they are observed. However, there are situations in which the recurrent events are not exactly observed. What is known is the number of recurrences since the subject's last visit; this type of data is known as panel count data (Lawless and Zhan, 1998; Sun and Wei, 2000). It would be desirable to extend the accelerated rates model to handle panel count data.

### Acknowledgments

The author would like to thank the Associate Editor and referees, whose comments improved the presentation of the article.

### Appendix A: Kernel Estimation of the Rate Function $d\mu_0(\cdot)$

Let us first consider the one-sample estimation problem of  $\mu(t) = E\{N^*(t)\}$ . A well-known estimator of  $\mu(t)$  is given by the Nelson–Aalen estimator:

$$\hat{\mu}(t) = \sum_{i=1}^n \int_0^t \frac{dN_i(u)}{\sum_{j=1}^n Y_j(u)},$$

where  $N_i(t) = N_i(t; 0)$  and  $Y_i(t) = Y_i(t; 0)$  ( $i = 1, \dots, n$ ). Then using arguments similar to those in Ramlau-Hansen (1983), an estimator for the one-sample rate function  $m(t) = d\mu(t)$  is given by

$$\hat{m}(t) = h^{-1} \int_0^\infty K\left(\frac{t-u}{h}\right) d\hat{\mu}(u), \quad (9)$$

where  $h$  is the bandwidth and  $K(z)$  is a kernel function. The estimator of  $d\mu_0(t)$  in model (1) follows by replacing  $\hat{\mu}(t)$  in (9) by  $\hat{\mu}_0(t; \hat{\beta})$ .

### Appendix B: Proofs of Asymptotic Results

#### B.1. Asymptotic Normality of $\hat{\beta}$ and $U(\beta_0)$ and Consistency of $V(\hat{\beta})$

We first impose the following regularity conditions

- (i) For  $i = 1, \dots, n$ ,  $N_i$  and  $Z_i$  are bounded;
- (ii)  $(N_i^*, C_i, \mathbf{Z}_i)$  ( $i = 1, \dots, n$ ) are independent and identically distributed random vectors;
- (iii)  $W$  has bounded variation and converges almost surely to a continuous function  $w$ ;
- (iv) For  $i = 1, \dots, n$ ,  $C_i(\beta_0)$  has a bounded density and  $\mu_0$  has a bounded second derivative;
- (v)  $\mu_0(\cdot)$  is not constant on  $[0, \infty)$ .

Conditions (i)–(iv) are the same as those made in the appendix of Lin et al. (1998). Condition (v) is needed here in order to ensure that  $\beta_0$  will be well-defined. These regularity conditions will be assumed throughout the rest of the Appendix.

Before proving the asymptotic normality of the estimated regression coefficient, we first prove the weak convergence of  $n^{-1/2}U(\beta_0)$ . Define

$$n^{-1/2}U(\beta_0, t) := n^{-1/2} \sum_{i=1}^n \int_0^t W(s; \beta_0) \{Z_i - \bar{Z}(s; \beta_0)\} e^{\beta_0^\top Z_i} d\tilde{N}_i(s; \beta_0).$$

Using arguments similar to those in the proof of Theorem 1 of Lin et al. (1998), we can show that  $n^{-1/2}U(\beta_0, \cdot)$  converges weakly to a Gaussian process with mean zero and covariance function

$$\mathbf{V}(t_1, t_2) = E \left[ \int_0^{t_1} w(s) \{Z_1 - \bar{z}(s)\} e^{\beta_0^\top Z_1} dM_1(s; \beta_0) \int_0^{t_2} w(s) \{Z_1 - \bar{z}(s)\}^\top \times e^{\beta_0^\top Z_1} dM_1(s; \beta_0) \right].$$

Note that  $\mathbf{V} = \mathbf{V}(\infty, \infty)$ . As an intermediate result, we have that  $n^{-1/2}U(\beta_0) := n^{-1/2}U(\beta_0, \infty)$  converges in distribution to a normal random variable with mean zero vector and covariance matrix  $\mathbf{V}$ .

It then follows from the arguments in Ying (1993) that under conditions (i)–(v),

$$\sup_{\|\beta - \beta_0\| \leq d_n} \{ \|U(\beta) - U(\beta_0) + \mathbf{A}n(\beta - \beta_0)\| / (n^{1/2} + n\|\beta - \beta_0\|) \} = o_P(1), \quad (A1)$$

almost surely for any sequence  $d_n \rightarrow 0$ .

Let  $\mathbf{u}(\beta)$  denote the limit of  $n^{-1}U(\beta)$ , and let  $C_R \equiv C_R(\beta_0)$  be a compact neighborhood of  $\beta_0$ . It is easy to show that  $n^{-1}U(\beta)$  converges uniformly to  $\mathbf{u}(\beta)$  on  $C_R$ . If  $\mathbf{u}(\beta) \neq 0$  for all  $\beta \in C_R \setminus \{\beta_0\}$ , then by arguments in the proof of Theorem 4 of Ying (1993),  $\hat{\beta} \rightarrow \beta_0$  almost surely. In addition, if  $\mathbf{A}$  is nonsingular, then by definition of  $\hat{\beta}$  and two applications of (A1),

$$n^{1/2}(\hat{\beta} - \beta_0) = \mathbf{A}^{-1}n^{-1/2}U(\beta_0) + o_P(1). \quad (A2)$$

This implies the asymptotic normality of  $\hat{\beta}$  and proves the weak convergence of  $n^{1/2}(\hat{\beta} - \beta_0)$ .

Using arguments similar to those in the proof of Theorem 3 of Lin et al. (1998), we can show that if conditions (i)–(v) hold, then for any  $\beta_n^\dagger$  that converges to  $\beta_0$ ,  $\mathbf{V}(\beta_n^\dagger) \rightarrow \mathbf{V}$ . This implies the consistency of  $\mathbf{V}(\hat{\beta})$ .

### B.2. Weak Convergence of $\hat{\mu}_0(t)$

By arguments analogous to those in Appendix B.1., we can derive an expression resembling (A1) for  $\hat{\mu}_0$ , i.e.,

$$\sup_{\|\beta - \beta_0\| \leq d_n, t \in [0, \tau]} \|n^{1/2}\{\hat{\mu}_0(t; \beta) - \hat{\mu}_0(t; \beta_0)\} - c^\top(t)n^{1/2}(\|\beta - \beta_0\|)\| = o_P(1), \quad (A3)$$

for all  $d_n \rightarrow 0$  and for  $\tau$  such that  $\lim n^{-1} \sum_{i=1}^n Y_i(\tau; \beta_0) > 0$ . Combining (A1) and (A3) yields that

$$n^{1/2}\{\hat{\mu}_0(t; \hat{\beta}) - \mu_0(t)\} = n^{1/2}\{\hat{\mu}_0(t; \beta_0) - \mu_0(t)\} + c^\top(t)\mathbf{A}^{-1}n^{-1/2}U(\beta_0) + o_P(1), \quad (A4)$$

uniformly in  $[0, \tau]$ . Using techniques as those in Appendix B.1., we can write

$$n^{1/2}\{\widehat{\mu}_0(t; \beta_0) - \mu_0(t)\} = n^{-1/2} \sum_{i=1}^n \int_0^t \frac{e^{\beta_0^\top \mathbf{Z}_i} dM_i(u; \beta_0)}{E\{Y_1(u; \beta_0)\}} + o_P(1). \quad (\text{A5})$$

Plugging (A5) into (A4) yields a representation of  $n^{1/2}\{\widehat{\mu}_0(t; \widehat{\beta}) - \mu_0(t)\}$  as a normalized average of mean-zero iid terms. To be precise,

$$n^{1/2}\{\widehat{\mu}_0(t; \widehat{\beta}) - \mu_0(t)\} = n^{-1/2} \sum_{i=1}^n \Psi_i(t) + o(1),$$

where

$$\Psi_i(t) = \int_0^t \frac{e^{\beta_0^\top \mathbf{Z}_i} dM_i(u; \beta_0)}{E\{Y_1(u; \beta_0)\}} + c^\top(t) \mathbf{A}^{-1} \int_0^\infty w(u) \{\mathbf{Z}_i - \bar{\mathbf{z}}(u)\} e^{\beta_0^\top \mathbf{Z}_i} dM_i(u; \beta_0).$$

The finite-dimensional convergence follows easily from the multivariate central limit theorem. Since  $\Psi_i$  ( $i = 1, \dots, n$ ) are composed of monotone functions, the weak convergence of  $n^{1/2}\{\widehat{\mu}_0(t; \widehat{\beta}) - \mu_0(t)\}$  for  $t \in [0, \tau]$  can be shown using results from empirical process theory (Pollard, 1990; Van der Vaart and Wellner, 1996). Finally, we can drop the tail restriction to  $\tau$  using arguments similar to those in Ying (1993).

### B.3. Asymptotic Equivalence of $n^{1/2}(\widehat{\beta} - \beta_0)$ and of $n^{1/2}(\widehat{\beta}^* - \widehat{\beta})$

Note that  $n^{-1/2} \mathbf{U}(\beta^*) = -n^{-1/2} \sum_{i=1}^n \mathbf{Q}_i(\widehat{\beta}) G_i$ . We can use arguments similar to those in the previous paragraphs to show that

$$n^{1/2}(\widehat{\beta}^* - \widehat{\beta}) = -\mathbf{A}^{-1} n^{-1/2} \sum_{i=1}^n \mathbf{Q}_i(\widehat{\beta}) G_i + o_P(1).$$

It is easy to see that conditional distribution of the random vector  $n^{-1/2} \sum_{i=1}^n \mathbf{Q}_i(\widehat{\beta}) G_i$  is normal with mean zero and covariance matrix  $\mathbf{V}(\widehat{\beta})$ . Since  $\mathbf{V}(\widehat{\beta})$  converges almost surely to  $\mathbf{V}$ , the conditional distribution of  $n^{1/2}(\widehat{\beta}^* - \widehat{\beta})$  is asymptotically equivalent to the unconditional distribution of  $n^{1/2}(\widehat{\beta} - \beta_0)$ .

### B.4. Asymptotic Equivalence of $n^{-1/2} \mathbf{U}_o^*(t)$ and $n^{-1/2} \mathbf{U}_o(t; \widehat{\beta})$

For  $i = 1, \dots, n$ , define  $H_i(t; \beta) = e^{\beta^\top \mathbf{Z}_i} M_i(t; \beta)$  and  $\widehat{H}_i(t; \beta) = e^{\beta^\top \mathbf{Z}_i} \widehat{M}_i(t; \beta)$ . Taking a Taylor series expansion about  $\beta_0$  and using arguments similar to those in the previous section, we can write  $n^{-1/2} \mathbf{U}_o(t; \widehat{\beta})$  as

$$\begin{aligned} n^{-1/2} \mathbf{U}_o(t; \widehat{\beta}) &= n^{-1/2} \mathbf{U}_o(t; \beta_0) - \mathbf{A}(t) \mathbf{A}^{-1} n^{-1/2} \mathbf{U}(\beta_0) + o_P(1), \\ &= n^{-1/2} \sum_{i=1}^n \int_0^t \{\mathbf{Z}_i - \bar{\mathbf{z}}(u)\} dH_i(u; \beta_0) \\ &\quad - \mathbf{A}(t) \mathbf{A}^{-1} n^{-1/2} \sum_{i=1}^n \int_0^\infty \{\mathbf{Z}_i - \bar{\mathbf{z}}(u)\} dH_i(u; \beta_0) + o_P(1), \end{aligned}$$

where  $\mathbf{A}(t)$  is a deterministic matrix with  $\mathbf{A}(\infty) = \mathbf{A}$ . Similarly, we can write  $n^{-1/2}\mathbf{U}_0^*(t)$  as

$$\begin{aligned} n^{-1/2}\mathbf{U}_0^*(t) &= n^{-1/2} \sum_{i=1}^n \int_0^t \{\mathbf{Z}_i - \bar{\mathbf{z}}(u)\} G_i dH_i(u; \hat{\beta}) \\ &\quad - \mathbf{A}(t)\mathbf{A}^{-1}n^{-1/2} \sum_{i=1}^n \int_0^\infty \{\mathbf{Z}_i - \bar{\mathbf{z}}(u)\} G_i dH_i(u; \hat{\beta}) + o_P(1). \end{aligned}$$

Since we have expressed  $n^{-1/2}\mathbf{U}_0(t; \hat{\beta})$  and  $n^{-1/2}\mathbf{U}_0^*(t)$  as normalized averages of independent and identically distributed terms, they both converge in finite dimension to Gaussian processes with mean zero. The weak convergence of these processes follows from results in empirical process theory. Using arguments similar to those in Appendix 2 of Lin et al. (2000), we can show that the conditional distribution of  $n^{-1/2}\mathbf{U}_0^*(t)$  and the unconditional distribution of  $n^{-1/2}\mathbf{U}_0(t; \hat{\beta})$  are the same.

## References

- P. K. Andersen and R. D. Gill, "Cox's regression model for counting processes: a large sample study," *Annals of Statistics* vol. 10 pp. 1100–1120, 1982.
- Y. Q. Chen and M.-C. Wang, "Analysis of accelerated hazards model," *Journal on American Statistical Association* 95, 608–618, 2000.
- D. R. Cox, "Regression models and life-tables (with discussion)," *Journal of the Royal Statistical Society Series B* vol. 34 pp. 187–220, 1972.
- D. R. Cox and D. Oakes, *Analysis of Survival Data*, Chapman and Hall: London, 1984.
- T. R. Fleming and D. P. Harrington, *Counting Processes and Survival Analysis*, Wiley: New York, 1991.
- Z. Jin, D. Y. Lin, L. J. Wei, and Z. Ying, "Rank-based inference for the accelerated failure time model," *Biometrika* vol. 90, 341–353, 2003.
- J. D. Kalbfleisch and R. L. Prentice, *The Statistical Analysis of Failure Time Data*, Wiley: New York, 1980.
- J. F. Lawless and C. Nadeau, "Some simple and robust methods for the analysis of recurrent events," *Technometrics* vol. 37 pp. 158–168, 1995.
- J. F. Lawless, C. Nadeau and R. J. Cook, "Analysis of mean and rate functions for recurrent events," in *Proceedings of the First Seattle Survival Analysis Symposium*, (D. Y. Lin and T. R. Fleming, eds.) Springer: New York, pp. 137–149, 1997.
- J. F. Lawless and M. Zhan, "Analysis of interval-grouped recurrent-event data using piecewise constant rate functions," *Canadian Journal of Statistics* vol. 26 pp. 549–565, 1998.
- D. Y. Lin and C. J. Geyer, "Computational methods for semiparametric linear regression with censored data," *Journal of Computational and Graphical Statistics* vol. 1 pp. 77–90, 1992.
- D. Y. Lin, J. A. Robins, and L. J. Wei, "Comparing two failure time distributions in the presence of dependent censoring," *Biometrika* vol. 83 pp. 381–393, 1996.
- D. Y. Lin, L. J. Wei, I. Yang and Z. Ying, "Semiparametric regression for the mean and rate functions of recurrent events," *Journal of the Royal Statistical Society Series B* vol. 62 pp. 711–730, 2000.
- D. Y. Lin, L. J. Wei and Z. Ying, "Checking the Cox model with cumulative sums of martingale-based residuals," *Biometrika*, vol. 80 pp. 557–572, 1993.
- D. Y. Lin, L. J. Wei, and Z. Ying, "Accelerated failure time models for counting processes," *Biometrika* vol. 85 pp. 609–618, 1998.
- D. Y. Lin, L. J. Wei, and Z. Ying, "Semiparametric transformation models for counting processes," *Journal of the American Statistical Association* vol. 96 pp. 620–628, 2001.

- M. I. Parzen, L. J. Wei, and Z. Ying, "A resampling method based on pivotal estimating functions," *Biometrika* vol. 81 pp. 341–350, 1994.
- M. S. Pepe and J. Cai, "Some graphical displays and marginal regression analyses for recurrent failure times and time dependent covariates," *Journal of the American Statistical Association* vol. 88 pp. 811–820, 1993.
- D. Pollard, *Empirical Processes: Theory and Applications*. IMS: Hayward, 1990.
- R. L. Prentice, "Linear rank tests with right censored data," *Biometrika* vol. 65 pp. 167–179, 1978.
- W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, *Numerical Recipes in C, 2nd Ed.*, Cambridge University Press: Cambridge, 1992.
- H. Ramlau-Hansen, "Smoothing counting process intensities by means of kernel functions," *Annals of Statistics* vol. 11 pp. 453–466, 1983.
- J. Sun and L. J. Wei, "Regression analysis of panel count data with covariate-dependent observation and censoring times," *Journal of the Royal Statistical Society Series B* vol. 62 pp. 293–302, 2000.
- T. M. Therneau and S. A. Hamilton, "RhDNase as an example of recurrent event analysis," *Statistics in Medicine* vol. 16 pp. 2029–2047, 1997.
- A. Van der Vaart and J. A. Wellner, *Weak Convergence and Empirical Processes*, Springer: New York, 1996.
- M. C. Wang, J. Qin, and C. T. Chiang, "Analyzing recurrent event data with informative censoring," *Journal of the American Statistical Association* vol. 96 pp. 1057–1065, 2001.
- L. J. Wei, "Testing goodness of fit for proportional hazards model with censored observations," *Journal of the American Statistical Association* vol. 79 pp. 649–652, 1984.
- L. J. Wei, Z. Ying, and D. Y. Lin, "Linear regression analysis of censored survival data based on rank tests," *Biometrika* vol. 77 pp. 845–851, 1990.
- Z. Ying, "A large sample study of rank estimation for censored regression data," *Annals of Statistics* vol. 21 pp. 76–99, 1993.