Accelerated Failure Time Models for Counting Processes

D. Y. Lin; L. J. Wei; Zhiliang Ying

Biometrika, Vol. 85, No. 3 (Sep., 1998), 605-618.

Stable URL:
http://links.jstor.org/sici?sici=0006-3444%28199809%2985%3A3%3C605%3AAFTMFC%3E2.0.CO%3B2-K

Biometrika is currently published by Biometrika Trust.

Biometrika (1998), **85**, 3, *pp.* 605–618
Printed in Great Britain

Accelerated failure time models for counting processes

BY D. Y. LIN
Department of Biostatistics, Box 357232, University of Washington, Seattle, Washington 98195, U.S.A.
danyu@biostat.washington.edu

L. J. WEI
Department of Biostatistics, Harvard University, 677 Huntington Avenue, Boston, Massachusetts 02115, U.S.A.
wei@biostat.harvard.edu

AND ZHILIANG YING
Department of Statistics, Hill Center, Busch Campus, Rutgers University, Piscataway, New Jersey 08855, U.S.A.
zying@stat.rutgers.edu

SUMMARY

We present a natural extension of the conventional accelerated failure time model for survival data to formulate the effects of covariates on the mean function of the counting process for recurrent events. A class of consistent and asymptotically normal rank estimators is developed for estimating the regression parameters of the proposed model. In addition, a Nelson–Aalen-type estimator for the mean function of the counting process is constructed, which is consistent and, properly normalised, converges weakly to a zero-mean Gaussian process. We assess the finite-sample properties of the proposed estimators and the associated inference procedures through Monte Carlo simulation and provide an application to a well-known bladder cancer study.

Some key words: Accelerated life model; Censoring; Cox regression; Log-rank statistic; Multiple events; Poisson process; Proportional hazards; Rank regression; Recurrent events; Survival data.

1. INTRODUCTION

In long-term follow-up studies, individual subjects may experience recurrent or repeated events. Examples in medical research include the sequence of asthmatic attacks, epileptic seizures, bleeding incidents, infection episodes or tumour recurrences in individual patients, while in industry recurrent events may result from breakdown of a certain type of machinery, such as computers and automobiles. In such studies, the investigators are often interested in estimating the frequency of recurrences over time as well as assessing the effects of covariates on the recurrence times.

One of the most significant advances in failure time data analysis since the seminal paper of Cox (1972) was the introduction of the counting process model with the Cox-type intensity function for recurrent events by Anderson & Gill (1982). These authors established an elegant large-sample estimation theory for their model by using the powerful

martingale theory. Recently, Pepe & Cai (1993), Lawless & Nadeau (1995) and Lawless, Nadeau & Cook (1997) studied multiplicative models for the rate and mean functions of arbitrary counting processes.

For classical survival data, an important alternative to the Cox proportional hazards model is the accelerated failure time model, which relates the logarithm of the failure time linearly to the covariates (Kalbfleisch & Prentice, 1980, pp. 32–4; Cox & Oakes, 1984, pp. 64–5). As pointed out by D. R. Cox (Reid, 1994, p. 450), 'accelerated life models are in many ways more appealing [than the proportional hazards model] because of their quite direct physical interpretation'. Semiparametric inference methods for the accelerated failure time model were proposed around 1980, e.g. Buckley & James (1979), Prentice (1978), and further studied in the early 1990's, e.g. Ritov (1990), Tsiatis (1990), Wei, Ying & Lin (1990), Lai & Ying (1991a, b), Ying (1993).

In this paper, we provide a natural extension of the aforementioned accelerated failure time model to accommodate recurrent events. This extension parallels the Andersen–Gill and Lawless–Nadeau–Cook extensions of the classical Cox proportional hazards model. Like Lawless et al., we formulate the mean functions of arbitrary counting processes rather than the intensity functions of Poisson processes. The proposed model retains the direct physical interpretation of the original accelerated failure time model in that the role of the covariates is to accelerate or decelerate the time to each recurrence. We estimate the vector of regression parameters under this model by generalising the rank-type estimating equations previously studied by Tsiatis (1990), Wei et al. (1990), Lai & Ying (1991a) and Ying (1993). In addition, we estimate the mean number of recurrences by extending the familiar Nelson–Aalen estimator. These generalisations entail some new technical challenges, which are tackled by modern empirical process theory. To avoid technical distractions, we describe the proposed estimators and the associated inference procedures in the next section while relegating the underlying theoretical development to the Appendix. In § 3, we present the results of some simulation studies along with an application to a bladder cancer study.

2. Inference procedures

For $i = 1, \ldots, n$ and $k = 1, 2, \ldots$, let T_{ik} be the kth event time for the ith subject. We assume that the subjects are independent, but do not impose any dependence structure on the recurrence times of the same subject. Define $N_i^*(t)$ as the number of events that have occurred on the ith subject by time t in the absence of censoring. That is,

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

where $I(.)$ is the indicator function. Suppose that the mean function of the counting process $N_i^*(t)$ associated with a p-vector of covariates Z_i takes the form

$$E\{N_i^*(t) | Z_i\} = \mu_0(e^{\beta_0' Z_i} t), \tag{1}$$

where β_0 is a p-vector of unknown regression parameters, and $\mu_0(.)$ is an unspecified continuous function. Write $\tilde{T}_{ik}(\beta) = T_{ik} e^{\beta' Z_i}$ and

$$\tilde{N}_i^*(t; \beta) = \sum_{k=1}^{\infty} I\{\tilde{T}_{ik}(\beta) \leqslant t\}.$$

Clearly, $\tilde{N}_i^*(t;\beta) = N_i^*(t\, e^{-\beta' Z_i})$, and model (1) is equivalent to

$$E\{\tilde{N}_i^*(t;\beta_0)\} = \mu_0(t). \tag{2}$$

According to model (1), the expected number of events by time t under $Z_i = z$ equals the expected number of events by time $e^{\beta_0' z} t$ under $Z_i = 0$. In other words, the set of covariates Z_i affects the frequency of recurrences over time by expanding or contracting the time scale on which the events occur by a multiplicative factor of $e^{\beta_0' Z_i}$ relative to that of a zero-valued covariate vector. If there is only a single event per subject with T_i denoting the failure time of the ith subject, then equation (2) implies that $\mathrm{pr}(T_i e^{\beta_0' Z_i} \leqslant t) = \mu_0(t)$ or

$$\log T_i = -\beta_0' Z_i + \varepsilon_i, \tag{3}$$

where the error terms ε_i $(i = 1, \ldots, n)$ have a common distribution. Equation (3) is the univariate accelerated failure time model (Kalbfleisch & Prentice, 1980, pp. 32–4).

In most applications, the follow-up time is subject to right censoring. Let C_i be the censoring time for the ith subject, which is assumed to be independent of T_{ik} $(k = 1, 2, \ldots)$ conditional on Z_i. The familiar counting process $N_i(t)$ as used by Andersen & Gill (1982) records the observed number of events over the follow-up period of the ith subject. That is,

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

where $a \wedge b = \min(a, b)$. For our purpose, it is more convenient to use the observed counting process on the transformed time scale:

$$\tilde{N}_i(t;\beta) := \sum_{k=1}^{\infty} I\{\tilde{T}_{ik}(\beta) \leqslant t \wedge \tilde{C}_i(\beta)\},$$

where $\tilde{C}_i(\beta) = C_i e^{\beta' Z_i}$. Let us also define $Y_i(t;\beta) = I\{\tilde{C}_i(\beta) \geqslant t\}$, and

$$M_i(t;\beta) = \tilde{N}_i(t;\beta) - \int_0^t Y_i(s;\beta)\, d\mu_0(s).$$

Note that $\tilde{N}_i(t;\beta) = N_i(t\, e^{-\beta' Z_i})$ and $Y_i(t;\beta) = I(C_i \geqslant t\, e^{-\beta' Z_i})$. Note also that

$$\tilde{N}_i(t;\beta) = \int_0^t Y_i(s;\beta)\, d\tilde{N}_i^*(s;\beta), \quad M_i(t;\beta) = \int_0^t Y_i(s;\beta)\, d\{\tilde{N}_i^*(s;\beta) - \mu_0(s)\}.$$

In view of equation (2), $M_i(t;\beta_0)$ $(i = 1, \ldots, n)$ have zero means. As a matter of fact, $M_i(t;\beta_0)$ $(i = 1, \ldots, n)$ are zero-mean martingale processes if $N_i^*(t)$ $(i = 1, \ldots, n)$ are non-homogeneous Poisson processes (Andersen & Gill, 1982). In this paper, we do not impose the Poisson structure because the independent increment assumption is rarely met in practice, especially in biomedical applications.

Motivated by the partial likelihood score function for the proportional intensity Poisson process model (Andersen & Gill, 1982) and the weighted rank estimating functions for model (3) (Prentice, 1978; Tsiatis, 1990; Wei et al., 1990), we propose the following class of estimating functions for β_0:

$$U(\beta) := \sum_{i=1}^{n} \int_0^{\infty} Q(t;\beta)\{Z_i - \bar{Z}(t;\beta)\}\, d\tilde{N}_i(t;\beta),$$

where $Q(t;\beta)$ is a specified weight function, and

$$\bar{Z}(t;\beta)=\frac{\sum_{i=1}^{n} Y_i(t;\beta)Z_i}{\sum_{i=1}^{n} Y_i(t;\beta)}.$$

We shall refer to $U(\beta)$ as the log-rank estimating function if $Q=1$ and as the Gehan estimating function if $Q(t;\beta)=n^{-1}Y_{\cdot}(t;\beta)$, where $Y_{\cdot}(t;\beta)=\sum Y_i(t;\beta)$.

As in the case of rank estimation for model (3), the estimating function $U(\beta)$ is a piecewise constant function of β. We define the estimator $\hat{\beta}$ as a zero-crossing of $U(\beta)$ or as a minimiser of $\|U(\beta)\|$, where $\|a\|=(a'a)^{\frac{1}{2}}$. When there are only a small number of covariates, direct grid search or the bisection method may be used to obtain $\hat{\beta}$. For high-dimensional covariate vectors, specialised numerical methods such as the technique of simulated annealing (Lin & Geyer, 1992) may be more efficient.

For the Gehan estimating function, it can be shown that

$$U(\beta)=n^{-1}\sum_{i=1}^{n}\sum_{j=1}^{n}\sum_{k=1}^{\infty} I(T_{ik}\leqslant C_i)(Z_i-Z_j)I\{\log C_j-\log T_{ik}\geqslant \beta'(Z_i-Z_j)\},$$

the rth component of which is monotone in β_r $(r=1,\ldots,p)$. Thus, one may obtain $\hat{\beta}$ by minimising the function

$$n^{-1}\sum_{i=1}^{n}\sum_{j=1}^{n}\sum_{k=1}^{\infty} I(T_{ik}\leqslant C_i)\{\log C_j-\log T_{ik}-\beta'(Z_i-Z_j)\}^{+},$$

where $a^{+}=\max(0,a)$. This minimisation problem can be easily solved by the linear programming technique (Barrodale & Roberts, 1973; Parzen, Wei & Ying, 1994). The resulting estimator may be slightly different from the minimiser of $\|U(\beta)\|$, though they are asymptotically equivalent.

Given $\hat{\beta}$, we estimate $\mu_0(t)$ by the Nelson–Aalen-type estimator $\hat{\mu}_0(t;\hat{\beta})$, where

$$\hat{\mu}_0(t;\beta)=\sum_{i=1}^{n}\int_0^t \frac{d\tilde{N}_i(s;\beta)}{Y_{\cdot}(s;\beta)}.$$

When no covariate is involved, $\hat{\mu}_0(t;0)$ reduces to the original Nelson–Aalen estimator (Fleming & Harrington, 1991, pp. 4–5).

By some simple algebraic manipulation,

$$U(\beta_0)=\sum_{i=1}^{n}\int_0^{\infty} Q(t;\beta_0)\{Z_i-\bar{Z}(t;\beta_0)\}\,dM_i(t;\beta_0), \tag{4}$$

and, for $t\leqslant \max_{1\leqslant i\leqslant n}\tilde{C}_i(\beta_0)$,

$$\hat{\mu}_0(t;\beta_0)-\mu_0(t)=\sum_{i=1}^{n}\int_0^t \frac{dM_i(s;\beta_0)}{Y_{\cdot}(s;\beta_0)}. \tag{5}$$

Since $M_i(t;\beta_0)$ $(i=1,\ldots,n)$ are zero-mean processes, both (4) and (5) are centred around 0 for large n. In fact, (4) and (5) would be martingale integrals if $N_i^{*}(t)$ were non-homogeneous Poisson processes. Without the martingale structure, it is more challenging to establish the asymptotic properties of (4) and (5).

We prove in the Appendix that $n^{-\frac{1}{2}}U(\beta_0)$ converges in distribution to a zero-mean normal random vector with a covariance matrix that can be consistently estimated by

$V(\beta_0)$ and $V(\hat{\beta})$, where

$$V(\beta) = n^{-1} \sum_{i=1}^{n} D_i(\beta) D_i(\beta)', \quad D_i(\beta) = \int_0^\infty Q(t; \beta)\{Z_i - \bar{Z}(t; \beta)\}\, d\hat{M}_i(t; \beta),$$

$$\hat{M}_i(t; \beta) = \tilde{N}_i(t; \beta) - \int_0^t Y_i(s; \beta)\, d\hat{\mu}_0(s; \beta).$$

These results parallel those of Pepe & Cai (1993) and Lawless et al. (1997) for multiplicative models. To test the null hypothesis $H_0: \beta = \beta_0$, one may use the quadratic form $n^{-1}U'(\beta_0)V^{-1}(\beta_0)U(\beta_0)$, which is asymptotically chi-squared on p degrees of freedom under H_0. This class of statistics generalises the weighted log-rank statistics for survival data (Prentice, 1978; Fleming & Harrington, 1991, Ch. 7) to recurrent events.

We also show in the Appendix that $n^{\frac{1}{2}}(\hat{\beta} - \beta_0)$ is asymptotically zero-mean normal. As evident from Theorem 2 given in the Appendix, an analytic estimation of the limiting covariance matrix for $\hat{\beta}$ would entail estimating the derivative of $\mu_0(.)$, which may not be done reliably in finite samples. Therefore, it is challenging to make inferences about subsets of β_0 when p is greater than 1. In the case of $p = 1$, it is straightforward to construct confidence intervals for β_0 by inverting the asymptotically standard-normal test statistic $n^{-\frac{1}{2}}U(\beta_0)/V^{\frac{1}{2}}(\beta_0)$ or $n^{-\frac{1}{2}}U(\beta_0)/V^{\frac{1}{2}}(\hat{\beta})$.

We propose to estimate the covariance matrix of $\hat{\beta}$ and to construct confidence intervals for individual components of β_0 by applying a resampling technique due to Parzen et al. (1994). Specifically, let $\hat{\beta}^*$ be the solution to

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

where $(G_1, \ldots, G_n)$ are independent standard normal variables. We may also obtain $\hat{\beta}^*$ by solving $U(\beta) = G$, where G is zero-mean normal with covariance matrix $nV(\hat{\beta})$. By the arguments of Parzen et al. (1994), $n^{\frac{1}{2}}(\hat{\beta} - \hat{\beta}^*)$ has the same limiting distribution as $n^{\frac{1}{2}}(\hat{\beta} - \beta_0)$. To approximate the distribution of $\hat{\beta}$, we obtain a large number of realisations of $\hat{\beta}^*$ by repeatedly generating the normal random samples $(G_1, \ldots, G_n)$ or the normal random vector G while fixing the data $\{N_i(.), C_i, Z_i\}$ $(i = 1, \ldots, n)$ at their observed values. The covariance matrix of $\hat{\beta}$ can then be estimated by the empirical covariance matrix of $\hat{\beta}^*$. In addition, confidence intervals for individual components of β_0 can be obtained from the percentiles of the empirical distribution of $\hat{\beta}^*$ or by the Wald method. The $(1 - 2\alpha)$ percentile interval is $[\hat{\beta}^*_{(\alpha)}, \hat{\beta}^*_{(1-\alpha)}]$, where $\hat{\beta}^*_{(\alpha)}$ and $\hat{\beta}^*_{(1-\alpha)}$ are the 100αth and $100(1-\alpha)$th percentiles of $\hat{\beta}^*$.

When model (1) holds, two estimators of β_0 with distinct weight functions, $\hat{\beta}_{Q_1}$ and $\hat{\beta}_{Q_2}$, say, should yield similar answers since they are both consistent. On the other hand, $\hat{\beta}_{Q_1}$ and $\hat{\beta}_{Q_2}$ tend to differ if model (1) is incorrect. Thus, we may check the adequacy of model (1) by comparing $\hat{\beta}_{Q_1}$ and $\hat{\beta}_{Q_2}$. This approach was previously taken by Lin (1991) and Wei et al. (1990) for the univariate proportional hazards and accelerated failure time models. Let $[\hat{\beta}^{*\prime}_{Q_1}, \hat{\beta}^{*\prime}_{Q_2}]'$ satisfy

$$\begin{bmatrix} U_{Q_1}(\hat{\beta}^*_{Q_1}] \\ U_{Q_2}(\hat{\beta}^*_{Q_2}) \end{bmatrix} = \begin{bmatrix} \sum_{i=1}^n D_{Q_1,i}(\hat{\beta}_{Q_1}) G_i \\ \sum_{i=1}^n D_{Q_2,i}(\hat{\beta}_{Q_2}) G_i \end{bmatrix},$$

where U_{Q_1}, U_{Q_2}, $D_{Q_1,i}$ and $D_{Q_2,i}$ are U and D_i associated with Q_1 and Q_2. Also, let Ω

be the empirical covariance matrix of $\hat{\beta}^*_{Q_1} - \hat{\beta}^*_{Q_2}$. Then the quadratic form

$$S := (\hat{\beta}_{Q_1} - \hat{\beta}_{Q_2})' \Omega^{-1} (\hat{\beta}_{Q_1} - \hat{\beta}_{Q_2})$$

is asymptotically chi-squared on p degrees of freedom under model (1).

Let $W(t) = n^{\frac{1}{2}} \{\hat{\mu}_0(t; \hat{\beta}) - \mu_0(t)\}$. We prove in the Appendix that $W(t)$ converges weakly to a zero-mean Gaussian process. As in the case of $\hat{\beta}$, it is difficult to estimate the limiting covariance function of $W(t)$ analytically. We again appeal to the resampling approach. We show in the Appendix that $W(t)$ has the same limiting distribution as

$$\hat{W}(t) := n^{\frac{1}{2}} \left\{ \hat{\mu}_0(t; \hat{\beta}) - \hat{\mu}_0(t; \hat{\beta}^*) + \sum_{i=1}^{n} H_i(t; \hat{\beta}) G_i \right\},$$

where

$$H_i(t; \beta) = \int_0^t \frac{d\hat{M}_i(s; \beta)}{Y_.(s; \beta)}.$$

Thus, we may use the simulated distribution of $\hat{W}(.)$ to make inferences about $\mu_0(.)$ along the lines of Lin, Fleming & Wei (1994).

As in Lin et al. (1994), we recommend that the log transformation be used in the construction of the confidence intervals/bands for $\mu_0(.)$, which not only restricts the resulting intervals/bands to be positive, but also improves the coverage probabilities in small samples. With the log transformation, the 95% confidence intervals for $\mu_0(t)$ based on the Wald and percentile methods are, respectively,

$$\hat{\mu}_0(t) \exp\{\pm 1{\cdot}96 n^{-\frac{1}{2}} \hat{\sigma}(t)/\hat{\mu}_0(t)\},$$

$$[\hat{\mu}_0(t) \exp\{-n^{-\frac{1}{2}} \xi_{0{\cdot}975}(t)/\hat{\mu}_0(t)\},\ \hat{\mu}_0(t) \exp\{-n^{-\frac{1}{2}} \xi_{0{\cdot}025}(t)/\hat{\mu}_0(t)\}],$$

where the argument $\hat{\beta}$ in $\hat{\mu}_0(t; \hat{\beta})$ is suppressed, $\hat{\sigma}(t)$ is the estimated standard error of $\hat{W}(t)$, and $\xi_{0{\cdot}025}(t)$ and $\xi_{0{\cdot}975}(t)$ are the estimated 2·5th and 97·5th percentile points of $\hat{W}(t)$. In addition, the 95% equal-precision confidence band for $\mu_0(t)$ $(\tau_1 \leqslant t \leqslant \tau_2)$ is

$$\hat{\mu}_0(t) \exp\{\pm \psi_{0{\cdot}95} n^{-\frac{1}{2}} \hat{\sigma}(t)/\hat{\mu}_0(t)\},$$

where $\psi_{0{\cdot}95}$ is the estimated 95th percentile of $\sup_{\tau_1 \leqslant t \leqslant \tau_2} |\hat{W}(t)/\hat{\sigma}(t)|$.

In many applications, the investigators are interested in estimating the frequency of recurrences associated with a given set of covariate values, say z. This is particularly useful in predicting the recurrence experience for individual patients. It is evident from equation (1) that the mean of $N^*(t)$ associated with z, denoted by $\mu(t|z)$, is equal to the baseline mean function $\mu_0(t)$ if the covariates are centred at z. Thus, one can use the above formulae for $\mu_0(t)$ to make inferences about $\mu(t|z)$ upon replacing the Z_i's by $(Z_i - z)$. Without transforming the data, one would estimate $\mu(t|z)$ by $\hat{\mu}(t|z) := \hat{\mu}_0(e^{\hat{\beta}'z} t; \hat{\beta})$, and approximate the distribution of $n^{\frac{1}{2}} \{\hat{\mu}(t|z) - \mu(t|z)\}$ by

$$\hat{W}(t; z) := n^{\frac{1}{2}} \left\{ \hat{\mu}_0(e^{\hat{\beta}'z} t; \hat{\beta}) - \hat{\mu}_0(e^{\hat{\beta}^{*\prime} z} t; \hat{\beta}^*) + \sum_{i=1}^{n} H_i(e^{\hat{\beta}'z} t; \hat{\beta}) G_i \right\}.$$

3. NUMERICAL RESULTS

3·1. *Simulation studies*

A series of simulation studies was conducted to assess the performance of the proposed inference procedures. We considered randomised trials with $n/2$ subjects in each of the

two groups, and consequently considered model (1) with Z being a single dichotomous covariate. Recurrent events were generated from both Poisson and non-Poisson processes. For the Poisson processes, the gap times between successive events are independent exponential variables; for the non-Poisson processes, the conditional distribution of the kth gap time X_k given X_{k-1} is the same as that of Gumbel's (1960) bivariate exponential distribution with correlation of 0·25. The hazard rate for each gap time was 1 or $e^{0.5}$ depending on whether $Z=0$ or 1. The follow-up time was censored by an independent Un[0, 3·5] random variable, which resulted in on average 2·3 and 2·6 observed events per subject in the Poisson and non-Poisson cases, respectively.

Table 1 summarises the main results of the simulation studies, showing the Monte Carlo estimates for the biases and standard errors of $\hat{\beta}$ and $\hat{\mu}_0(t)$ $(t=1, 3)$, the means of the standard error estimators for $\hat{\beta}$ and $\hat{\mu}_0(t)$, and the coverage probabilities of the 95% confidence intervals for β_0 and $\mu_0(t)$ based on the percentile and Wald methods. Each entry was based on 1000 simulated datasets. For each dataset, 1000 samples of $(G_1, \ldots, G_n)$ were generated to approximate the distributions of $\hat{\beta}$ and $\hat{\mu}_0(t)$. The estimates of β_0 were obtained by the bisection method with the accuracy of $\pm 0{\cdot}001$. The true value of β_0 is 0·5; the true values of $\{\mu_0(1), \mu_0(3)\}$ are (1, 3) and approximately (1·18, 3·35) in the Poisson and non-Poisson cases, respectively.

Table 1. *Summary of the simulation studies on the estimation of* β_0 *and* $\mu_0(t)$

	Poisson process				Non-Poisson process			
	Log-rank		Gehan		Log-rank		Gehan	
	$n=50$	100	$n=50$	100	$n=50$	100	$n=50$	100
$\lvert\text{Bias}(\hat{\beta})\rvert$	0·01	0·00	0·01	0·01	0·01	0·01	0·01	0·01
SE$(\hat{\beta})$	0·20	0·15	0·21	0·15	0·24	0·18	0·24	0·18
Mean of est. SE$(\hat{\beta})$	0·21	0·15	0·22	0·15	0·26	0·18	0·26	0·18
Coverage of 95% CI's								
percentile method	0·94	0·94	0·95	0·94	0·93	0·93	0·94	0·93
Wald method	0·96	0·94	0·96	0·95	0·95	0·94	0·96	0·94
$\lvert\text{Bias}(\hat{\mu}_0(1))\rvert$	0·01	0·00	0·01	0·00	0·00	0·00	0·00	0·00
SE$(\hat{\mu}_0(1))$	0·17	0·13	0·18	0·13	0·24	0·17	0·24	0·17
Mean of est. SE$(\hat{\mu}_0(1))$	0·18	0·13	0·18	0·13	0·24	0·17	0·24	0·17
Coverage of 95% CI's								
percentile method	0·94	0·93	0·94	0·94	0·93	0·94	0·93	0·93
Wald method	0·96	0·94	0·95	0·94	0·94	0·95	0·94	0·94
$\lvert\text{Bias}(\hat{\mu}_0(3))\rvert$	0·02	0·00	0·02	0·01	0·02	0·00	0·03	0·01
SE$(\hat{\mu}_0(3))$	0·46	0·35	0·47	0·35	0·59	0·42	0·59	0·42
Mean of est. SE$(\hat{\mu}_0(3))$	0·46	0·34	0·47	0·34	0·56	0·41	0·57	0·42
Coverage of 95% CI's								
percentile method	0·91	0·93	0·92	0·93	0·92	0·93	0·92	0·93
Wald method	0·93	0·94	0·94	0·94	0·93	0·94	0·93	0·94

As is evident from Table 1, the proposed inference procedures perform well in practical situations. Specifically, both $\hat{\beta}$ and $\hat{\mu}_0(t)$ are virtually unbiased. Their standard error estimators are very accurate. The confidence intervals for β_0 and $\mu_0(t)$ have reasonable coverage probabilities, although the percentile method tends to be slightly anti-conservative in small samples.

3·2. *Bladder tumour data*

We now illustrate the proposed methods with the well-known bladder cancer data reported by Byar (1980). These data were obtained from a randomised clinical trial con-

ducted by the Veterans Administration Co-operative Urological Group. One hundred and eighteen patients with superficial bladder tumours were admitted to this study between 1971 and 1976. The tumours were removed transurethrally and patients were randomly assigned to one of three treatments: placebo, pyridoxine or thiotepa. The average follow-up was about 31 months in all three treatment groups, but some patients were followed as long as five years. Bladder tumours tend to recur repeatedly. By the end of the follow-up, 62 patients had experienced at least one recurrence of tumours, among whom 39 had two or more recurrences and 28 had a third recurrence; the largest number of recurrences for a patient was 9.

As pointed out by Byar (1980), the goal of the analysis should be to determine the effect of treatment on the frequency of tumour recurrences. Our illustration focuses on the comparison between the placebo and thiotepa groups. There are 48 patients with a total of 87 observed recurrences and 38 patients with a total of 45 observed recurrences in the placebo and thiotepa groups, respectively. The investigators suspected that the frequency of recurrences might be associated with the number and sizes of the tumours present initially at the time of randomisation. Thus, we consider model (1) with $Z_i = (Z_{1i}, Z_{2i}, Z_{3i})'$, where Z_{1i} indicates by the value 1 versus 0 whether the ith patient was on placebo or thiotepa, Z_{2i} denotes the initial number of tumours for the ith patient, and Z_{3i} denotes the diameter, measured in centimetres, of the largest initial tumour for the ith patient. Both Z_{2i} and Z_{3i} range from 1 to 8.

The estimates of the regression parameters are presented in Table 2. The bisection method with the accuracy of $\pm 0{\cdot}00001$ was employed to find the estimates of β_0, and 10 000 simulations were used to approximate the distribution of $\hat{\beta}$. The results differ slightly between the log-rank and Gehan estimating functions and also between the percentile and Wald confidence intervals. The treatment assignment is marginally significant at the 5% level. The initial number of tumours is highly significant whereas the size of the largest initial tumour is non-significant. Incidentally, the chi-squared goodness-of-fit statistic S has an observed value of 0·995 on 3 degrees of freedom, providing no evidence against the assumed accelerated failure time model.

Table 2. *Regression analysis of the multiple tumour recurrence data for patients with bladder cancer*

Estimating function	Covariate	Parameter estimate	Estimated SE	95% confidence intervals Percentile	Wald
Log-rank	Treatment	0·542	0·312	(0·076, 1·269)	(−0·071, 1·154)
	Initial number	0·204	0·066	(0·102, 0·357)	(0·074, 0·334)
	Initial size	−0·038	0·084	(−0·237, 0·094)	(−0·203, 0·127)
Gehan	Treatment	0·657	0·314	(0·125, 1·354)	(0·042, 1·273)
	Initial number	0·218	0·086	(0·098, 0·445)	(0·048, 0·387)
	Initial size	−0·022	0·101	(−0·219, 0·183)	(−0·220, 0·176)

For comparison, we also analysed these data using the Andersen–Gill model. The corresponding parameter estimates are 0·524, 0·201 and −0·040, with robust standard error estimates of 0·262, 0·064 and 0·076. These numbers are very close to those of the log-rank estimating functions shown in Table 2. Previously, Wei, Lin & Weissfeld (1989) formulated the marginal distributions of the first four recurrences with proportional hazards models. Their estimate for the overall treatment effect is 0·549, with a robust standard error estimate of 0·285.

Figure 1 shows the estimation for the mean frequency of recurrences from month 5 to month 60 for a patient on thiotepa who has a single initial tumour that is 1 centimetre in diameter. It is not surprising that the mean function of recurrences for such a patient is fairly low since thiotepa is effective in reducing tumour recurrences and a smaller number of initial tumours is associated with a lower recurrence rate.

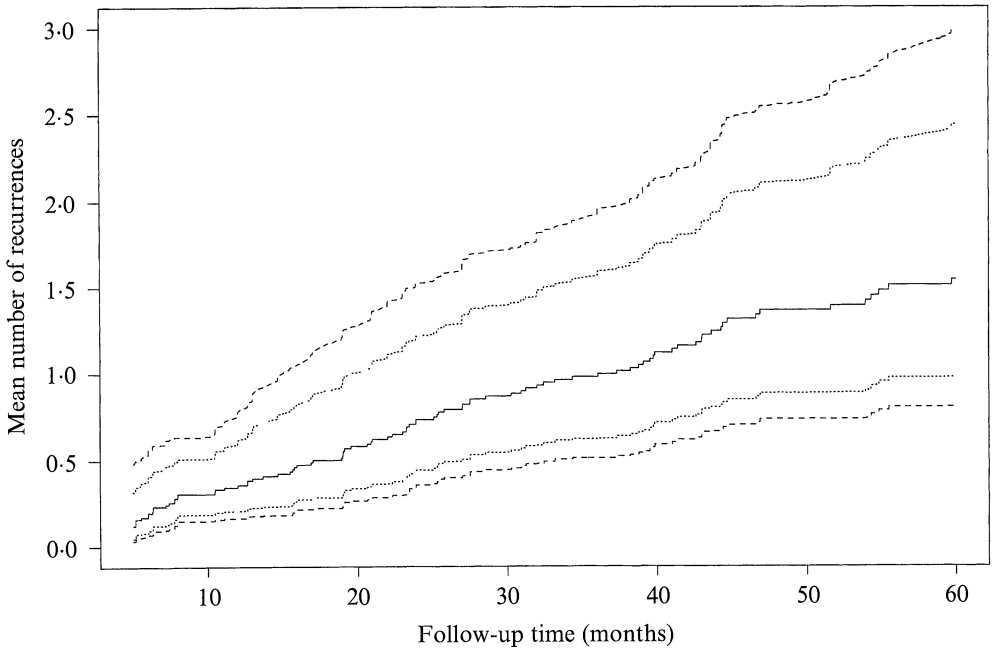

Fig. 1. Bladder tumour data: Estimated mean frequency of tumour recurrences over follow-up time for a patient on thiotepa who has a single initial tumour that is 1 centimetre in diameter. The point estimates are shown by the solid line, pointwise 95% Wald confidence intervals by dotted lines and 95% equal-precision confidence bands by dashed lines. The confidence intervals and bands are based on 1000 simulations.

4. Remarks

In general, it is hard to solve the estimating equation $U(\beta) = 0$ when the dimension of covariates is high. As mentioned in § 2, the Gehan estimating equation $U(\beta) = 0$ or $U(\beta) = G$ can be solved efficiently and reliably through the conventional linear programming technique. If the model fits the data reasonably well, then the solution to the Gehan estimating equation will be similar to the solutions to other estimating equations and may be used as the initial search value for the latter. It suffices for most practical purposes to make inferences based on the Gehan estimating function.

Lin & Wei (1992) provided a different approach to analysing multiple events data with accelerated failure time models. They formulated the marginal distribution for the time to each type of event measured from study entry with a univariate accelerated failure time model in the form of (3) and derived the joint distribution for the regression parameter estimators of the marginal models. Their method can only handle a small and equal number of recurrences per subject, and does not provide global estimation of the underlying mean function. The approach taken in this paper is more natural and efficient for handling recurrent events, especially when there are long strings of events or when the

numbers of events vary substantially among the study subjects. The generalisation of the accelerated failure time model provided in this paper parallels the Andersen–Gill/Lawless–Nadeau–Cook generalisation of the Cox model to counting processes, while the Lin & Wei method shares the spirit of the Wei et al. (1989) approach to analysing genuine multivariate failure time data.

We have so far restricted our attention to time-invariant covariates. For survival data, Cox & Oakes (1984, pp. 64–8), Robins & Tsiatis (1992) and Lin & Ying (1995) studied the accelerated failure time model with time-dependent covariates. To accommodate time-dependent covariates in the accelerated failure time model with recurrent events, we generalise equation (1) to

$$E\{N_i^*(t)|Z_i\} = \mu_0\left(\int_0^t e^{\beta_0' Z_i(s)}\, ds\right), \tag{6}$$

which may also be formulated in terms of equation (2) with the redefinition of

$$\tilde{T}_{ik}(\beta) = \int_0^{T_{ik}} e^{\beta' Z_i(s)}\, ds.$$

A similar modification is made to $\tilde{C}_i(\beta)$. In addition, redefine

$$U(\beta) = \sum_{i=1}^n \int_0^\infty \left\{Q_i(t;\beta) - \frac{\sum_{j=1}^n Y_j(t;\beta)Q_j(t;\beta)}{Y_\cdot(t;\beta)}\right\} d\tilde{N}_i(t;\beta),$$

$$D_i(\beta) = \int_0^\infty \left\{Q_i(t;\beta) - \frac{\sum_{j=1}^n Y_j(t;\beta)Q_j(t;\beta)}{Y_\cdot(t;\beta)}\right\} d\hat{M}_i(t;\beta),$$

where $Q_i(t;\beta)$ $(i=1,\ldots,n)$ are p-dimensional processes involving Z_i and possibly β. Then the arguments of Lin & Ying (1995), along with those given in the Appendix, can be used to show that the results of § 2 continue to hold for time-dependent covariates.

Lawless et al. (1997) considered the proportional means model

$$E\{dN_i^*(t)|Z_i\} = e^{\beta_0' Z_i(t)}\, d\mu_0(t), \tag{7}$$

and provided a class of estimators for β_0, which includes the Andersen–Gill estimator as a special case; Pepe & Cai (1993) studied similar models for rate functions. In general, models (6) and (7) are different. It would be worthwhile to develop methods for determining which of the two models is more appropriate for a given dataset.

Models (6) and (7) coincide if covariates are time-invariant and μ_0 is a straight line through the origin, i.e.

$$E\{N_i^*(t)|Z_i\} = e^{\beta_0' Z_i} t. \tag{8}$$

For the bladder tumour data, the fact that $\hat{\mu}_0(.)$ looks linear in Fig. 1 may explain why the Andersen–Gill estimates are similar to those of Table 2. It is interesting to study the asymptotic efficiency of the proposed $\hat{\beta}$ relative to the Lawless–Nadeau–Cook-type estimator under model (8). It can be shown that the asymptotic relative efficiency is equal to 1 at $\beta_0 = 0$ provided that the same weight function is used for both estimators. For nonzero β_0, the asymptotic relative efficiency is slightly below 1 when the underlying process is Poisson and the log-rank weight function is used for both estimators. The asymptotic efficiencies for non-Poisson processes require further investigation. Incidentally, for exponential survival times, the rank estimator under the accelerated failure

time model is more efficient than the maximum partial likelihood estimator under the equivalent proportional hazards model (Lai & Ying, 1991a).

Acknowledgement

This work was supported by the U.S. National Institutes of Health, National Science Foundation and National Security Agency. The authors are grateful to the Editor, an Associate Editor and a referee for their helpful comments.

Appendix

Asymptotic properties of $U(\beta)$, $\hat{\beta}$ and $\hat{\mu}_0(t;\hat{\beta})$

We assume the following regularity conditions:

(i) N_i and Z_i are bounded,

(ii) (N_i^*, C_i, Z_i) $(i=1,\ldots,n)$ are independent and identically distributed,

(iii) Q has bounded variation and converges almost surely to a continuous function q,

(iv) $C_i(\beta_0)$ has a bounded density and μ_0 has a bounded second derivative.

In this Appendix, the integration is taken from 0 to ∞ unless otherwise indicated, the summation is always taken over $i=1$ to n, and the limit taken as $n\to\infty$. We first study the process

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

Theorem 1. *Under conditions* (i)–(iii), $n^{-\frac{1}{2}}U(\beta_0,.)$ *converges weakly to a zero-mean Gaussian process with covariance function*

$$B(t,t^\dagger)=E\left[\int_0^t q(s)\{Z_1-\bar{z}(s)\}\,dM_1(s;\beta_0)\int_0^{t^\dagger} q(x)\{Z_1-\bar{z}(x)\}'\,dM_1(x;\beta_0)\right],$$

where $\bar{z}(t)=\lim\bar{Z}(t;\beta_0)$.

Proof. We first prove the weak convergence on a finite interval $[0,\tau]$ such that $\lim n^{-1}Y_{\cdot}(\tau;\beta_0)>0$. Let

$$U_M(t)=n^{-\frac{1}{2}}\sum M_i(t;\beta_0),\quad U_{MZ}(t)=n^{-\frac{1}{2}}\sum Z_iM_i(t;\beta_0).$$

For each t, both $U_M(t)$ and $U_{MZ}(t)$ are sums of n independent and identically distributed zero-mean terms. Thus, the finite-dimensional convergence of (U_M, U_{MZ}) follows easily from the standard multivariate central limit theorem. In addition, $\{M_i(t;\beta_0);\, i=1,\ldots,n\}$ and $\{Z_iM_i(t;\beta_0);\, i=1,\ldots,n\}$ can be written as sums/products of monotone functions and are therefore 'manageable' (Pollard, 1990, p. 38; Bilias, Gu & Ying, 1997, Theorem 2.1). It then follows from the functional central limit theorem (Pollard, 1990, p. 53) that (U_M, U_{MZ}) is tight and converges weakly to a zero-mean Gaussian process, denoted by $(\mathscr{W}_M, \mathscr{W}_{MZ})$. By the Skorokhod–Dudley–Wichura theorem (Shorack & Wellner, 1986, p. 47), we can construct in another probability space an equivalent process of (U_M, U_{MZ}) such that the weak convergence becomes almost sure convergence. Since $\bar{Z}(t;\beta_0)$ and $Q(t;\beta_0)$ are of bounded variation and converge almost surely to $\bar{z}(t)$ and $q(t)$, we can show through integration by parts that $n^{-\frac{1}{2}}U(\beta_0,t)$ converges weakly to the zero-mean Gaussian process

$$\int_0^t q(s)\,d\mathscr{W}_{MZ}(s)-\int_0^t q(s)\bar{z}(s)\,d\mathscr{W}_M(s).$$

A straightforward variance-covariance calculation verifies that $B(t,t^\dagger)$ is the covariance function of the limiting process. Finally, we can argue along the lines of Ying (1993) that the weak conver-

gence result holds without the tail restriction. □

To study the asymptotic properties of $\hat{\beta}$, we need to establish the asymptotic linearity of $U(\beta)$ in a neighbourhood of β_0. Let $u(\beta)$ be the limit of $n^{-1}U(\beta)$ and $\mathcal{N}(\beta_0)$ be a compact neighbourhood of β_0 on which $\|U(\beta)\|$ is minimised to obtain $\hat{\beta}$. It is not difficult to show that $n^{-1}U(\beta)$ converges to $u(\beta)$ uniformly on $\mathcal{N}(\beta_0)$.

THEOREM 2. *Under conditions* (i)–(iv), *for any sequence* $d_n \to 0$,

$$\sup_{\|\beta-\beta_0\|\leq d_n} \{\|U(\beta)-U(\beta_0)+An(\beta-\beta_0)\|/(n^{\frac{1}{2}}+n\|\beta-\beta_0\|)\} = o(1) \tag{A1}$$

almost surely, where

$$A = \int q(t)E[Y_1(t;\beta_0)\{Z_1-\bar{z}(t)\}^{\otimes 2}]\, d\{\dot{\mu}_0(t)t\},$$

$a^{\otimes 2} = aa'$ *and* $\dot{\mu}_0(t) = d\mu_0(t)/dt$. *Furthermore, if* $u(\beta) \neq 0$ *for all* $\beta \in \mathcal{N}(\beta_0)$ *but* $\beta \neq \beta_0$ *and* A *has full rank, then* $\hat{\beta}$ *is strongly consistent and* $n^{\frac{1}{2}}(\hat{\beta}-\beta_0)$ *converges in distribution to zero-mean normal with covariance matrix* $A^{-1}BA^{-1}$, *where* $B = B(\infty,\infty)$.

Proof. Write $U(\beta)-U(\beta_0)$ as

$$\begin{aligned}&\Big[\sum \int Q(t;\beta)\{Z_i-\bar{Z}(t;\beta)\}\{d\tilde{N}_i(t;\beta)-Y_i(t;\beta)\,d\mu_0(t\,e^{(\beta_0-\beta)'Z_i})\}\\ &\quad-\sum \int Q(t;\beta_0)\{Z_i-\bar{Z}(t;\beta_0)\}\{d\tilde{N}_i(t;\beta_0)-Y_i(t;\beta_0)\,d\mu_0(t)\}\Big]\\ &\quad+\sum \int Q(t;\beta)\{Z_i-\bar{Z}(t;\beta)\}Y_i(t;\beta)\,d\{\mu_0(t\,e^{(\beta_0-\beta)'Z_i})-\mu_0(t)\}.\end{aligned} \tag{A2}$$

Applying the technique of Ying (1993, Theorem 1), we can show that the first term of (A2) is of order $o(n^{\frac{1}{2}})$. By Taylor series expansion,

$$\mu_0(t\,e^{(\beta_0-\beta)'Z_i})-\mu_0(t) = \{\dot{\mu}_0(t)+o(1)\}tZ_i'(\beta_0-\beta).$$

Thus, the second term of (A2) is

$$\sum \int Q(t;\beta)\{Z_i-\bar{Z}(t;\beta)\}Y_i(t;\beta)Z_i'\,d\{\dot{\mu}_0(t)t\}(\beta_0-\beta)+o(n\|\beta-\beta_0\|) = An(\beta_0-\beta)+o(n\|\beta-\beta_0\|)$$

almost surely. Combining the preceding approximation with (A2) gives (A1).

Now, since $n^{-1}U(\beta)\to u(\beta)$ uniformly in $\mathcal{N}(\beta_0)$ and $u(\beta)\neq 0$ for $\beta\neq\beta_0$, we have $\hat{\beta}\to\beta_0$. In addition, the definition of $\hat{\beta}$ and (A1) imply that $n^{-1}U(\hat{\beta}) = o(1)$. Thus, it follows from (A1) that $n^{\frac{1}{2}}(\hat{\beta}-\beta_0) = A^{-1}n^{-\frac{1}{2}}U(\beta_0)+o(1)$ when A is invertible. The asymptotic normality for $\hat{\beta}$ then follows from Theorem 1. □

The consistency of the covariance matrix estimators $V(\beta_0)$ and $V(\hat{\beta})$ is a special case of the following more general result.

THEOREM 3. *Suppose that the assumptions of Theorem* 2 *are satisfied. If* $\beta_n^*\to\beta_0$, *then* $V(\beta_n^*)\to B$.

Proof. By the uniform strong law of large numbers (Pollard, 1990, p. 41),

$$n^{-1}\sum \tilde{N}_i(t;\beta)\to E\tilde{N}_1(t;\beta),\quad n^{-1}Y_.(t;\beta)\to EY_1(t;\beta)$$

uniformly in t and β. These approximations, together with similar ones for $\bar{Z}$, can be used to show that

$$n^{-1}\sum \left\| D_i(\beta_n^*)-\int q(t)\{Z_i-\bar{z}(t)\}\,dM_i(t;\beta_0)\right\|^2\to 0.$$

Thus, to prove that $V(\beta_n^*) \to B$, it suffices to show that

$$n^{-1} \sum \left[\int q(t)\{Z_i - \bar{z}(t)\}\, dM_i(t;\, \beta_0) \right]^{\otimes 2} \to B,$$

but the latter readily follows from the strong law of large numbers. □

Finally, we deal with the weak convergence of W and $\hat{W}$.

THEOREM 4. *Under conditions* (i)–(iv), *both* $W(.)$ *and* $\hat{W}(.)$ *converge weakly to a zero-mean Gaussian process with covariance function*

$$\sigma(t, t^\dagger) = E\left(\left[\int_0^t \frac{dM_1(s;\, \beta_0)}{E\{Y_1(s;\, \beta_0)\}} + b'(t) A^{-1} \int_0^\infty q(s)\{Z_1 - \bar{z}(s)\}\, dM_1(s;\, \beta_0) \right] \right.$$
$$\left. \times \left[\int_0^{t^\dagger} \frac{dM_1(s;\, \beta_0)}{E\{Y_1(s;\, \beta_0)\}} + b'(t^\dagger) A^{-1} \int_0^\infty q(s)\{Z_1 - \bar{z}(s)\}\, dM_1(s;\, \beta_0) \right] \right),$$

where

$$b(t) = -\int_0^t \bar{z}(s)\, d\{\dot{\mu}_0(s)s\}.$$

Proof. An asymptotic linearity similar to (A1) of Theorem 2 holds for $\hat{\mu}_0$, that is

$$\sup_{t \in [0,\tau],\, \|\beta - \beta_0\| \leqslant d_n} \| n^{\frac{1}{2}} \{\hat{\mu}_0(t;\, \beta) - \hat{\mu}_0(t;\, \beta_0)\} - b'(t) n^{\frac{1}{2}} (\beta - \beta_0) \| = o(1) \tag{A3}$$

for any $d_n \to 0$ and for τ such that $\lim n^{-1} Y_\cdot(\tau;\, \beta_0) > 0$. Applying (A1) and (A3) with $\beta = \hat{\beta}$, we have, uniformly in $t \in [0, \tau]$,

$$W(t) = n^{\frac{1}{2}} \{\hat{\mu}_0(t;\, \beta_0) - \mu_0(t)\} + b'(t) A^{-1} n^{-\frac{1}{2}} U(\beta_0) + o(1)$$
$$= n^{-\frac{1}{2}} \sum \int_0^t \frac{dM_i(s;\, \beta_0)}{E\{Y_1(s;\, \beta_0)\}} + b'(t) A^{-1} n^{-\frac{1}{2}} \sum \int_0^\infty q(s)\{Z_i - \bar{z}(s)\}\, dM_i(s;\, \beta_0) + o(1), \tag{A4}$$

where the second equality follows from the arguments given in the proof of Theorem 1. In view of (A4), the finite-dimensional convergence of $W(.)$ to the desired Gaussian process follows from the multivariate central limit theorem together with a straightforward covariance calculation. As in the proof of Theorem 1, the tightness can be established via the modern empirical process theory (Pollard, 1990).

Applying (A3) twice with $\beta = \hat{\beta}$ and $\beta = \hat{\beta}^*$, we obtain

$$\hat{W}(t) = b'(t) n^{\frac{1}{2}} (\hat{\beta} - \hat{\beta}^*) + n^{-\frac{1}{2}} \sum n H_i(t;\, \hat{\beta}) G_i + o(1).$$

On the other hand, applying (A1) twice and noting that $n^{-\frac{1}{2}} U(\hat{\beta}) = o(1)$, we have

$$n^{-\frac{1}{2}} U(\hat{\beta}^*) = A n^{\frac{1}{2}} (\hat{\beta} - \hat{\beta}^*) + o(1).$$

Thus,

$$\hat{W}(t) = b'(t) A^{-1} n^{-\frac{1}{2}} \sum D_i(\hat{\beta}) G_i + n^{-\frac{1}{2}} \sum n H_i(t;\, \hat{\beta}) G_i + o(1).$$

Then, by the multivariate central limit theorem and a simple covariance calculation, $\hat{W}$ converges in finite-dimensional distributions to the limiting Gaussian process of W. The modern empirical process theory can again be used to prove the tightness. □

REFERENCES

ANDERSEN, P. K. & GILL, R. D. (1982). Cox's regression model for counting processes: a large sample study. *Ann. Statist.* **10**, 1100–20.

BARRODALE, I. & ROBERTS, F. (1973). An improved algorithm for discrete l_1 linear approximation. *SIAM J. Numer. Anal.* **10**, 839–48.

BILIAS, Y., GU, M. & YING, Z. (1997). Towards a general asymptotic theory for Cox model with staggered entry. *Ann. Statist.* **25**, 662–82.

BUCKLEY, J. & JAMES, I. (1979). Linear regression with censored data. *Biometrika* **66**, 429–36.

BYAR, D. P. (1980). The Veterans Administration study of chemoprophylaxis for recurrent stage I bladder tumors: comparisons of placebo, pyridoxine, and topical thiotepa. In *Bladder Tumors and Other Topics in Urological Oncology*, Ed. M. Pavone-Macaluso, P. H. Smith and F. Edsmyn, pp. 363–70. New York: Plenum.

COX, D. R. (1972). Regression models and life-tables (with Discussion). *J. R. Statist. Soc.* B **34**, 187–20.

COX, D. R. & OAKES, D. (1984). *Analysis of Survival Data.* London: Chapman & Hall.

FLEMING, T. R. & HARRINGTON, D. P. (1991). *Counting Processes and Survival Analysis.* New York: Wiley.

GUMBEL, E. J. (1960). Bivariate exponential distributions. *J. Am. Statist. Assoc.* **55**, 698–707.

KALBFLEISCH, J. D. & PRENTICE, R. L. (1980). *The Statistical Analysis of Failure Time Data.* New York: Wiley.

LAI, T. L. & YING, Z. (1991a). Rank regression methods for left-truncated and right-censored data. *Ann. Statist.* **19**, 531–56.

LAI, T. L. & YING, Z. (1991b). Large sample theory of a modified Buckley–James estimator for regression analysis with censored data. *Ann. Statist.* **19**, 1370–402.

LAWLESS, J. F. & NADEAU, C. (1995). Some simple and robust methods for the analysis of recurrent events. *Technometrics* **37**, 158–68.

LAWLESS, J. F., NADEAU, C. & COOK. R. J. (1997). Analysis of mean and rate functions for recurrent events. In *Proceedings of the First Seattle Symposium in Biostatistics: Survival Analysis*, Ed. D. Y. Lin and T. R. Fleming, pp. 37–49. New York: Springer-Verlag.

LIN, D. Y. (1991). Goodness-of-fit analysis for the Cox regression model based on a class of parameter estimators. *J. Am. Statist. Assoc.* **86**, 725–8.

LIN, D. Y., FLEMING, T. R. & WEI, L. J. (1994). Confidence bands for survival curves under the proportional hazards model. *Biometrika* **81**, 73–81.

LIN, D. Y. & GEYER, C. J. (1992). Computational methods for semiparametric linear regression with censored data. *J. Comp. Graph. Statist.* **1**, 77–90.

LIN, D. Y. & YING, Z. (1995). Semiparametric inference for the accelerated life model with time-dependent covariates. *J. Statist. Plan Inf.* **44**, 47–63.

LIN, J. S. & WEI, L. J. (1992). Linear regression analysis for multivarite failure time observations. *J. Am. Statist. Assoc.* **87**, 1091–7.

PARZEN, M. I., WEI, L. J. & YING, Z. (1994). A resampling method based on pivotal estimating functions. *Biometrika* **81**, 341–50.

PEPE, M. S. & CAI, J. (1993). Some graphical displays and marginal regression analyses for recurrent failure times and time dependent covariates. *J. Am. Statist. Assoc.* **88**, 811–20.

POLLARD, D. (1990). *Empirical Processes: Theory and Applications.* Hayward, CA: Institute of Mathematical Statistics.

PRENTICE, R. L. (1978). Linear rank tests with right censored data. *Biometrika*, **65**, 167–79.

REID, N. (1994). A conversation with Sir David Cox. *Statist. Sci.* **9**, 439–55.

RITOV, Y. (1990). Estimation in a linear regression model with censored data. *Ann. Statist.* **18**, 303–28.

ROBINS, J. M. & TSIATIS, A. A. (1992). Semiparametric estimation of an accelerated failure time model with time-dependent covariates. *Biometrika* **79**, 311–9.

SHORACK, G. R. & WELLNER, J. A. (1986). *Empirical Processes with Applications to Statistics.* New York: Wiley.

TSIATIS, A. A. (1990). Estimating regression parameters using linear rank tests for censored data. *Ann. Statist.* **18**, 354–72.

WEI, L. J., LIN, D. Y. & WEISSFELD, L. (1989). Regression analysis of multivariate incomplete failure time data by modeling marginal distributions. *J. Am. Statist. Assoc.* **84**, 1065–73.

WEI, L. J., YING, Z. & LIN, D. Y. (1990). Linear regression analysis of censored survival data based on rank tests. *Biometrika* **77**, 845–51.

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

[*Received May* 1997. *Revised December* 1997]