Acta Mathematicae Applicatae Sinica, English Series
Vol. 35, No. 3 (2019) 549–563
https://doi.org/10.1007/s10255-019-0831-9
http://www.ApplMath.com.cn & www.SpringerLink.com

Acta Mathematicae Applicatae Sinica, English Series

On a General Class of Semiparametric Hazards Regression Models for Recurrent Gap Times

Qin JIANG[1,2,†], **Hui ZHAO**[3], **Hong QIN**[1]

[1]Department of Statistics, Central China Normal University, Wuhan 430079, China
(E-mail: jiangqin999@126.com)
[2]Department of Mathematics, Huanggang Normal University, Huanggang 438000, China
[3]Department of Statistics, Zhongnan University of Economics and Law, Wuhan 430073, China

Abstract In the article, we investigate a general class of semiparametric hazards regression models for recurrent gap times. The general class includes the proportional hazards model, the accelerated failure time model and the accelerated hazards models as special cases. The model is flexible in modelling recurrent gap times since a covariate effect is identified as having two separate components, namely a time-scale change on hazard progression and a relative hazards ratio. In order to infer the model parameters, the procedure is proposed based on estimating equations. The asymptotic properties of the proposed estimators are established and the finite sample properties are investigated via simulation studies. In addition, a lack of fit test is presented to assess the adequacy of the model and an application of data from a bladder cancer study is reported for illustration.

Keywords gap times; model checking; recurrent events; estimating equations; semiparametric model
2000 MR Subject Classification 62N02; 62G05

1 Introduction

Recurrent events are frequently encountered in biomedical studies, reliability studies and social sciences when each subject experiences a particular event repeatedly over time. Here the event could be, for instance, the occurrence of a certain disease, of hospitalization, or of breakdowns of an automobile. It is challenging to analyze such recurrent events data on account of the dependence of the recurrent event times within each individual and the presence of censoring such as the loss to follow-up. For recurrent event data, various statistical methods have been proposed, such as some intensity based methods (see [2,27,39]), some frailty model approaches (see [11,22,23,40]) and some marginal means and rates models (see [7,13,17,30,34,35]).

In many applications, researchers are naturally interested in the gap times between recurrent events, and inferring effects of covariates such as age and treatment on the gap times. Many semiparametric hazards regression models are proposed for gap times. For example, Huang and Chen[10], Schaubel and Cai[29] and Darlington and Dixon[8] studied proportional hazards models. Sun, Park, and Sun[33] considered additive hazards models. Chang[5] and Strawderman[31] considered accelerated failure time models for logarithm transformed gap times. Lu[19] studied the semiparametric linear transformation models for the gap times, which include the proportional hazards and proportional odds models as special cases, but do not contain the additive hazards model as a special case. Kang et al.[12] considered a class of transformed hazards models for recurrent gap time data, including both the proportional and additive hazards models

Manuscript received April 26, 2018. Revised November 1, 2018.
This paper was supported by the National Natural Science Foundation of China (No.11471135, 11861030), by the Natural Science Foundation of Hubei Province (No. 2018CFC825) and by National Statistical Scientific Research Project (No. 2018LY17).

as special cases. Moreover, some nonparametric models were discussed for the gap time (see [9,16,25,36]). Luo and Huang[20,21] demonstrated that many existing methods for recurrent gap time data can be viewed as weighted risk-set methods.

In this paper, we pay attention to the analysis of recurrent gap times with a general flexible class of semiparametric hazards regression models, which was studied by Chen and Jewell[6] for univariate survival data. In the general class of models, a covariate's effect is identified as having two separate components, namely a time-scale change on hazard progression and a relative hazards ratio. The class includes the proportional hazards model, the accelerated failure time model and the accelerated hazards models as special cases. Estimation of model parameters and associated statistical properties are also investigated under a set of mild regularity conditions.

The remainder of the paper is organized as follows. In Section 2, we introduce data structure and the models. Estimation procedures are presented for the model parameters. The asymptotic properties of the proposed estimators are established. In Section 3, a technique is developed for checking the adequacy of the general model. In Section 4, some results are reported from simulation studies conducted for evaluating the proposed methods. In Section 5, the methodology is applied to a bladder cancer study data of Byar[3], followed by concluding remarks in Section 6. The details of the proofs are relegated to Appendix.

2 Model and Estimation Procedure

2.1 The Model

Suppose that a total of n subjects are observed over time and each subject experiences recurrences of the same event (see [4,10,12,33]). For subject i, let T_{ij} denote the time from the $(j-1)$th to the jth occurrence of the event, $j = 1, 2, \cdots,\ i = 1, 2, \cdots, n$. Obviously, $T_{i1} + \cdots + T_{ij}$ is the jth recurrent event time. Also let Z_i denote the p-dimensional vector of time-independent covariates associated with subject i, C_i denote the follow-up or censoring time and $N_i = \{T_{ij} : j = 1, 2, \cdots\}$. Assume that $\{(N_i, C_i, Z_i);\ i = 1, 2, \cdots, n\}$ are independent and identically distributed, and N_i is independent of C_i given Z_i. Let M_i be the index of observed gap times for subject i, such that

$$\sum_{j=1}^{M_i-1} T_{ij} \le C_i, \qquad \sum_{j=1}^{M_i} T_{ij} > C_i.$$

Then observed data are $\{T_{i1}, \cdots, T_{i,M_i-1}, C_i, Z_i\}$ which mean the first $M_i - 1$ gap times are observed, but T_{i,M_i} is censored at

$$T^+_{i,M_i} = C_i - \sum_{j=1}^{M_i-1} T_{ij}.$$

From Huang and Chen[10], Sun et al.[33] and Kang et al.[12], we suppose that each individual recurrent event process is a renewal process. This implies that the observed complete gap times $\{T_{ij},\ j = 1, \cdots, M_i - 1\}$ are identically distributed for given (C_i, M_i, T^+_{i,M_i}).

Let $\lambda(t|Z_i)$ be the hazard function of T_{ij} given Z_i. In the following, we investigate a general class of semiparametric hazards regression models defined as

$$\lambda(t|Z_i) = \lambda_0(te^{\beta'_{10}Z_i}) \exp(\beta'_{20}Z_i) \tag{1}$$

for T_{ij} given Z_i, where β_{10},β_{20} are the true p-vectors regression parameters, and $\lambda_0(t)$ is an an unspecified baseline hazards function.

To simplify our discussion without losing generality, as an example we only consider the covariate Z_i is binary. As seen in the following development, results can be easily extended to the situation of multiple covariates.

Suppose that $Z_i = 1$ stands for the treatment group and $Z_i = 0$ for the control group, as in a randomised clinical trial. Then Model (1) reduces to

$$\lambda_1(t) = \lambda_0(te^{\beta_{10}})\exp(\beta_{20}), \tag{2}$$

where $\lambda_1(\cdot)$ is the treatment group's hazard function, $\lambda_0(\cdot)$ is the control group's hazard function, and $\beta_{10}, \beta_{20} \in R$ are the parameters. Clearly, Model (2) includes the proportional hazards model, the accelerated failure time model and the accelerated hazards model as subclasses of models: when $\beta_{10} = 0$, Model (2) becomes the proportional hazards model with a proportionality constant of $\exp(\beta_{20})$; when $\beta_{10} = \beta_{20}$, model (2) becomes the accelerated failure time model for counting processes with a time scale change of $\exp(\beta_{10}) = \exp(\beta_{20})$; when $\beta_{20} = 0$, model (2) reduces to accelerated hazards model with the hazards progression time ratio of $\exp(\beta_{10})$.

The two parameters β_{10} and β_{20} can be interpreted as measuring two different effects the covariate may have on recurrent gap times. The first parameter β_{10} identifies the acceleration or deceleration of hazard progression in the treatment group, while β_{20} characterises the relative hazard after adjusting for the different hazard progressions in the treatment and the control groups. Therefore, model (1) implies that the treatment can alter both the magnitude of hazard and the pace of progression simultaneously.

Correctly identifying and estimating these two components may better describe a given recurrent gap times data, although the main value of the general Model (2) may be in quantifying and highlighting differences between the three subclasses of models that are commonly used individually.

Flexibility of the general model may lead to concern of identifiability. Following the arguments of Chen and Jewell[6] and Sun and Su[34], we show 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* $\lambda_0(t) = \Sigma_{k=-\infty}^{+\infty} c_k t^k$ *for any* $t \in [0, t_0]$, *then* β_{10} *and* β_{20} *are identifiable if and only if there exist* $k_1, k_2 \in \{0, \pm 1, \pm 2, \cdots\}$ *such that* $k_1 \neq k_2$ *and* $c_{k_1}, c_{k_2} \neq 0$.

In the paper, we assume that β_{10} and β_{20} are identifiable.

2.2 Inference Method

Our inference procedure is based on the establishment of a connection between a subset of the observed gap times and clustered survival data. Let $\triangle_i = I(M_i > 1),\ S_i = \max(M_i - 1, 1)$, and

$$X_{ij} = \begin{cases} T_{ij}, & \text{if } \triangle_i = 1, \\ T_{ij}^{+}, & \text{if } \triangle_i = 0, \end{cases} j = 1, 2, \cdots, S_i.$$

Then $\{X_{ij}, \triangle_i, Z_i,\ j = 1, \cdots, S_i\}$ $(i = 1, \cdots, n)$ can be treated as clustered survival data. Since the cluster size is informative, the censored gap time needs to be removed for $M_i > 1$.

Define $N_{ij}(t) = \Delta_i I(X_{ij} \le t)$ and $Y_{ij}(t; \beta_1) = I(X_{ij} \ge te^{-\beta_1' Z_i})$. Denote $\widetilde{N}_{ij}(t; \beta_1) = N_{ij}(te^{-\beta_1' Z_i})$. Let $\beta = (\beta_1', \beta_2')'$, $\beta_0 = (\beta_{10}', \beta_{20}')'$ and

$$M_{ij}(t; \beta, \Lambda_0) = \widetilde{N}_{ij}(t; \beta_1) - \int_0^t Y_{ij}(s; \beta_1)\exp\{(\beta_2 - \beta_1)' Z_i\} d\Lambda_0(s).$$

According to Model (1), we know that $M_{ij}(t; \beta_0, \Lambda_0)$ $(j = 1, \cdots, S_i;\ i = 1, \cdots, n)$ are zero-mean stochastic processes but not martingales, since the hazard function is not modeled conditionally on the recurrent event process. Based on this fact and using the generalized estimating equation approach[14], it is reasonable to estimate $\Lambda_0(t)$ and β_0 with the following equations

$$\sum_{i=1}^{n}\frac{1}{S_i}\sum_{j=1}^{S_i}\int_0^t dM_{ij}(s;\beta,\Lambda_0)=0, \tag{3}$$

$$\sum_{i=1}^{n}\frac{1}{S_i}\sum_{j=1}^{S_i}\int_0^t Z_i dM_{ij}(s;\beta,\Lambda_0)=0, \tag{4}$$

$$\sum_{i=1}^{n}\frac{1}{S_i}\sum_{j=1}^{S_i}\int_0^t W(s,Z_i;\beta)dM_{ij}(s;\beta,\Lambda_0)=0, \tag{5}$$

for all $0 \le t \le \tau$, where τ is a prespecified constant satisfying $P(C_i \ge \tau e^{-\beta_{10}' Z_i}) > 0$, $i = 1,2,\cdots,n$, $W(s,Z_i;\beta)$ is a known p-dimensional weight function of s, Z_i and β, but not in the span of the function 1 and Z_i. We denote the estimator of $\Lambda_0(t)$ by $\widehat{\Lambda}_0(t,\beta)$, which can be specified from (3) in the following explicit expression:

$$\widehat{\Lambda}_0(t,\beta)=\sum_{i=1}^{n}\frac{1}{S_i}\sum_{j=1}^{S_i}\int_0^t \frac{d\widetilde{N}_{ij}(s;\beta_1)}{\sum_{i=1}^{n}\frac{1}{S_i}\sum\limits_{j=1}^{S_i}Y_{ij}(s;\beta_1)\exp\{(\beta_2-\beta_1)'Z_i\}}, \qquad 0\le t\le\tau. \tag{6}$$

In order to estimate β_0, replacing $\Lambda_0(t)$ with its estimator $\widehat{\Lambda}_0(t,\beta)$ in (4) and (5), we can use the following two estimating equations:

$$U_1(\beta)=\sum_{i=1}^{n}\frac{1}{S_i}\sum_{j=1}^{S_i}\int_0^\tau \big\{Z_i-\overline{Z}(t;\beta)\big\}d\widetilde{N}_{ij}(t,\beta_1), \tag{7}$$

$$U_2(\beta)=\sum_{i=1}^{n}\frac{1}{S_i}\sum_{j=1}^{S_i}\int_0^\tau \big\{W(t,Z_i;\beta)-\overline{W}(t;\beta)\big\}d\widetilde{N}_{ij}(t,\beta_1), \tag{8}$$

where

$$\overline{Z}(t;\beta)=\frac{\sum\limits_{i=1}^{n}\frac{1}{S_i}\sum\limits_{j=1}^{S_i}Y_{ij}(t;\beta_1)\exp\{(\beta_2-\beta_1)'Z_i\}Z_i}{\sum\limits_{i=1}^{n}\frac{1}{S_i}\sum\limits_{j=1}^{S_i}Y_{ij}(t;\beta_1)\exp\{(\beta_2-\beta_1)'Z_i\}},$$

$$\overline{W}(t;\beta)=\frac{\sum\limits_{i=1}^{n}\frac{1}{S_i}\sum\limits_{j=1}^{S_i}Y_{ij}(t;\beta_1)\exp\{(\beta_2-\beta_1)'Z_i\}W(t,Z_i;\beta)}{\sum\limits_{i=1}^{n}\frac{1}{S_i}\sum\limits_{j=1}^{S_i}Y_{ij}(t;\beta_1)\exp\{(\beta_2-\beta_1)'Z_i\}}.$$

Denote $Z_i^*(t;\beta) = (Z_i', W(t,Z_i;\beta)')'$, $\overline{Z}^*(t;\beta) = (\overline{Z}(t,\beta)', \overline{W}(t;\beta)')'$ and $U(\beta) = (U_1(\beta)', U_2(\beta)')'$. Thus, we rewrite (7) and (8) as the following equation

$$U(\beta)=\sum_{i=1}^{n}\frac{1}{S_i}\sum_{j=1}^{S_i}\int_0^\tau \big\{Z_i^*(t;\beta)-\overline{Z}^*(t;\beta)\big\}d\widetilde{N}_{ij}(t,\beta_1). \tag{9}$$

That is, we can estimate β_0 by (9). However, $U(\beta)$ is a discontinuous function of β_1, it is usually unavailable to get a unique solution for β in $U(\beta)=0$. We may define a solution $\widehat{\beta}$ as a zero-crossing of (9) or as the minimiser of the Euclidean norm of $\|U(\beta)\|$ (see [6,15,34]).

In our numerical studies, we use the "fminsearch" function in MATLAB to obtain the minimiser of$\|U(\beta)\|$. The Nelder-Mead direct search algorithm is adopted and its accuracy is set to be less than 0.0001.

2.3 Asymptotic Properties

In general, in order to establish the asymptotic properties of β, we first establish the asymptotic properties of $U(\beta_0)$. The asymptotic properties of $U(\beta_0)$ are summarized in the following theorems with the proofs given in Appendix.

Theorem 1. *Under Conditions (C1)–(C5) stated in Appendix, $n^{-1/2}U(\beta_0)$ is asymptotically normal with mean zero and covariance matrix $\Sigma = E\{d_i d_i'\}$ which can be consistently estimated by $\widehat{\Sigma}$, where*

$$
\begin{aligned}
d_i &= \frac{1}{S_i}\sum_{j=1}^{S_i}\int_0^\tau \{Z_i^*(t;\beta_0) - \overline{z}^*(t)\}dM_{ij}(t;\beta_0),\\
\widehat{\Sigma} &= n^{-1}\sum_{i=1}^{n} D_i(\widehat{\beta})D_i(\widehat{\beta})',\\
\overline{z}^*(t) &= (\overline{z}(t)', \overline{w}(t)')',\\
D_i(\beta) &= \frac{1}{S_i}\sum_{j=1}^{S_i}\int_0^\tau \{Z_i^*(t;\beta) - \overline{Z}^*(t;\beta)\}d\widehat{M}_{ij}(t;\beta),\\
\widehat{M}_{ij}(t,\beta) &= \widetilde{N}_{ij}(t;\beta_1) - \int_0^t Y_{ij}(s;\beta_1)\exp\{(\beta_2-\beta_1)'Z_i\}d\widehat{\Lambda}_0(s),
\end{aligned}
$$

and $\overline{z}(t)$ and $\overline{w}(t)$ are the limits of $\overline{Z}(t;\beta)$ and $\overline{W}(t;\beta)$ respectively.

In the following, let

$$
\begin{aligned}
S^{(0)}(t;\beta) &= n^{-1}\sum_{i=1}^{n}\frac{1}{S_i}\sum_{j=1}^{S_i} Y_{ij}(t;\beta_1)\exp\{(\beta_2-\beta_1)'Z_i\},\\
S^{(1)}(t;\beta) &= n^{-1}\sum_{i=1}^{n}\frac{1}{S_i}\sum_{j=1}^{S_i} Y_{ij}(t;\beta_1)\exp\{(\beta_2-\beta_1)'Z_i\}Z_i,\\
S_z^{(2)}(t;\beta) &= n^{-1}\sum_{i=1}^{n}\frac{1}{S_i}\sum_{j=1}^{S_i} Y_{ij}(t;\beta_1)\exp\{(\beta_2-\beta_1)'Z_i\}Z_i^{\otimes 2},\\
S_w^{(2)}(t;\beta)) &= n^{-1}\sum_{i=1}^{n}\frac{1}{S_i}\sum_{j=1}^{S_i} Y_{ij}(t;\beta_1)\exp\{(\beta_2-\beta_1)'Z_i\}W_i(t,Z_i;\beta)Z_i'.
\end{aligned}
$$

Next, the asymptotic properties of $\widehat{\beta}$ are read as follows.

Theorem 2. *Under the Hypotheses (C1)–(C6) stated in Appendix, the estimator $\widehat{\beta}$ is strongly consistent and $n^{1/2}(\widehat{\beta}-\beta_0)$ converges in distribution to zero-mean normal with covariance matrix $A^{-1}\Sigma(A^{-1})'$ that can be consistently estimated by $\widehat{A}^{-1}\widehat{\Sigma}(\widehat{A}^{-1})'$, where*

$$
\begin{aligned}
\widehat{A} &= \begin{pmatrix} \widehat{A}_{11} & \widehat{A}_{12}\\ \widehat{A}_{21} & \widehat{A}_{22}\end{pmatrix},\\
\widehat{A}_{11} &= \int_0^\tau [S_z^{(2)}(t,\widehat{\beta}) - \overline{Z}(t,\widehat{\beta})^{\otimes 2}S^{(0)}(t,\widehat{\beta})]d\{\widehat{\lambda}_0(t)t\},
\end{aligned}
$$

$$\widehat{A}_{12}=\int_0^\tau [S_z^{(2)}(t,\widehat{\beta})-\overline{Z}(t,\widehat{\beta})S^{(1)}(t,\widehat{\beta})']d\{\widehat{\Lambda}_0(t)\},$$
$$\widehat{A}_{21}=\int_0^\tau [S_w^{(2)}(t,\widehat{\beta})-\overline{W}(t,\widehat{\beta})\overline{Z}(t,\widehat{\beta})S^{(0)}(t,\widehat{\beta})']d\{\widehat{\lambda}_0(t)t\},$$
$$\widehat{A}_{22}=\int_0^\tau [S_w^{(2)}(t,\widehat{\beta})-\overline{W}(t,\widehat{\beta})S^{(1)}(t,\widehat{\beta})']d\{\widehat{\Lambda}_0(t)\},$$

and $\widehat{\lambda}_0(t)=h^{-1}\int K\big(\frac{u-t}{h}\big)d\widehat{\Lambda}_0(u)$ *is a density-type estimator of* $\dot{\Lambda}_0(t)$, h *is the bandwidth and* $K(\cdot)$ *is a kernel function with a compact support*[28].

Based on the above result, we see that the estimator $\widehat{A}$ requires an estimate for the derivative of $\Lambda_0(t)$. Because such density-type estimator $\widehat{\lambda}_0(t)$ tends to be numerically unstable, the resulting variance estimator for $\widehat{\beta}$ will also be unreliable. In order to obtain a stable variance estimate of $\widehat{\beta}$, via Parzen et al.[24] and Sun and Su[34], we adapt a resampling technique. Specifically, let $\widehat{\beta}^*$ be the solution to

$$U(\beta)=\sum_{i=1}^n D_i(\widehat{\beta})G_i, \tag{10}$$

where $\{G_1,\cdots,G_n\}$ are independent standard normal variables. Following the arguments of Parzen et al.[24], Lin et al.[15] and Sun and Su[34], the asymptotic distribution of $n^{1/2}(\widehat{\beta}-\beta_0)$ can be approximated by the conditional distribution of $n^{1/2}(\widehat{\beta}^*-\beta_0)$ given the data $\{N_{ij}(\cdot),Y_{ij},Z_i,S_i\}$ $(i=1,\cdots,n,\ j=1,\cdots,S_i)$. To approximate the distribution of $\widehat{\beta}$, we produce a large number of realisations of $\widehat{\beta}^*$ by repeatedly generating the random samples $\{G_1,\cdots,G_n\}$ while fixing the data $\{N_{ij}(\cdot),Y_{ij},Z_i,S_i\}(i=1,\cdots,n,\ j=1,\cdots,S_i)$ at their observed values. The covariance matrix of $\widehat{\beta}$ can then be approximated by the empirical covariance matrix of $\widehat{\beta}^*$. Hence, confidence intervals for β_0 can be constructed using the empirical distribution of $\widehat{\beta}^*$.

Theorem 3. $\widehat{\Lambda}_0(t)$ *converges almost surely to* $\Lambda_0(t)$ *uniformly on* $[0,\tau]$, *and the process* $V(t)=n^{1/2}\{\widehat{\Lambda}_0(t)-\Lambda_0(t)\}$, $0\le t\le\tau$, *converges weakly to a mean-zero Gaussian process whose covariance function at* (s,t) *can be consistently estimated by* $\widehat{\Gamma}(s,t)=n^{-1}\sum_{i=1}^n\widehat{\Psi}_i(s)\widehat{\Psi}_i(t)$, *where*

$$\widehat{\Psi}_i(t)=\frac{1}{S_i}\sum_{j=1}^{S_i}\Big(\int_0^t\frac{d\widehat{M}_{ij}(s,\widehat{\beta})}{S^{(0)}(s;\widehat{\beta})}-\widehat{H}(t)'\widehat{A}^{-1}\int_0^\tau\{Z_i^*(u,\widehat{\beta})-\overline{Z}^*(u,\widehat{\beta})\}d\widehat{M}_{ij}(u;\widehat{\beta})\Big),$$
$$\widehat{H}_1(t)=\int_0^t\overline{Z}(u,\widehat{\beta})d\{\widehat{\lambda}_0(u)u\},\qquad \widehat{H}_2(t)=\int_0^t\frac{S^{(1)}(u,\widehat{\beta})}{S^{(0)}(u,\widehat{\beta})}d\{\widehat{\Lambda}_0(u)\},$$

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

As in the case of $\widehat{\beta}$, it is difficult to estimate the asymptotic covariance function of $V(t)$ analytically. Following Lin et al.[15] and Sun and Su[34], we show that the asymptotic distribution of $V(t)$ can be approximated by the conditional distribution of $\widehat{V}(t)$, where

$$\widehat{V}(t)=n^{1/2}\{\widehat{\Lambda}_0(t;\widehat{\beta})-\widehat{\Lambda}_0(t;\widehat{\beta}^*)\}+n^{-1/2}\sum_{i=1}^n\frac{1}{S_i}\sum_{j=1}^{S_i}\int_0^t\frac{d\widehat{M}_{ij}(s;\widehat{\beta})}{S^{(0)}(s;\widehat{\beta})}G_i.$$

Here $\widehat{\beta}^*$ is the solution to (10). Thus, we may make inference about $\Lambda_0(t)$ via the simulated distribution of $\widehat{V}(t)$.

3 Model Checking

To assess the adequacy of hazards model, we employ some existing goodness-of-fit methods[18]. In the following, we develop a lack-of-fit test for assessing the adequacy of Model (1). Following from Lin et al.[17] and Sun and Su[34], we consider the following cumulative sums of residuals:

$$\mathcal{F}(t,z;\widehat{\beta}) = n^{-1/2}\sum_{i=1}^{n}\int_0^t I(Z_i \le z)d\widehat{M}_i(s;\widehat{\beta}),$$

where $d\widehat{M}_i(s;\beta) = \frac{1}{S_i}\sum\limits_{j=1}^{S_i} d\widehat{M}_{ij}(s;\beta)$ and the event $I(Z_i \le z)$ means that each of the components of Z_i is not larger than the respective component of z. We show in Appendix that the null distribution of $\mathcal{F}(t,z;\widehat{\beta})$ can be approximated by

$$\begin{aligned}\widetilde{\mathcal{F}}(t,z) =& n^{-1/2}\sum_{i=1}^{n}\int_0^t \Big\{I(Z_i \le z) - \frac{S(u,z;\widehat{\beta})}{S^{(0)}(u;\widehat{\beta})}\Big\}d\widehat{M}_i(u;\widehat{\beta}) \\ & - \widehat{B}(t,z)'\widehat{A}^{-1}n^{-1/2}\sum_{i=1}^{n}\int_0^\tau (Z_i^*(u;\widehat{\beta}) - \overline{Z}^*(u;\widehat{\beta}))d\widehat{M}_i(u;\widehat{\beta}),\end{aligned} \tag{11}$$

where

$$\begin{aligned} S(u,z;\beta) =& n^{-1}\sum_{i=1}^{n}\frac{1}{S_i}\sum_{j=1}^{S_i} Y_{ij}(u;\beta_1)\exp\{(\beta_2-\beta_1)'Z_i\}I(Z_i \le z), \\ \widehat{B}_1(t,z) =& n^{-1}\sum_{i=1}^{n}\frac{1}{S_i}\sum_{j=1}^{S_i}\int_0^t Y_{ij}(u;\widehat{\beta}_1)\exp\{(\widehat{\beta}_2-\widehat{\beta}_1)'Z_i\} \\ & \cdot I(Z_i \le z)\big\{Z_i - \overline{Z}(u;\widehat{\beta})\big\}d\{\widehat{\lambda}_0(u)u\}, \\ \widehat{B}_2(t,z) =& n^{-1}\sum_{i=1}^{n}\frac{1}{S_i}\sum_{j=1}^{S_i}\int_0^t Y_{ij}(u;\widehat{\beta}_1)\exp\{(\widehat{\beta}_2-\widehat{\beta}_1)'Z_i\}Z_i \\ & \cdot \Big\{I(Z_i \le z) - \frac{S(u,z;\widehat{\beta})}{S^{(0)}(u;\widehat{\beta})}\Big\}d\widehat{\Lambda}_0(u), \end{aligned}$$

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

It is also difficult to estimate the asymptotic covariance function of $\mathcal{F}(t,z;\widehat{\beta})$ analytically. We again use the resampling approach to approximate the null distribution of $\mathcal{F}(t,z;\widehat{\beta})$ by the conditional distribution of $\widehat{\mathcal{F}}(t,z)$, where

$$\begin{aligned}\widehat{\mathcal{F}}(t,z) =& \big\{\mathcal{F}(t,z;\widehat{\beta}) - \mathcal{F}(t,z;\widehat{\beta}^*)\big\} \\ & + n^{-1/2}\sum_{i=1}^{n}\int_0^t \Big\{I(Z_i \le z) - \frac{S(u,z;\widehat{\beta})}{S^{(0)}(u;\widehat{\beta})}\Big\}d\widehat{M}_i(u;\widehat{\beta})G_i.\end{aligned}$$

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

It can be shown that under model (1), $\mathcal{F}(t,z;\widehat{\beta})$ is equivalent in distribution to $\widetilde{\mathcal{F}}(t;z)$, which converges to a zero-mean Gaussian process. Actually, $\widetilde{\mathcal{F}}(t;z)$ was used here only to show the asymptotic normality of $\mathcal{F}(t,z;\widehat{\beta})$ theoretically. However, it is difficult to estimate the asymptotic covariance function of $\mathcal{F}(t,z;\widehat{\beta})$ or $\widetilde{\mathcal{F}}(t;z)$ analytically. Then a resampling technique based on $\widehat{\mathcal{F}}(t,z)$ was applied and the p-value of the test can be obtained by comparing the observed value of $\sup_{0\le t\le\tau,z}|\mathcal{F}(t,z;\widehat{\beta})|$ to a large number of realizations of $\sup_{0\le t\le\tau,z}|\widehat{\mathcal{F}}(t,z)|$.

4 Simulation Studies

In this section, we conduct some simulation studies to examine the finite-sample behavior of the proposed inference procedure. In the study, a heterogeneous mixture of individual renewal processes is used with Model (1). Specifically, the baseline hazard function is chosen to be log-logistic distributions with scale parameter of 1, that is $\lambda_0(t)=1/(1+t)$. Firstly, we generate the baseline gap time T_{ij}^0 having distribution function $\Phi(A_i+B_{ij})$, where Φ is the cumulative distribution function of the standard normal distribution, A_i and B_{ij} are independent normal random variables with mean zeros and variances ρ and $1-\rho$, respectively, with $\rho\in[0,1]$. Here the parameter ρ dictates the heterogeneity of between individual, and $1-\rho$ controls the heterogeneity of between-episodes within an individual.

Secondly, given the baseline gap times, general gap times T_{ij} are taken as

$$\exp(-\beta_{10}'Z_i)\big((T_{ij}^0)^{\exp\{(\beta_{10}-\beta_{20})'Z_i\}}-1\big),$$

with $\beta_{10}=-1,\ \beta_{20}=1$ and $\beta_{10}=0,\ \beta_{20}=1$ respectively, where Z_i is a Bernoulli random variable with success probability 0.5. The censoring time C_i is taken as the minimum of the uniform distribution on (0, 3) and 1.

For each simulation study, we consider $\rho=0,0.25$ and 0.5. Two choices are considered for the weight function $W(t,Z_i;\beta)$:

(i) $W_1(t,Z_i;\beta)=n^{-1}\sum_{k=1}^{n}\frac{1}{S_k}\sum_{j=1}^{S_k}Y_{kj}(t,\beta_1)\exp(-\beta_1'Z_k)Z_i$;

(ii) $W_2(t,Z_i;\beta)=\frac{t}{1+t}Z_i$.

1000 simulation samples are considered and 1,000 resamplings are generated for each simulation sample.

Table 1 presents the simulation results on estimation of β_{10} and β_{20} with the sample sizes $n=$ 100 and 200, according to two different weight functions. In Table 1, Bias stands for the sample means of the point estimates $\widehat{\beta}$ minus the true value, SSE is the sampling standard errors of $\widehat{\beta}$, ESE is the sampling means of the estimated standard errors of $\widehat{\beta}$, and CP stands for the 95% empirical coverage probability for $\beta_0=(\beta_{10}',\beta_{20}')'$ based on the empirical distribution of $\widehat{\beta}^*$. From Table 1, we easily see the proposed estimation procedure performs well for the situations considered here. Specifically, the proposed estimators are practically unbiased, and both the variance estimation and coverage probabilities seem reasonable. Moreover the results become better when the sample size increases from 100 to 200.

Table 1. Summary of the Simulation Study

			$\beta_{10} = -1$				$\beta_{20} = 1$			
n	ρ	W	BIAS	SSE	ESE	CP	BIAS	SSE	ESE	CP
100	0	W_1	−0.0714	0.1651	0.1664	0.919	0.0082	0.3404	0.3351	0.937
	0.25	W_1	−0.0721	0.1646	0.1693	0.922	0.0091	0.3449	0.3393	0.933
	0.5	W_1	−0.0708	0.1593	0.1796	0.924	0.0081	0.3463	0.3272	0.928
	0	W_2	−0.0725	0.1587	0.1783	0.931	0.0191	0.3344	0.3130	0.928
	0.25	W_2	−0.0763	0.1594	0.1763	0.942	0.0214	0.3357	0.3198	0.932
	0.5	W_2	−0.0802	0.1517	0.1905	0.943	0.0267	0.3457	0.3266	0.925
200	0	W_1	−0.0707	0.1362	0.1483	0.927	0.0076	0.3344	0.3213	0.935
	0.25	W_1	−0.0671	0.1396	0.1497	0.921	0.0072	0.3385	0.3261	0.931
	0.5	W_1	−0.0663	0.1385	0.1499	0.929	0.0073	0.3397	0.3183	0.933
	0	W_2	−0.0687	0.1182	0.1435	0.947	0.0136	0.2304	0.2129	0.925
	0.25	W_2	−0.0665	0.1194	0.1535	0.951	0.0021	0.2285	0.2171	0.936
	0.5	W_2	−0.0662	0.1166	0.1744	0.967	0.0070	0.2272	0.2618	0.963
			$\beta_{10} = 0$				$\beta_{20} = 1$			
100	0	W_1	−0.0013	0.0359	0.0236	0.948	0.0257	0.2685	0.2627	0.943
	0.25	W_1	−0.0009	0.0295	0.0218	0.946	0.0178	0.2642	0.2628	0.948
	0.5	W_1	−0.0006	0.0344	0.0303	0.943	0.0206	0.2721	0.2684	0.946
	0	W_2	−0.0022	0.0324	0.0212	0.951	0.0261	0.2754	0.2625	0.933
	0.25	W_2	−0.0008	0.0375	0.0227	0.944	0.0208	0.2679	0.2622	0.943
	0.5	W_2	−0.0009	0.0332	0.0253	0.943	0.0212	0.2613	0.2583	0.946
200	0	W_1	−0.0007	0.0254	0.0201	0.941	0.0203	0.2217	0.2201	0.939
	0.25	W_1	−0.0005	0.0223	0.0213	0.942	0.0116	0.2050	0.1798	0.944
	0.5	W_1	−0.0006	0.0227	0.0198	0.947	0.0171	0.2017	0.1884	0.945
	0	W_2	−0.0008	0.0219	0.0199	0.948	0.0163	0.2059	0.1736	0.937
	0.25	W_2	−0.0006	0.0228	0.0216	0.945	0.0153	0.2113	0.1822	0.942
	0.5	W_2	−0.0007	0.0235	0.0206	0.938	0.0162	0.2012	0.1906	0.947

5 An Example of Application

In this section, the proposed methodology is applied to a bladder cancer study conducted by the Veterans Administration Cooperative Urological Research Group (see [2,3,37]). The study consisted of 85 patients with superficial bladder tumors, a number of whom experienced recurrences of the tumors. The patients were randomly allocated to one of two treatments, placebo (47) and thiotepa (38). This dataset can be found in the R Package “survival”. A primary interest for this data is to assess the effects of thiotepa treatment on the recurrent event (see [10,33,37]).

In this section, we aim at assessing the effects of the thiotepa on the gap time between recurrent tumors in the general model (1). We set the treatment indicator Z to be 0 for placebo or 1 for thiotepa. To estimate standard errors of $\widehat{\beta}$, 1000 resamplings are used. We adopt the weight function $W_2(t, Z_i; \beta)$ as the simulations. The results are shown in Table 2.

In Table 2, the signs of parameters estimators $(\widehat{\beta}_{10}, \widehat{\beta}_{20}) = (-2.6779, -0.9921)$, are negative, which indicate the treatment thiotepa can delay the recurrence of the bladder tumor. Moreover, according to p-value of β_{10}, we find that thiotepa treatment has an significant effect on the time-scale change. However, the p-value of β_{20} suggests that thiotepa treatment seems to be no significant proportional effect. Finally, we apply the model checking techniques given in Section 3 to assess the adequacy of Model (1) for the data. We calculate the statistic $\mathcal{F}(t, z; \widehat{\beta})$ and obtain p-value of test statistic $\sup_{0 \le t \le \tau, z} |\mathcal{F}(t, z; \widehat{\beta})|$ is 0.3061 based on 1000 realizations of the corresponding statistic $\sup_{0\le t\le\tau,z}|\widehat{F}(t,z)|$. This result indicates that we have no sufficient proof to reject the model assumption and that the model is reasonable to fit the data in some extent. In this example, β_{20} is not significant, we can just consider the accelerated hazards models.

Table 2. Summary of Regression Analysis for Bladder Cancer Data

Parameters	Estimation	Std	p-Value
β_{10}	-2.6779	0.3332	0.0000
β_{20}	-0.9921	0.7096	0.1646

6 Concluding Remarks

In this article we study a general class of semiparametric hazards regression models for recurrent gap time data. The models are flexible and include some commonly used models as special cases. An estimation procedure is proposed for the model parameters, based on an established connection between observed gap times and clustered survival data with informative cluster size. The asymptotic properties of the estimators are established. Simulation studies are conducted to verify the finite sample behaviors and the results show that the proposed method works well.

Note that in (8), there is a weight function that needs to be specified. We use two different weight functions W_1 and W_2 in the simulation studies, and there are many potential candidates for the weight function[6]. Usually the weight function was used to improve the efficiency of the estimation since the estimate based on the generalized estimating equations is not efficient. It is clear that W cannot be any data-dependent function as W needs to be chosen such that the estimating function has zero expectation. In general, one would like to choose W that gives the most efficient estimate of covariate effects. However, sometimes, this may not be possible and it is usually quite difficult. Chen and Jewell[6] gave some suggestions in their simulation part. It may be a valuable research direction in the future to develop some procedures for the selection of an appropriate weight function for a given data set.

7 Appendix: the Proofs of Asymptotic Properties

Now, we give the similar regularity conditions defined by Anderson and Gill[1] in the following:

(C1) $(N_i(\cdot), C_i, Z_i(\cdot))$ are independent and identically distributed for $i = 1, \cdots, n$.

(C2) $P(Y_i(\tau;\beta_{10}) = 1) > 0$.

(C3) Z_i and $W(t, Z_i; \beta_0)$ are bounded on $[0,\tau]$ for $i = 1, \cdots, n$.

(C4) $\lambda_0(\cdot)$ is bounded twice continuously differentiable.

(C5) $C_i e^{\beta_{10}Z_i}$ has a bounded density.

(C6) A is nonsingular, where

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

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

$$A_{12} = \int_0^\tau [s_z^{(2)}(t) - \overline{z}(t) s^{(1)}(t)'] d\{\Lambda_0(t)\},$$

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

$$A_{22} = \int_0^\tau [s_w^{(2)}(t) - \overline{w}(t)s^{(1)}(t)']d\{\Lambda_0(t)\}.$$

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

Proof of Theorem 1. Clearly, an easy algebraic manipulation yields that

$$U(\beta_0) = \sum_{i=1}^n \frac{1}{S_i}\sum_{j=1}^{S_i}\int_0^\tau \{Z_i^*(t;\beta_0) - \overline{Z}^*(t;\beta_0)\}dM_{ij}(t;\beta_{10}).$$

By Theorem 1 of Lin et al.[15], it's easy to obtain that

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

where

$$d_i = \frac{1}{S_i}\sum_{j=1}^{S_i}\int_0^\tau \{Z_i^*(u,\beta_0) - \overline{z}^*(u)\}dM_{ij}(u;\beta_0).$$

It follows from the multivariate central limit theorem that $n^{-1/2}U(\beta_0)$ 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 $\widehat{\Sigma}$.

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 $\widehat{\beta}$. □

Proof of Theorem 2. We write

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

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

$$U_1(\beta_1,\beta_2) - U_1(\beta_1,\beta_{20}) = -A_{12}n(\beta_2 - \beta_{20}) + o(n\|\beta-\beta_0\|).$$

On the other hand, we note that

$$\begin{aligned} U_1(\beta_1,\beta_{20}) - U_1(\beta_{10},\beta_{20}) =& \sum_{i=1}^n \frac{1}{S_i}\sum_{j=1}^{S_i}\int_0^\tau \big(Z_i(t;\beta_1,\beta_{20}) - \overline{Z}(t;\beta_1,\beta_{20})\big) \\ &\times \big\{d\widetilde{N}_{ij}(t,\beta_1) - Y_{ij}(t,\beta_1)\exp\{(\beta_{20}-\beta_1)'Z_i\}\big\}d\Lambda_0(te^{(\beta_{10}-\beta_1)'Z_i}) \\ &- \sum_{i=1}^n \frac{1}{S_i}\sum_{j=1}^{S_i}\int_0^\tau \big(Z_i(t;\beta_{10},\beta_{20}) - \overline{Z}(t;\beta_{10},\beta_{20})\big) \\ &\times \big\{d\widetilde{N}_{ij}(t,\beta_{10}) - Y_{ij}(t,\beta_{10})\exp\{(\beta_{20}-\beta_{10})'Z_i\}\big\}d\Lambda_0(t) \\ &+ \sum_{i=1}^n \frac{1}{S_i}\sum_{j=1}^{S_i}\int_0^\tau (Z_i(t;\beta_1,\beta_{20}) - \overline{Z}(t;\beta_1,\beta_{20}) \\ &\times Y_{ij}(t,\beta_1)e^{(\beta_{20}-\beta_1)'Z_i}d[\Lambda_0(te^{(\beta_{10}-\beta_1)'Z_i}) - \Lambda_0(t)]. \end{aligned} \tag{12}$$

Applying the technique of Ying[38] and Lin et al.[15], one shows that the first term on the right hand side of equality (12) is of order $o(n^{1/2})$. By a Taylor series expansion, we obtain that

$$\Lambda_0(te^{(\beta_{10}-\beta_1)'Z_i}) - \Lambda_0(t) = \{\lambda_0(t) + o(1)\}t(\beta_{10}-\beta_1)'Z_i.$$

Therefore, the second term can be rewritten as

$$\sum_{i=1}^{n}\frac{1}{S_i}\sum_{j=1}^{S_i}\int_0^t[Z_i(t;\beta_1,\beta_{20})-\overline{Z}^{(}t;\beta_1,\beta_{20})]\cdot Y_{ij}(t,\beta_1)e^{(\beta_{20}-\beta_1)'Z_i}Z_i'd\{t\lambda_0(t)\}(\beta_{10}-\beta_1)$$
$$+o(n\|\beta_1-\beta_{10}\|)=-A_{11}n(\beta_1-\beta_{10})+o(n\|\beta_1-\beta_{10}\|)$$

almost surely. It then follows from analogy with Theorem 1 in [38] that, for any sequence $\varepsilon_n\to 0$,

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

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

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

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

$$\sup_{\|\beta-\beta_0\|\le\varepsilon_n}\left\{\|U(\beta)-U(\beta_0)+A\cdot n(\beta-\beta_0)\|/(n^{1/2}+n\|\beta-\beta_0\|)\right\}=o(1)\tag{13}$$

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.[15], we can get that $\widehat{\beta}$ is strongly consistent under the regularity Conditions (C1)–(C5). In addition, by the definition of $\widehat{\beta}$, condition (C6) and (13), $n^{1/2}(\widehat{\beta}-\beta_0)$ is asymptotically normal with mean zero and covariance matrix $A^{-1}\Sigma(A^{-1})'$, which can be consistently estimated by $\widehat{A}^{-1}\widehat{\Sigma}(\widehat{A}^{-1})'$. □

Proof of Theorem 3. Note that

$$\begin{aligned}\widehat{\Lambda}_0(t)-\Lambda_0(t)=&\widehat{\Lambda}_0(t;\widehat{\beta}_1,\widehat{\beta}_2)-\widehat{\Lambda}_0(t;\widehat{\beta}_1,\beta_{20})+\widehat{\Lambda}_0(t;\widehat{\beta}_1,\beta_{20})\\&-\widehat{\Lambda}_0(t;\beta_0)+\widehat{\Lambda}_0(t;\beta_0)-\Lambda_0(t).\end{aligned}$$

With a Taylor series expansion, the uniform strong law of large numbers[26] and an easy algebraic manipulation, we have

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

uniformly in $t\in[0,\tau]$. Applying the results of Theorem 2, we obtain

$$n^{1/2}(\widehat{\Lambda}_0(t;\widehat{\beta}_1,\beta_{20})-\widehat{\Lambda}_0(t;\beta_0))=-h_1(t)'(\widehat{\beta}_1-\beta_{10})+o_p(1),$$

uniformly in $t\in[0,\tau]$, where $h_1(t)=\int_0^t\overline{z}(u)d\{\lambda_0(u)u\}$ and $h_2(t)=\int_0^t\frac{s^{(1)}(u)}{s^{(0)}(u)}d\{\Lambda_0(u)\}$. Clearly, some algebraic manipulations yield

$$n^{1/2}\{\widehat{\Lambda}_0(t,\beta_0)-\Lambda_0(t)\}=n^{-1/2}\sum_{i=1}^{n}\frac{1}{S_i}\sum_{j=1}^{S_i}\int_0^t\frac{dM_{ij}(u;\beta_0)}{s^{(0)}(u)}+o_p(1)$$

uniformly in $t\in[0,\tau]$. Hence, we have

$$n^{1/2}\{\widehat{\Lambda}_0(t)-\Lambda_0(t)\}=n^{-1/2}\sum_{i=1}^{n}\Psi_i(t)+o_p(1),\tag{14}$$

where

$$\Psi_i(t) = \frac{1}{S_i}\sum_{j=1}^{S_i}\int_0^t \frac{dM_{ij}(u;\beta_0)}{s^{(0)}(u)} - h(t)'A^{-1}\frac{1}{S_i}\sum_{j=1}^{S_i}\int_0^\tau \{Z_i^*(u,\beta_0) - \overline{z}^*(u)\}dM_{ij}(u;\beta_0)$$

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

Because $\Psi_i(t)$ are independent zero-mean random variables for each t, via the multivariate central limit theorem, we know $n^{-1/2}\sum_{i=1}^n \Psi_i(t)$ converges in finite dimensional distributions to a zero-mean Gaussian process. By the modern empirical theory as in Lin et al.[17], we know $n^{-1/2}\sum_{i=1}^n \Psi_i(t)$ is tight. Thus, $n^{1/2}(\widehat{\Lambda}_0(t) - \Lambda_0(t))$ converges weakly to a zero-mean Gaussian process with covariance function $\Gamma(s,t)$ that can be consistently estimated by $\widehat{\Gamma}(s,t)$ at (s,t), by the arguments of Lin et al.[17]. □

Proof of (11) in Section 3. Rewrite

$$\begin{aligned}
&\mathcal{F}(t,z;\widehat{\beta})\\
=&n^{-1/2}\sum_{i=1}^n\frac{1}{S_i}\sum_{j=1}^{S_i}\int_0^t I(Z_i\le z)\big\{d\widetilde{N}_{ij}(u;\widehat{\beta}_1) - Y_{ij}(u;\widehat{\beta}_1)e^{(\beta_{20}-\widehat{\beta}_1)'Z_i}d\Lambda_0(ue^{(\beta_{10}-\widehat{\beta}_1)'Z_i})\big\}\\
&-n^{-1/2}\sum_{i=1}^n\frac{1}{S_i}\sum_{j=1}^{S_i}\int_0^t I(Z_i\le z)\big\{d\widetilde{N}_{ij}(u;\beta_{10}) - Y_{ij}(u;\beta_{10})e^{(\beta_{20}-\beta_{10})'Z_i}d\Lambda_0(u)\big\}\\
&-n^{-1/2}\sum_{i=1}^n\frac{1}{S_i}\sum_{j=1}^{S_i}\int_0^t I(Z_i\le z)Y_{ij}(u;\widehat{\beta}_1)e^{(\widehat{\beta}_2-\widehat{\beta}_1)'Z_i}d\big[\widehat{\Lambda}_0(u,\widehat{\beta}) - \Lambda_0(u)\big]\\
&-n^{-1/2}\sum_{i=1}^n\frac{1}{S_i}\sum_{j=1}^{S_i}\int_0^t I(Z_i\le z)Y_{ij}(u;\widehat{\beta}_1)\big[e^{(\widehat{\beta}_2-\widehat{\beta}_1)'Z_i}d\Lambda_0(u) - e^{(\beta_{20}-\widehat{\beta}_1)'Z_i}d\Lambda_0(ue^{(\beta_{10}-\widehat{\beta}_1)'Z_i})\big]\\
&+n^{-1/2}\sum_{i=1}^n\frac{1}{S_i}\sum_{j=1}^{S_i}\int_0^t I(Z_i\le z)dM_{ij}(u;\beta_0).
\end{aligned} \tag{15}$$

Applying the technique of Ying[38] and Lin et al.[15], we can show that the first and the second term on the right-hand side of equality (15) is of order o(1) uniformly in t and z. Similar to (14), the third term on the right-hand side of equality (15) is equivalent to

$$-n^{-1/2}\sum_{i=1}^n\frac{1}{S_i}\sum_{j=1}^{S_i}\int_0^t\frac{s(u,z)}{s^{(0)}(u)}dM_{ij}(u;\beta_0) + \widetilde{b}(t,z)'n^{1/2}(\widehat{\beta}-\beta_0) + o_p(1),$$

where

$$\widetilde{b}_1(t,z) = n^{-1}\sum_{i=1}^n\frac{1}{S_i}\sum_{j=1}^{S_i}\int_0^t Y_{ij}(u;\beta_{10})e^{(\beta_{20}-\beta_{10})'Z_i}I(Z_i\le z)\overline{z}(u)d\{\lambda_0(u)u\},$$

$$\widetilde{b}_2(t,z) = n^{-1}\sum_{i=1}^n\frac{1}{S_i}\sum_{j=1}^{S_i}\int_0^t Y_{ij}(u;\beta_{10})e^{(\beta_{20}-\beta_{10})'Z_i}I(Z_i\le z)\frac{s^{(1)}(u)}{s^{(0)}(u)}d\{\Lambda_0(u)\}$$

and $\widetilde{b}(t,z) = (\widetilde{b}_1(t,z)', \widetilde{b}_2(t,z)')'$. By a Taylor series expansion, we obtain the fourth term on the right-hand side of (15) equals

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

where

$$b_1^*(t,z) = n^{-1}\sum_{i=1}^{n}\frac{1}{S_i}\sum_{j=1}^{S_i}\int_0^t Y_{ij}(u;\beta_{10})e^{(\beta_{20}-\beta_{10})'Z_i}I(Z_i \le z)Z_i d\{\lambda_0(u)u\},$$

$$b_2^*(t,z) = n^{-1}\sum_{i=1}^{n}\frac{1}{S_i}\sum_{j=1}^{S_i}\int_0^t Y_{ij}(u;\beta_{10})e^{(\beta_{20}-\beta_{10})'Z_i}I(Z_i \le z)Z_i d\{\Lambda_0(u)\}$$

and $b^*(t,z) = (b_1^*(t,z)', b_2^*(t,z)')'$. Set $b(t,z) = \widetilde{b}(t,z) - b^*(t,z)$. Therefore, we have

$$\begin{aligned}\mathcal{F}(t,z;\widehat{\beta}) =& n^{-1/2}\sum_{i=1}^{n}\frac{1}{S_i}\sum_{j=1}^{S_i}\int_0^t \Big\{I(Z_i \le z) - \frac{s(u,z)}{s^{(0)}(u)}\Big\}dM_{ij}(u;\beta_0)\\ &- b(t,z)'A^{-1}n^{-1/2}\sum_{i=1}^{n}\frac{1}{S_i}\sum_{j=1}^{S_i}\int_0^\tau \{Z_i^*(u;\beta_0) - \overline{z}^*\}dM_{ij}(u;\beta_0)\end{aligned}$$

uniformly in t and z. Obviously, it is a sum of i.i.d. zero-mean terms for fixed t and z. By the multivariate central limit theorem, we all know $\mathcal{F}(t,z;\widehat{\beta})$ converges in finite dimensional distributions to a zero-mean Gaussian process. Using the modern empirical theory as in Lin et al.[17], we can show that $\mathcal{F}(t,z;\widehat{\beta})$ is tight. By the arguments of Lin et al.[17], this Gaussian process can be approximated by $\widetilde{\mathcal{F}}(t,z)$ given by (11). □

References

[1] Andersen, P.K., Gill, R.D. Coxs regression model for counting processes: a large sample study. *Ann. Statist.*, 10: 1100–1120 (1982)

[2] Andrews, D.F., Herzberg, A.M. Data; A Collection of Problems from Many Fields for the Student and Research Worker. Springer-Verlag, New York, 1985

[3] Byar, D. P. The veterans administration study of chemoprophylaxis for recurrent stage I bladder tumors: comparison of placebo, pyridoxine, and topical thiotepa. In: Bladder Tumors and Other Topics in Urological Oncology (Edited by M. Pavone-Macaluso, P. H. Smith, and F. Edsmyn), Plenum, New York, 1980, 363–370

[4] Chang, S.H., Wang, M.C. Conditional regression analysis for recurrence time data. *J. Amer. Statist. Assoc.*, 94, 1221–1230 (1999)

[5] Chang, S.H. Estimating marginal effects in accelerated failure time models for serial sojourn among repeated event. *Lifetime Data Anal.*, 10: 175–190 (2004)

[6] Chen, Y.Q., Jewell, N.P. On a general class of semiparametric hazards regression models. *Biometrika*, 88: 687–702 (2001)

[7] Cook, R.J., Lawless, J.F. The Statistical Analysis of Recurrent Events. Springer, New York, 2007

[8] Darlington, G.A., Dixon, S.N. Event-weighted proportional hazards modelling for recurrent gap time data. *Stat. Med.*, 32, 124–130 (2013)

[9] Du, P. Nonparametric modeling of the gap time in recurrent event data. *Lifetime Data Anal.*, 15: 256–277 (2009)

[10] Huang, Y., Chen, Y.Q. Marginal regression of gaps between recurrent events. *Lifetime Data Anal.*, 9: 293–303 (2003)

[11] Huang, X., Liu, L. A joint frailty model for survival and gap times between recurrent events. *Biometrics*, 63: 389–397 (2007)

[12] Kang, F.Y., Sun, L.Q., Zhao, X.Q. A class of transformed hazards models for recurrent gap times. *Computational Statistics and Data Analysis*, 83: 151–167 (2015)

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

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

[15] Lin, D.Y., Wei, L.J., Ying, Z. Accelerated failure time models for counting processes. *Biometrika*, 85: 609–618 (1998)

[16] Lin, D.Y., Sun, W., Ying, Z. Nonparametric estimation of the gap time distribution for serial events with censored data. *Biometrika*, 86: 59–70 (1999)

[17] Lin, D.Y., Wei, L.J., Yang, L., Ying, Z. Semiparametric regression for the mean and rate function of recurrent events. *J.R. Stat. Soc. Ser. B*, 69: 711–730 (2000)

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

[19] Lu, W. Marginal regression of multivariate event times based on linear transformation models. *Lifetime Data Anal.*, 11: 389–404 (2005)

[20] Luo, X., Huang, C.Y. Analysis of recurrent gap time data using the weighted risk-set method and the modified within-cluster resampling method. *Stat. Med.*, 30: 301–311 (2011)

[21] Luo, X., Huang, C. Y., Wang, L. Quantile regression for recurrent gap time data. *Biometrics*, 69(2): 375–385 (2013)

[22] Murphy, S.A. Asymptotic theory for the frailty model. *Ann. Statist.*, 23: 182–198 (1995)

[23] Nielsen, G.G., Gill, R.D., Andersen, P.K., Sorensen, T.I.A. A counting process approach to maximum likelihood estimation in frailty models. *Scand. J. Statist.*, 19: 25–43 (1992)

[24] Parzen, M.I., Wei, L.J., Ying, Z. A resampling method based on pivotal estimating functions. *Biometrika*, 81: 341–350 (1994)

[25] Peña, E.A., Strawderman, R., Hollander, M. Nonparametric estimation with recurrent event data. *J. Amer. Statist. Assoc.*, 96: 1299–1315 (2001)

[26] Pollard, D. Empirical Processes: Theory and Applications. Institute of Mathematical Statistics, Hayward, CA, 1990

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

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

[29] Schaubel, D.E., Cai, J. Regression analysis for gap time hazard functions of sequentially ordered multiplivariate failure time data. *Biometrika*, 91: 291–303 (2004)

[30] Schaubel, D.E., Zeng, D., Cai, J. A semiparametric additive rates model for recurrent event data. *Lifetime Data Anal.*, 12: 389–406 (2006)

[31] Strawderman, R.L. The accelerated gap times model. *Biometrika*, 92: 1299–1315 (2005)

[32] Stute, W. The oscillation behavior of empirical processes. *Ann. Statist.*, 10: 86–107 (1982)

[33] Sun, L., Park, D.H., Sun, J. The additive hazards model for recurrent gap times. *Statist. Sinica*, 16: 919–932 (2006)

[34] Sun, L., Su, B. A class of accelerated means regression models for recurrent event data. *Lifetime Data Anal.*, 14: 357–375 (2008)

[35] Sun, L., Tong, X., Zhou, X. A class of Box-Cox transformation models for recurrent event data. *Lifetime Data Anal.*, 17: 280–301 (2011)

[36] Wang, M.C., Chang, S.H. Nonparametric estimation of a recurrent survival function. *J. Amer. Statist. Assoc.*, 94: 146–153 (1999)

[37] Wei, L.J., Lin, D.Y., Weissfeld, L. Regression analysis of multivariate incomplete failure time data by modeling marginal distributions. *J. Amer. Statist. Assoc.*, 84: 1065–1073 (1989)

[38] Ying, Z. A large sample study of rank estimation for censored regression data. *The Annals of Statistics*, 21: 76-99 (1993)

[39] Zeng, D., Lin, D.Y. Efficient estimation of semiparametric transformation models with random effects for recurrent events. *Biometrika*, 93: 627–640 (2006)

[40] Zeng, D., Lin, D.Y. Semiparametric transformation models for counting processes. *J. Amer. Statist. Assoc.*, 102: 167–180 (2007)