Research Article

Statistics in Medicine

Received XXXX

(www.interscience.wiley.com) DOI: 10.1002/sim.0000

Induced smoothing for rank-based regression with recurrent gap time data

Tianmeng Lyu[a], Xianghua Luo[ab*†], Gongjun Xu[c], Chiung-Yu Huang[d]

Various semiparametric regression models have recently been proposed for the analysis of gap times between consecutive recurrent events. Among them, the semiparametric accelerated failure time (AFT) model is especially appealing owing to its direct interpretation of covariate effects on the gap times. In general, estimation of the semiparametric AFT model is challenging because the rank-based estimating function is a non-smooth step function. As a result, solutions to the estimating equations do not necessarily exist. Moreover, the popular resampling-based variance estimation for the AFT model requires solving rank-based estimating equations repeatedly and hence can be computationally cumbersome and unstable. In this paper, we extend the induced smoothing approach to the AFT model for recurrent gap time data. Our proposed smooth estimating function permits the application of standard numerical methods for both the regression coefficients estimation and the standard error estimation. Large-sample properties and an asymptotic variance estimator are provided for the proposed method. Simulation studies show that the proposed method outperforms the existing non-smooth rank-based estimating function methods in both point estimation and variance estimation. The proposed method is applied to the data analysis of repeated hospitalizations for patients in the Danish Psychiatric Center Register.

Keywords: accelerated failure time model; gap times; Gehan-type weight; induced smoothing; recurrent events

1. Introduction

Recurrent event data are frequently encountered in clinical and epidemiological studies, where each subject can experience an event of interest repeatedly. Examples of recurrent events include rehospitalizations experienced by patients with psychiatric disorders [1], recurrent infections after hematopoietic cell transplantations [2], and many others. Depending on the nature of recurrent events and the research interest, the focus of statistical analysis can be placed on the time-to-event data by modeling the intensity or rate function of the counting process or on the gap times between consecutive events. For the former, various nonparametric and semiparametric methods have been developed in the literature. Some ponparametric

[a] Division of Biostatistics, School of Public Health, University of Minnesota, Minneapolis, MN, U.S.A.
[b] Biostatistics Core, Masonic Cancer Center, University of Minnesota, Minneapolis, MN, U.S.A.
[c] Department of Statistics, University of Michigan, Ann Arbor, MI, U.S.A.
[d] Division of Biostatistics and Bioinformatics, Sidney Kimmel Comprehensive Cancer Center, Johns Hopkins University, Baltimore, MD, U.S.A.
* Correspondence to: X. Luo, Division of Biostatistics, MMC303, School of Public Health, University of Minnesota, Minneapolis, MN 55455, U.S.A.
† Email: luox0054@umn.edu

methods include the estimation of the cumulative rate function [3, 4] and techniques for estimating the rate function [5]. Several authors [6, 7, 8, 9] considered Cox-type models which assume that the effects of covariates are multiplicative on the intensity or rate functions of the underlying counting process, whereas others considered additive intensity or rate models [10, 11].

Alternatively, the focus can be placed on the gap times between recurrent events. As discussed in [12], the unique sequential ordering structure of recurrent gap time data generates difficulty in model estimation. First, due to the correlation among gap times of the same subject, the recurrent gap times beyond the first gap are subject to induced informative censoring even when the total censoring time is completely random. Second, the last censored gap time is expected to be longer than the previous uncensored gap times. Lastly, unlike the clustered survival data where the cluster size is typically assumed to be non-informative, the number of recurrent gap times of a subject is usually informative since subjects who are at a higher risk tend to have more gap times. Therefore, it is not appropriate to naively treat recurrent gap time data as independently censored clustered survival data and apply methods for clustered survival data to recurrent gap time data. Several authors [12, 13] have developed nonparametric methods to estimate the distribution of recurrent gap times, while others [14, 15] studied nonparametric estimation of the gap time hazard function in the presence of covariates. Semiparametric regression models for recurrent gap time data include proportional hazards (PH) models [16], accelerated failure time (AFT) models [17, 18], linear transformation models [19], additive hazards models [20], and more recently, quantile regression models [21] and transformed hazards models [22].

Among the various recurrent gap time models, the AFT model is particularly appealing as it provides a direct interpretation of the covariate effects on the (transformed) length of gap times. Nevertheless, similar to the AFT models for univariate survival data [23, 24, 25, and reference therein], the estimation of the AFT model for recurrent gap time data [17] usually relies on rank-based estimating functions which are non-smooth step functions of regression parameters. It is well known that solving non-smooth, rank-based estimating equations could be computationally challenging since the solution to a non-smooth estimating equation typically does not exist. In addition to the difficulties in point estimation, variance estimation for the semiparametric AFT models has also been found challenging. This is because the asymptotic variance depends on the slope of the estimating function which can not be evaluated directly when the estimating function is non-smooth. Popular alternatives for variance estimation include the bootstrap method [26] and the perturbation method [27, 17]. However, both methods require solving rank-based estimating equations for numerous times, and hence can be computationally inefficient and unstable since they depend heavily on the point estimation from the non-smooth estimating functions, which is not guaranteed to succeed, for each resampling.

To tackle the difficulties in variance estimation for the AFT models with univariate survival data, Zeng and Lin [28] proposed new resampling methods which only require evaluating the estimating functions repeatedly rather than solving them. These methods [28] can greatly improve the efficiency in computing for the variance estimation; however, the challenge in the point estimation remains unresolved. Alternatively, efforts have been made on improving the point and variance estimation simultaneously by approximating the rank-based estimation function by a continuously differentiable estimating function so that the standard numerical methods can be applied in the inference procedure. In particular, Brown and Wang [29] proposed the so-called induced smoothing technique for the rank-based estimating function for univariate survival data with Gehan's weight. Later, it was extended to general weights [30]. Similar smoothing techniques have been extended to clustered survival data [31, 32]. To our knowledge, no efforts have been made on improving the estimation of the AFT model with recurrent gap time data in literature. In this paper we propose to extend the induced smoothing technique to the AFT model for recurrent gap time data.

The rest of the paper is organized as follows. In Section 2, we first introduce the notation and setting of the AFT model for recurrent gap time data. We then briefly introduce the non-smooth rank-based estimating functions. In Section 3, we present the proposed induced smoothing method for the recurrent-gap-time AFT model followed by its large-sample properties and an asymptotic variance estimator. In Section 4, we conduct simulation studies to compare the proposed induced smoothing method with the existing rank-based estimating function method with various variance estimation methods. A real data analysis using the patient contact data from the Danish Psychiatric Central Register is presented in Section 5. Some concluding remarks are provided in Section 6.

2. The AFT model and rank-based estimating functions

2.1. *The AFT model for recurrent gap time data*

Consider a study with n subjects being recruited after each experienced an initial event and being followed on the recurrence of the event. Let $i = 1, \ldots, n$ index the subjects and $j = 0, 1, \ldots$ index the recurrent events of the ith subject, with $j = 0$ indicating the initial event. Let T_{ij} denote the gap time between the $(j-1)$th event and the jth event for subject i. Among the various regression models for recurrent gap times, the AFT model is of particular interest because of its direct interpretation of covariate effects on the (transformed) gap time variable. Let $\mathbf{Z}_i$ be the $p \times 1$ vector of baseline covariates. We impose the usual linear model for the logarithm-transformed gap times:

$$\log(T_{ij}) = \boldsymbol{\beta}_0^{\intercal}\mathbf{Z}_i + \epsilon_{ij}, \tag{1}$$

where $\boldsymbol{\beta}_0$ is the true $p \times 1$ vector of regression parameters and has the usual interpretation of covariate effects as in linear models. The error terms within each subject, $\epsilon_{ij}, j = 1, 2, \ldots$, are assumed to have an unknown common marginal distribution, and the correlation structure among the error terms is left unspecified. In this way, the correlation between two gap times ϵ_{ij} and $\epsilon_{ij'}$ is allowed to depend on j and j'. Finally, we assume that the error vectors $\epsilon_i = (\epsilon_{i1}, \epsilon_{i2}, \ldots)^{\intercal}$, $i = 1, \ldots, n$, are independently and identically distributed (i.i.d.) across subjects.

Note that the identical marginal distribution condition assumed for Model (1) is weaker than the shared frailty model which assumes that the error terms of the same subject are i.i.d. given a subject-specific frailty variable. Under the shared frailty model, each pair of gap times in the set $\{\log(T_{ij}), j = 1, \ldots\}$ are required to have the same correlation. The identical marginal distribution condition for Model (1) leaves the within-subject correlation structure fully unspecified, hence Model (1) allows more sophisticated correlation structure in real data, such as the autoregressive (AR) and the unstructured correlation.

In most applications, the observation of recurrent events is subject to right censoring due to loss of follow-up or end of study. Let C_i be the censoring time of the recurrent event process for the ith subject, which is assumed to be independent of $\{T_{ij}; j \geq 1\}$ conditional on $\mathbf{Z}_i$. Let m_i denote the number of observed events so that m_i satisfies $\sum_{j=1}^{m_i} T_{ij} \leq C_i$ and $\sum_{j=1}^{m_i+1} T_{ij} > C_i$, where $\sum_1^0 = 0$. We further define the censoring indicator for the jth event $\delta_{ij} = \mathrm{I}(\sum_{l=1}^{j} T_{il} \leq C_i)$, where $\mathrm{I}(\cdot)$ is an indicator function. Let X_{ij} denote the observed gap time such that $X_{ij} = T_{ij}$ for $j = 1, \ldots, m_i$ and $X_{i,m_i+1} = C_i - \sum_{l=1}^{m_i} X_{il}$. Define the transformed observed gap time $Y_{ij} = \log(X_{ij})$. The observed data of subject i consist of $\{(X_{ij}, \delta_{ij}); j = 1, \ldots, m_i + 1, \mathbf{Z}_i, C_i\}$.

2.2. *Rank-based estimating function*

We begin by considering the simple yet inefficient method that only uses times to first event in model estimation; that is, ignoring gap times of higher orders. Define the residuals $e_{ij}(\boldsymbol{\beta}) = \log(X_{ij}) - \boldsymbol{\beta}^{\intercal}\mathbf{Z}_i$. Let $N_{ij}(\boldsymbol{\beta}, t) = \delta_{ij} I\{e_{ij}(\boldsymbol{\beta}) \leq t\}$ and $R_{ij}(\boldsymbol{\beta}, t) = I\{e_{ij}(\boldsymbol{\beta}) \geq t\}$ be the counting process and at-risk process on the time scale of the residual, corresponding to subject i's jth gap time. An unbiased weighted rank-based estimating function for $\boldsymbol{\beta}$ based on the time-to-first event data takes the form [24, 34, 35]:

$$\sum_{i=1}^{n} w(\boldsymbol{\beta}, e_{i1}(\boldsymbol{\beta}))\delta_{i1}\left[\mathbf{Z}_i - \frac{\frac{1}{n}\sum_{l=1}^{n}\mathbf{Z}_l \mathrm{I}\{e_{l1}(\boldsymbol{\beta}) \geq e_{i1}(\boldsymbol{\beta})\}}{\frac{1}{n}\sum_{l=1}^{n}\mathrm{I}\{e_{l1}(\boldsymbol{\beta}) \geq e_{i1}(\boldsymbol{\beta})\}}\right]$$

or, equivalently,

$$\sum_{i=1}^{n} \int_{-\infty}^{\infty} w(\boldsymbol{\beta}, t) \left\{ \mathbf{Z}_i - \frac{S_1(\boldsymbol{\beta}, t)}{S_0(\boldsymbol{\beta}, t)} \right\} dN_{i1}(\boldsymbol{\beta}, t), \tag{2}$$

where $S_0(\boldsymbol{\beta}, t) = n^{-1} \sum_{i=1}^{n} R_{i1}(\boldsymbol{\beta}, t)$, $S_1(\boldsymbol{\beta}, t) = n^{-1} \sum_{i=1}^{n} \mathbf{Z}_i R_{i1}(\boldsymbol{\beta}, t)$, and $w(\boldsymbol{\beta}, t)$ is the weight function. Common choices of $w(\boldsymbol{\beta}, t)$ include $w(\boldsymbol{\beta}, t) \equiv 1$ for log-rank (LR) weight [23] and $w(\boldsymbol{\beta}, t) \equiv S_0(\boldsymbol{\beta}, t)$ for Gehan's weight [36]. Note that the estimating function in (2) is constructed based on the linear rank statistic and can be viewed as the sum of the weighted difference between the covariate of a subject with an event (subject i) and the expected covariate among those who are in the "risk set" at the transformed event time of this subject, $\{l : e_{l1}(\boldsymbol{\beta}) \geq e_{i1}(\boldsymbol{\beta})\}$.

To improve the efficiency of estimation, one can make use of information beyond the first gap time. However, as discussed earlier, methods for clustered survival data cannot be directly applied to the recurrent gap time data due to the unique sequential structure of recurrent events. It was demonstrated in [37] that, when the underlying recurrent gap times of a subject are exchangeable, the weighted-risk set (WRS) technique can be applied to a reduced dataset to avoid biases in estimation caused by induced informative censoring and the biased sampling of the last censored gap time. Specifically the last censored gap time is not used in the construction of the estimating functions if the number of uncensored gap times of a subject is at least one. For the ease of discussion, we define $m_i^* = \max\{m_i, 1\}$, then $m_i^* = 1$ if subject i has no observed recurrent events and m_i^* equals the number of observed recurrent events m_i if $m_i \geq 1$. Note that $X_{i1} = C_i$ if $m_i = 0$ and $X_{ij} = T_{ij}$ for $j = 1, .., m_i^*$ if $m_i \geq 1$. Thus, the reduced data used in the WRS estimations are $\{(X_{ij}, \delta_{ij}); j = 1, \ldots, m_i^*, \mathbf{Z}_i, C_i\}$ from each subject. The WRS method assigns a weight $1/m_i^*$ to each of the remaining m_i^* gap times of a subject to ensure that overall contribution of each subject to the estimation to be the same to avoid the possible bias caused by informative cluster sizes.

In the same spirit as the WRS method in [37], we first define the *averaged counting process* and the *averaged at-risk process* for the AFT model:

$$\begin{aligned} N_i^*(\boldsymbol{\beta}, t) &= \frac{1}{m_i^*} \sum_{j=1}^{m_i^*} N_{ij}(\boldsymbol{\beta}, t), \\ R_i^*(\boldsymbol{\beta}, t) &= \frac{1}{m_i^*} \sum_{j=1}^{m_i^*} R_{ij}(\boldsymbol{\beta}, t). \end{aligned}$$

Note that these two averaged processes are based on the individual counting processes N_{ij} and R_{ij} defined earlier, which are all on the scale of the residual of the log-transformed gap times. Hence, the two averaged processes $N_i^*(\boldsymbol{\beta}, t)$ and $R_i^*(\boldsymbol{\beta}, t)$ defined here are different than those in [37]. Let $S_0^*(\boldsymbol{\beta}, t) = n^{-1} \sum_{i=1}^{n} R_i^*(\boldsymbol{\beta}, t)$ and $S_1^*(\boldsymbol{\beta}, t) = n^{-1} \sum_{i=1}^{n} \mathbf{Z}_i R_i^*(\boldsymbol{\beta}, t)$. Then, we can replace $n^{-1} \sum_{i=1}^{n} N_{i1}(\boldsymbol{\beta}, t)$, $n^{-1} \sum_{i=1}^{n} \mathbf{Z}_i N_{i1}(\boldsymbol{\beta}, t)$, $S_0(\boldsymbol{\beta}, t)$, and $S_1(\boldsymbol{\beta}, t)$ in (2) with their respective multivariate counterparts $n^{-1} \sum_{i=1}^{n} N_i^*(\boldsymbol{\beta}, t)$, $n^{-1} \sum_{i=1}^{n} \mathbf{Z}_i N_i^*(\boldsymbol{\beta}, t)$, $S_0^*(\boldsymbol{\beta}, t)$, and $S_1^*(\boldsymbol{\beta}, t)$ and construct a new estimating equation:

$$U(\boldsymbol{\beta}) = \sum_{i=1}^{n} \int_{-\infty}^{\infty} w^*(\boldsymbol{\beta}, t) \left\{ \mathbf{Z}_i - \frac{S_1^*(\boldsymbol{\beta}, t)}{S_0^*(\boldsymbol{\beta}, t)} \right\} dN_i^*(\boldsymbol{\beta}, t), \tag{3}$$

where the weight function $w^*(\boldsymbol{\beta}, t)$ is required to converge to the same limit as $w(\boldsymbol{\beta}, t)$ as $n \to \infty$. It can be shown that (3) is equivalent to

$$\sum_{i=1}^{n} \frac{1}{m_i^*} \sum_{j=1}^{m_i^*} w^*(\boldsymbol{\beta}, e_{ij}(\boldsymbol{\beta})) \delta_{ij} \left[\mathbf{Z}_i - \frac{S_1^*\{\boldsymbol{\beta}, e_{ij}(\boldsymbol{\beta})\}}{S_0^*\{\boldsymbol{\beta}, e_{ij}(\boldsymbol{\beta})\}} \right]. \tag{4}$$

It is easy to show that the empirical processes $n^{-1} \sum_{i=1}^{n} N_i^*(\boldsymbol{\beta}, t)$, $n^{-1} \sum_{i=1}^{n} \mathbf{Z}_i N_i^*(\boldsymbol{\beta}, t)$, $S_0^*(\boldsymbol{\beta}, t)$, and $S_1^*(\boldsymbol{\beta}, t)$

converge to the same limits as their respective univariate counterparts and that the mapping defined by U in (3) is compactly differentiable with respect to the supremum norm. As a result, we can prove that $U(\boldsymbol{\beta})$ and its univariate counterpart in (2) converge weakly to the same limiting distribution and converge uniformly to the same limit. The latter ensures the consistency of the solution, denoted by $\hat{\boldsymbol{\beta}}$, to the estimating equation $U(\boldsymbol{\beta}) = 0$.

Note that, while Chang [17] was the first to consider the AFT model for recurrent event data, it is worthwhile to point out that the estimating function proposed in [17] is a special case of (4) with the unit or log-rank weight function, $w^*(\boldsymbol{\beta}, t) = 1$:

$$U_{\mathrm{LR}}(\boldsymbol{\beta}) = \sum_{i=1}^{n} \frac{1}{m_i^*} \sum_{j=1}^{m_i^*} \delta_{ij} \left[\mathbf{Z}_i - \frac{S_1^*\{\boldsymbol{\beta}, e_{ij}(\boldsymbol{\beta})\}}{S_0^*\{\boldsymbol{\beta}, e_{ij}(\boldsymbol{\beta})\}} \right]. \tag{5}$$

The existence of a strongly consistent and asymptotically normal sequence of solutions to $U_{\mathrm{LR}}(\boldsymbol{\beta}) = 0$ was established in [17]; however, the involvement of the unknown parameter $\boldsymbol{\beta}$ in the indicator function renders the estimating function in (5) a non-smooth step function of $\boldsymbol{\beta}$. Hence, a solution $\hat{\boldsymbol{\beta}}_{\mathrm{LR}}$ such that $U_{\mathrm{LR}}(\hat{\boldsymbol{\beta}}_{\mathrm{LR}}) = 0$ may not exist for a finite sample. An alternative approach is to estimate $\boldsymbol{\beta}$ by minimizing the norm of $U_{\mathrm{LR}}(\boldsymbol{\beta})$, that is $\|U_{\mathrm{LR}}(\boldsymbol{\beta})\| = U_{\mathrm{LR}}(\boldsymbol{\beta})^\intercal U_{\mathrm{LR}}(\boldsymbol{\beta})$. However, because monotonicity in $U_{\mathrm{LR}}(\boldsymbol{\beta})$ with respect to $\boldsymbol{\beta}$ is not guaranteed, there may exist multiple solutions to the minimization problem. Therefore, the point estimation based on the non-smooth estimating function in (5) could be computationally challenging in applications.

Because the asymptotic variance of the point estimator depends on the slope of the estimating function in (5), it is difficult to estimate the variance directly when the estimating function is non-smooth. In the literature, resampling-based methods are commonly used for variance estimation. Among them, the bootstrap method is popular due to the ease of implementation. As an alternative, Chang [17] adopted the perturbation technique proposed by Parzen *et al.* [27] to estimate the variance of $\hat{\boldsymbol{\beta}}_{\mathrm{LR}}$. Briefly, since it has been proved that $n^{-1/2}U_{\mathrm{LR}}(\boldsymbol{\beta})$ converges in distribution to a multivariate normal distribution with mean 0 and covariance $V_{\mathrm{LR}}(\boldsymbol{\beta})$, one can first generate a large number of random vectors R's from a multivariate normal distribution with mean 0 and covariance $\hat{V}_{\mathrm{LR}}(\boldsymbol{\beta})$, where $\hat{V}_{\mathrm{LR}}(\boldsymbol{\beta})$ is a consistent estimator of $V_{\mathrm{LR}}(\boldsymbol{\beta})$. Then, one can solve the equation $U_{\mathrm{LR}}(\boldsymbol{\beta}) = R$ to obtain $\hat{\boldsymbol{\beta}}_{\mathrm{LR}}(R)$ for each R. The variance of $\hat{\boldsymbol{\beta}}_{\mathrm{LR}}$ can be approximated by the sample variance of $\hat{\boldsymbol{\beta}}_{\mathrm{LR}}(R)$'s.

Note that both the bootstrap and the perturbation method require solving the estimating equation for a large number of times, which causes the computational burden to increase in a great amount, especially when the estimating function is non-smooth. In addition, the two variance estimation methods rely on the success of each resampling's point estimation whose challenges have been discussed previously.

3. The proposed induced smooth estimating function

Since the rank-based estimating functions discussed in Section 2.2 are non-smooth, causing difficulties in parameter estimation, we propose a monotonic, smooth estimating function in this section. We want to reemphasize that although Johnson and Strawderman [31] have proposed a smooth estimating function for the clustered survival data AFT model, their method cannot be directly applied to the recurrent gap time data because of the unique structure of this type of data.

For univariate survival data, it has been proved that, when using Gehan's weight, the estimating function in (2) is monotonic and corresponds to a convex objective function [38]. If the parameter is estimated by minimizing the objective function, then the set of minimizers would be convex although the minimizer may not be unique. Later, it was showed that applying an *induced smoothing* technique on the rank-based estimating function with Gehan's weight leads to an estimating function which is both smooth and monotonic, essential for improving the computation for both the point and variance estimation [29]. We now consider extending the induced smoothing technique to the setting of recurrent gap time data. We start with the rank-based estimating function for the recurrent gap time data in (4) by using a Gehan-type weight, defined as $w^*(\boldsymbol{\beta},t) = S_0^*(\boldsymbol{\beta},t)$, which converges to the same limit as Gehan's weight for univariate survival data $w(\boldsymbol{\beta},t) = S_0(\boldsymbol{\beta},t)$. The estimating function then becomes

$$\begin{aligned} U_{\mathrm{G}}(\boldsymbol{\beta}) &= \sum_{i=1}^{n}\int_{-\infty}^{\infty} S_0^*(\boldsymbol{\beta},t)\left\{\mathbf{Z}_i - \frac{S_1^*(\boldsymbol{\beta},t)}{S_0^*(\boldsymbol{\beta},t)}\right\} dN_i^*(\boldsymbol{\beta},t) \\ &= \sum_{i=1}^{n}\frac{1}{m_i^*}\sum_{j=1}^{m_i^*} S_0^*\{\boldsymbol{\beta},e_{ij}(\boldsymbol{\beta})\}\delta_{ij}\left[\mathbf{Z}_i - \frac{S_1^*\{\boldsymbol{\beta},e_{ij}(\boldsymbol{\beta})\}}{S_0^*\{\boldsymbol{\beta},e_{ij}(\boldsymbol{\beta})\}}\right] \\ &= \frac{1}{n}\sum_{i=1}^{n}\sum_{j=1}^{m_i^*}\sum_{l=1}^{n}\sum_{k=1}^{m_l^*}\frac{\delta_{ij}}{m_i^* m_l^*}(\mathbf{Z}_i - \mathbf{Z}_l)\mathrm{I}\left\{e_{lk}(\boldsymbol{\beta}) \geq e_{ij}(\boldsymbol{\beta})\right\} \end{aligned} \quad (6)$$

Then, we can apply the induced smoothing technique to the estimating function with the Gehan-type weight in (6) as follows. Let W be a $p\times 1$ independent standard normal vector, then a smoothed estimating function can be proposed by replacing $U_{\mathrm{G}}(\boldsymbol{\beta})$ with $\mathrm{E}_W[U_{\mathrm{G}}(\tilde{\boldsymbol{\beta}})]$, where $\tilde{\boldsymbol{\beta}} = \boldsymbol{\beta} + n^{-1/2}W$, and E_W denotes the expectation with respect to W. This leads to a smooth, monotonic estimating function:

$$U_{\mathrm{G}}^{(s)}(\boldsymbol{\beta}) = \mathrm{E}_W\left[U_{\mathrm{G}}(\tilde{\boldsymbol{\beta}})\right] = \frac{1}{n}\sum_{i=1}^{n}\sum_{j=1}^{m_i^*}\sum_{l=1}^{n}\sum_{k=1}^{m_l^*}\frac{\delta_{ij}}{m_i^* m_l^*}(\mathbf{Z}_i - \mathbf{Z}_l)\mathrm{E}_W\left[\mathrm{I}\{e_{lk}(\tilde{\boldsymbol{\beta}}) \geq e_{ij}(\tilde{\boldsymbol{\beta}})\}\right].$$

It is easy to show that

$$\begin{aligned} \mathrm{E}_W\left[I\left\{e_{lk}(\tilde{\boldsymbol{\beta}}) \geq e_{ij}(\tilde{\boldsymbol{\beta}})\right\}\right] &= \mathrm{E}_W\left[\mathrm{I}\left\{Y_{lk} - (\boldsymbol{\beta}+n^{-1/2}W)^{\mathsf{T}}\mathbf{Z}_l \geq Y_{ij} - (\boldsymbol{\beta}+n^{-1/2}W)^{\mathsf{T}}\mathbf{Z}_i\right\}\right] \\ &= \mathrm{E}_W\left[\mathrm{I}\left\{(\boldsymbol{\beta}+n^{-1/2}W)^{\mathsf{T}}(\mathbf{Z}_l - \mathbf{Z}_i) \leq Y_{lk} - Y_{ij}\right\}\right] \\ &= \Phi\left\{\frac{Y_{lk} - Y_{ij} - \boldsymbol{\beta}^{\mathsf{T}}(\mathbf{Z}_l - \mathbf{Z}_i)}{r_{il}}\right\}, \end{aligned}$$

where $\Phi(\cdot)$ is the cumulative distribution function of a standard normal random variable and $r_{il}^2 = n^{-1}(\mathbf{Z}_l - \mathbf{Z}_i)^{\mathsf{T}}(\mathbf{Z}_l - \mathbf{Z}_i)$. Let $h_{lk,ij}(\boldsymbol{\beta}) = \{Y_{lk} - Y_{ij} - \boldsymbol{\beta}^{\mathsf{T}}(\mathbf{Z}_l - \mathbf{Z}_i)\}/r_{il}$, then we have

$$\mathrm{E}_W\left[I\{e_{lk}(\tilde{\boldsymbol{\beta}}) \geq e_{ij}(\tilde{\boldsymbol{\beta}})\}\right] = \Phi(h_{lk,ij}(\boldsymbol{\beta})).$$

Thus, the resulting smooth estimating function can be expressed as

$$U_{\mathrm{G}}^{(s)}(\boldsymbol{\beta}) = \frac{1}{n}\sum_{i=1}^{n}\sum_{j=1}^{m_i^*}\sum_{l=1}^{n}\sum_{k=1}^{m_l^*}\frac{\delta_{ij}}{m_i^* m_l^*}(\mathbf{Z}_i - \mathbf{Z}_l)\Phi(h_{lk,ij}(\boldsymbol{\beta})). \quad (7)$$

Let $\dot{U}_{\mathrm{G}}^{(s)}(\boldsymbol{\beta}) = \partial\left\{\frac{1}{n}U_{\mathrm{G}}^{(s)}(\boldsymbol{\beta})\right\}/\partial\boldsymbol{\beta}$, then

$$\dot{U}_{\mathrm{G}}^{(s)}(\boldsymbol{\beta}) = \frac{1}{n^2}\sum_{i=1}^{n}\sum_{j=1}^{m_i^*}\sum_{l=1}^{n}\sum_{k=1}^{m_l^*}\frac{\delta_{ij}}{m_i^* m_l^*}\frac{1}{r_{il}}\phi(h_{lk,ij}(\boldsymbol{\beta}))(\mathbf{Z}_i - \mathbf{Z}_l)(\mathbf{Z}_i - \mathbf{Z}_l)^{\mathsf{T}},$$

where $\phi(\cdot)$ is the probability density function of a standard normal random variable. It can be easily shown that the smooth estimating function in (7) is the derivative of the convex objective function

$$L_{\mathrm{G}}^{(s)}(\boldsymbol{\beta}) = \frac{1}{n}\sum_{i=1}^{n}\sum_{j=1}^{m_i^*}\sum_{l=1}^{n}\sum_{k=1}^{m_l^*}\frac{\delta_{ij}}{m_i^* m_l^*}\left[\{e_{lk}(\boldsymbol{\beta}) - e_{ij}(\boldsymbol{\beta})\}\,\Phi(h_{ij,lk}(\boldsymbol{\beta})) + r_{il}\phi(h_{ij,lk}(\boldsymbol{\beta}))\right]. \quad (8)$$

The estimator $\hat{\boldsymbol{\beta}}_{\mathrm{G}}^{(s)}$ is obtained by minimizing the objective function $L_{\mathrm{G}}^{(s)}(\boldsymbol{\beta})$. The consistency and asymptotic normality of $\hat{\boldsymbol{\beta}}_{\mathrm{G}}^{(s)}$ are stated in the following theorem and proved in Appendix.

Theorem 1. Under the regularity conditions in the Appendix, $\hat{\boldsymbol{\beta}}_{\mathrm{G}}^{(s)}$ is a strongly consistent estimator of $\boldsymbol{\beta}_0$ and $\sqrt{n}(\hat{\boldsymbol{\beta}}_{\mathrm{G}}^{(s)} - \boldsymbol{\beta}_0)$ converges in distribution to $N(\mathbf{0}, \Sigma)$, where $\Sigma = A^{-1}V(A^{-1})^{\intercal}$, and $V = \lim_{n\to\infty} \mathrm{Var}\left\{n^{-1/2}U_{\mathrm{G}}(\boldsymbol{\beta}_0)\right\}, A = \partial/\partial\boldsymbol{\beta}\left\{\lim_{n\to\infty} n^{-1}U_{\mathrm{G}}(\boldsymbol{\beta}_0)\right\}$.

Note that the smooth estimator $\hat{\boldsymbol{\beta}}_{\mathrm{G}}^{(s)}$ has the same asymptotic properties as the estimator $\hat{\boldsymbol{\beta}}_{\mathrm{G}}$ (defined in the Appendix) based on the non-smooth estimating function with the Gehan-type weight.

Since the proposed estimating function in (7) is smooth and thus differentiable, one can use $\dot{U}_{\mathrm{G}}^{(s)}(\hat{\boldsymbol{\beta}}_{\mathrm{G}}^{(s)})$ to estimate $A_{\mathrm{G}}(\boldsymbol{\beta}_0)$ [29, 31]. Hence, we propose to use $\dot{U}_{\mathrm{G}}^{(s)}(\hat{\boldsymbol{\beta}}_{\mathrm{G}}^{(s)})^{-1}\hat{V}_{\mathrm{G}}\left\{\dot{U}_{\mathrm{G}}^{(s)}(\hat{\boldsymbol{\beta}}_{\mathrm{G}}^{(s)})^{-1}\right\}^{\intercal}$ to estimate the asymptotic variance of $\sqrt{n}\left(\hat{\boldsymbol{\beta}}_{\mathrm{G}}^{(s)} - \boldsymbol{\beta}_0\right)$, where $\hat{V}_G$ is the sample variance of $\{n^{-1/2}U_{\mathrm{G},b}^{(s)}(\hat{\boldsymbol{\beta}}_{\mathrm{G}}^{(s)}), b = 1, \ldots, N_B\}$ and $U_{\mathrm{G},b}^{(s)}(\hat{\boldsymbol{\beta}}_{\mathrm{G}}^{(s)})$ is the smooth estimating function based on the bth bootstrap sample at $\boldsymbol{\beta} = \hat{\boldsymbol{\beta}}_{\mathrm{G}}^{(s)}$.

4. Simulation

Simulation studies were conducted to assess the performance of the proposed smooth estimating function as compared to the non-smooth rank-based estimating function with various variance estimation methods. For each simulation scenario, 1000 datasets were generated, each with a sample size of $n = 100$ or $n = 200$. All resampling sizes (number of bootstraps or perturbations) were set to be 200.

We began by generating the log gap times $\log(T_{ij}), i = 1, \ldots, n, j = 1, 2, \ldots,$ from the AFT model:

$$\log(T_{ij}) = \beta_1 Z_{i1} + \beta_2 Z_{i2} + \epsilon_{ij}, \tag{9}$$

where $\beta_1 = \beta_2 = 0.5$, and $\epsilon_{ij} = \alpha_i + \epsilon_{ij}^*$. The covariate Z_1 had a Bernoulli distribution with success probability equal to 0.5 and Z_2 followed a uniform distribution on the interval $[0, 1]$. The frailties α_i followed a normal distribution with mean -1 and variance ρ. Two types of distributions of the random errors ϵ_{ij}^* were examined: normal distribution and logistic distribution, and the parameters of the distributions were determined so that ϵ_{ij}^* had mean zero and variance $1 - \rho$. Two values of the variance parameter, $\rho = 0.2, 0.4$, were considered to achieve different levels of within-subject correlations. Note that Model (9) implies a uniform correlation structure and the within subject correlation is ρ. It is easy to prove that the above shared frailty model satisfies the identical marginal distribution condition assumed in Model (1). The censoring times C_i were generated from uniform distributions to yield desirable censoring rates (i.e., percent of subjects without any observed events), $c_p = 25\%$ and 50%.

To show that the proposed method is valid when the data have more complicated correlation structure, we considered scenarios where $\log(T_{ij})$ follow a first-order autoregression or AR (1) model:

$$\log(T_{ij}) = \beta_1 Z_{i1} + \beta_2 Z_{i2} + \alpha + \omega_{ij},$$

where $\alpha = -1$, $\omega_{ij} = \rho\omega_{ij-1} + v_{ij}$ and v_{ij} followed a normal distribution with mean zero and variance $1 - \rho^2$ for $j = 2, \ldots$. We started by generating ω_{i1} from a normal distribution with mean zero and variance 1. Two levels of ρ, 0.2 and 0.4, were considered. It can easily be proved that the above AR(1) model also satisfies the identical marginal distribution condition in Model (1).

With the simulated data, we first compared the performance of the non-smooth estimating equation with either the log-rank weight in (5) [17] or Gehan's weight in (6) to the performance of the proposed smooth estimating equation in (7). The simulation results for data with an uniform correlation structure in normal or logistic random errors, and data with the

Table 1. Results for simulated data with uniform correlation structure and normal random error. The standard error for the non-smooth models, with the log-rank weight and Gehan's weight, is estimated by the bootstrap method or the perturbation method (Parzen et al.); the standard error for the proposed method is estimated by the bootstrap method and the asymptotic variance (ASV) estimator; n is the sample size; c_p is the percent of subjects without events; ρ is the within-subject correlation; $\bar{m}$ is the average number of gap times, observed or censored per subject; Bias is the relative bias computed as the difference of the mean estimated parameter and the true value divided by the true value; SD is the Monte-Carlo standard deviation; ASE is the mean standard error; CP is the proportion of the 95% confidence intervals covering the true value.

					Non-smooth, log-rank weight						Non-smooth, Gehan's weight			
							Bootstrap		Perturbation				Bootstrap	
n	c_p	ρ	$\bar{m}$		Bias	SD	ASE	CP	ASE	CP	Bias	SD	ASE	CP
100	0.25	0.2	3.41	β_1	-0.005	0.211	0.210	0.943	0.212	0.943	0.004	0.184	0.186	0.949
				β_2	-0.026	0.378	0.361	0.927	0.363	0.926	-0.022	0.321	0.323	0.946
		0.4	3.80	β_1	0.005	0.213	0.215	0.958	0.218	0.964	0.005	0.193	0.193	0.958
				β_2	0.017	0.372	0.373	0.938	0.376	0.937	0.021	0.321	0.336	0.952
	0.50	0.2	1.88	β_1	0.004	0.239	0.250	0.958	0.267	0.965	0.006	0.217	0.235	0.964
				β_2	-0.037	0.424	0.436	0.956	0.449	0.957	-0.006	0.394	0.409	0.960
		0.4	2.00	β_1	0.023	0.248	0.248	0.942	0.264	0.958	0.017	0.231	0.233	0.954
				β_2	-0.000	0.438	0.436	0.950	0.447	0.947	0.004	0.404	0.406	0.945
200	0.25	0.2	3.41	β_1	0.002	0.143	0.149	0.956	0.150	0.957	0.009	0.123	0.130	0.956
				β_2	-0.018	0.276	0.257	0.926	0.258	0.931	-0.004	0.231	0.225	0.937
		0.4	3.80	β_1	0.004	0.154	0.152	0.940	0.153	0.944	0.001	0.133	0.134	0.946
				β_2	-0.001	0.261	0.264	0.942	0.264	0.947	-0.002	0.225	0.233	0.960
	0.50	0.2	1.88	β_1	-0.005	0.169	0.173	0.944	0.177	0.951	-0.007	0.153	0.162	0.959
				β_2	-0.009	0.302	0.303	0.943	0.308	0.949	0.003	0.275	0.281	0.953
		0.4	2.00	β_1	0.002	0.169	0.174	0.956	0.177	0.962	0.002	0.158	0.162	0.956
				β_2	0.003	0.301	0.304	0.946	0.310	0.945	0.010	0.280	0.281	0.939

					Proposed, smooth						Naive			
							Bootstrap		ASV		Univariate		Clustered	
n	c_p	ρ	$\bar{m}$		Bias	SD	ASE	CP	ASE	CP	Bias	SD	Bias	SD
100	0.25	0.2	3.41	β_1	0.005	0.184	0.186	0.949	0.182	0.945	0.003	0.217	0.039	0.176
				β_2	-0.022	0.321	0.323	0.944	0.315	0.939	-0.004	0.378	0.017	0.291
		0.4	3.80	β_1	0.005	0.193	0.193	0.956	0.189	0.949	0.010	0.221	0.066	0.234
				β_2	0.022	0.321	0.336	0.950	0.326	0.949	0.024	0.371	0.058	0.400
	0.50	0.2	1.88	β_1	0.008	0.218	0.235	0.964	0.228	0.950	0.016	0.233	0.070	0.220
				β_2	-0.006	0.394	0.408	0.961	0.394	0.954	-0.005	0.423	0.022	0.377
		0.4	2.00	β_1	0.020	0.231	0.233	0.952	0.227	0.943	0.015	0.244	0.120	0.269
				β_2	0.005	0.404	0.406	0.943	0.393	0.934	0.003	0.426	0.056	0.469
200	0.25	0.2	3.41	β_1	0.009	0.123	0.130	0.955	0.129	0.953	0.009	0.148	0.049	0.122
				β_2	-0.004	0.231	0.225	0.936	0.222	0.936	0.018	0.274	0.033	0.211
		0.4	3.80	β_1	0.001	0.133	0.134	0.946	0.133	0.939	0.001	0.155	0.062	0.163
				β_2	-0.002	0.225	0.233	0.960	0.230	0.953	0.002	0.260	0.034	0.280
	0.50	0.2	1.88	β_1	-0.005	0.154	0.162	0.958	0.160	0.956	-0.000	0.163	0.056	0.157
				β_2	0.004	0.275	0.281	0.953	0.277	0.949	0.003	0.295	0.048	0.271
		0.4	2.00	β_1	0.003	0.158	0.162	0.957	0.160	0.952	0.002	0.166	0.120	0.192
				β_2	0.011	0.280	0.281	0.939	0.277	0.934	0.010	0.292	0.109	0.333

AR(1) correlation structure are presented in Tables 1, 2 and 3, respectively. For the point estimates, we report the relative bias (Bias) and the Monte-Carlo empirical standard deviation of the point estimates (SD). For each variance estimation method, we report the average standard errors (ASE) and the coverage percentage (CP) of the 95% confidence intervals (CIs).

The simulation results show that the average point estimates based on the non-smooth and smooth estimating functions

Table 2. Results for simulated data with uniform correlation structure and logistic random error. The standard error for the non-smooth models, with the log-rank weight and Gehan's weight, is estimated by the bootstrap method or the perturbation method (Parzen et al.); the standard error for the proposed method is estimated by the bootstrap method and the asymptotic variance (ASV) estimator; n is the sample size; c_p is the percent of subjects without events; ρ is the within-subject correlation; $\bar{m}$ is the average number of gap times, observed or censored per subject; Bias is the relative bias computed as the difference of the mean estimated parameter and the true value divided by the true value; SD is the Monte-Carlo standard deviation; ASE is the mean standard error; CP is the proportion of the 95% confidence intervals covering the true value.

					Non-smooth, log-rank weight						Non-smooth, Gehan's weight			
							Bootstrap		Perturbation				Bootstrap	
n	c_p	ρ	$\bar{m}$		Bias	SD	ASE	CP	ASE	CP	Bias	SD	ASE	CP
100	0.25	0.2	3.39	β_1	0.002	0.214	0.204	0.939	0.206	0.942	0.018	0.182	0.181	0.945
				β_2	-0.018	0.369	0.354	0.938	0.355	0.930	-0.006	0.313	0.314	0.946
		0.4	3.80	β_1	0.014	0.210	0.211	0.950	0.213	0.955	0.020	0.189	0.189	0.946
				β_2	0.003	0.372	0.367	0.939	0.369	0.940	0.011	0.329	0.329	0.952
	0.50	0.2	1.88	β_1	0.016	0.235	0.237	0.952	0.257	0.958	0.022	0.224	0.226	0.948
				β_2	0.039	0.405	0.412	0.955	0.421	0.951	0.055	0.381	0.389	0.954
		0.4	2.00	β_1	0.000	0.246	0.241	0.949	0.257	0.955	0.007	0.229	0.229	0.948
				β_2	-0.006	0.433	0.423	0.932	0.433	0.931	-0.005	0.398	0.397	0.949
200	0.25	0.2	3.40	β_1	0.013	0.152	0.145	0.942	0.146	0.944	0.018	0.127	0.126	0.948
				β_2	-0.012	0.262	0.252	0.944	0.253	0.945	-0.002	0.221	0.218	0.953
		0.4	3.80	β_1	0.006	0.150	0.150	0.949	0.150	0.947	0.007	0.135	0.132	0.949
				β_2	0.012	0.262	0.261	0.954	0.262	0.951	0.017	0.228	0.230	0.957
	0.50	0.2	1.88	β_1	0.016	0.158	0.165	0.957	0.168	0.958	0.016	0.149	0.156	0.965
				β_2	0.024	0.286	0.289	0.948	0.294	0.947	0.029	0.270	0.270	0.951
		0.4	1.99	β_1	0.005	0.175	0.170	0.931	0.173	0.934	0.008	0.160	0.160	0.944
				β_2	0.013	0.306	0.296	0.943	0.301	0.943	0.004	0.274	0.275	0.949

					Proposed, smooth						Naive			
							Bootstrap		ASV		Univariate		Clustered	
n	c_p	ρ	$\bar{m}$		Bias	SD	ASE	CP	ASE	CP	Bias	SD	Bias	SD
100	0.25	0.2	3.39	β_1	0.019	0.182	0.181	0.943	0.177	0.942	0.013	0.216	0.047	0.172
				β_2	-0.005	0.314	0.314	0.945	0.306	0.941	-0.005	0.377	0.035	0.287
		0.4	3.80	β_1	0.020	0.189	0.189	0.947	0.186	0.941	0.007	0.219	0.072	0.226
				β_2	0.012	0.329	0.329	0.953	0.320	0.947	-0.001	0.371	0.070	0.392
	0.50	0.2	1.88	β_1	0.024	0.225	0.226	0.947	0.219	0.936	0.027	0.240	0.089	0.226
				β_2	0.056	0.381	0.389	0.955	0.376	0.938	0.051	0.400	0.108	0.373
		0.4	2.00	β_1	0.009	0.229	0.229	0.948	0.223	0.942	0.014	0.241	0.130	0.275
				β_2	-0.003	0.398	0.397	0.948	0.384	0.941	-0.002	0.425	0.081	0.446
200	0.25	0.2	3.40	β_1	0.018	0.127	0.126	0.949	0.125	0.946	0.010	0.151	0.049	0.123
				β_2	-0.001	0.221	0.218	0.953	0.215	0.950	-0.007	0.264	0.037	0.211
		0.4	3.80	β_1	0.008	0.135	0.132	0.949	0.131	0.946	0.000	0.152	0.069	0.168
				β_2	0.018	0.227	0.230	0.958	0.228	0.955	0.014	0.256	0.065	0.288
	0.50	0.2	1.88	β_1	0.017	0.149	0.156	0.967	0.154	0.964	0.020	0.161	0.080	0.152
				β_2	0.030	0.270	0.270	0.951	0.265	0.944	0.025	0.284	0.074	0.262
		0.4	1.99	β_1	0.009	0.160	0.160	0.943	0.158	0.938	0.015	0.166	0.125	0.193
				β_2	0.005	0.274	0.275	0.949	0.271	0.948	0.007	0.291	0.087	0.323

are all virtually unbiased. We noticed that under the simulation scenarios that we used, the non-smooth method with the log-rank weight failed to converge for about half a percent of the simulated datasets (the results in the tables are based on the simulated datasets with converged point estimates).

As for the variance estimation, the asymptotic variance estimator of the proposed smooth estimating function gives satisfactory variance estimation with the ASE being close to the Monte-Carlo empirical SD and the bootstrap ASE and

Table 3. Results for simulated data with AR(1) correlation structure. The standard error for the non-smooth models, with the log-rank weight and Gehan's weight, is estimated by the bootstrap method or the perturbation method (Parzen et al.); the standard error for the proposed method is estimated by the bootstrap method and the asymptotic variance (ASV) estimator; n is the sample size; c_p is the percent of subjects without events; ρ is the correlation parameter in the AR(1) correlation structure; $\bar{m}$ is the average number of gap times, observed or censored per subject; Bias is the relative bias computed as the difference of the mean estimated parameter and the true value divided by the true value; SD is the Monte-Carlo standard deviation; ASE is the mean standard error; CP is the proportion of the 95% confidence intervals covering the true value.

					Non-smooth, log-rank weight						Non-smooth, Gehan's weight			
							Bootstrap		Perturbation				Bootstrap	
n	c_p	ρ	$\bar{m}$		Bias	SD	ASE	CP	ASE	CP	Bias	SD	ASE	CP
100	0.25	0.2	3.22	β_1	-0.028	0.213	0.209	0.943	0.211	0.944	-0.009	0.183	0.185	0.954
				β_2	-0.031	0.367	0.365	0.934	0.367	0.938	-0.038	0.319	0.324	0.933
		0.4	3.38	β_1	-0.029	0.207	0.209	0.943	0.211	0.945	-0.014	0.182	0.187	0.948
				β_2	-0.021	0.373	0.363	0.939	0.364	0.942	-0.010	0.325	0.326	0.942
	0.50	0.2	1.85	β_1	0.008	0.251	0.249	0.951	0.265	0.957	0.010	0.230	0.233	0.957
				β_2	-0.027	0.427	0.433	0.950	0.445	0.947	0.007	0.388	0.403	0.958
		0.4	1.92	β_1	0.012	0.261	0.247	0.929	0.265	0.940	0.015	0.241	0.232	0.934
				β_2	-0.021	0.430	0.433	0.948	0.445	0.952	0.006	0.394	0.402	0.945
200	0.25	0.2	3.23	β_1	-0.021	0.150	0.147	0.935	0.148	0.940	-0.010	0.128	0.128	0.941
				β_2	-0.024	0.259	0.257	0.936	0.257	0.942	-0.029	0.221	0.223	0.950
		0.4	3.38	β_1	-0.019	0.149	0.148	0.948	0.149	0.949	-0.012	0.126	0.131	0.947
				β_2	-0.030	0.263	0.256	0.938	0.257	0.934	-0.032	0.225	0.226	0.947
	0.50	0.2	1.85	β_1	-0.022	0.168	0.174	0.946	0.177	0.951	-0.012	0.155	0.161	0.963
				β_2	-0.027	0.299	0.303	0.952	0.308	0.950	0.000	0.273	0.278	0.950
		0.4	1.92	β_1	-0.018	0.177	0.173	0.941	0.176	0.944	-0.014	0.162	0.161	0.942
				β_2	-0.008	0.291	0.302	0.954	0.307	0.955	0.006	0.266	0.278	0.963

					Proposed, smooth						Naive			
							Bootstrap		ASV		Univariate		Clustered	
n	c_p	ρ	$\bar{m}$		Bias	SD	ASE	CP	ASE	CP	Bias	SD	Bias	SD
100	0.25	0.2	3.22	β_1	-0.009	0.183	0.185	0.955	0.181	0.950	0.007	0.218	-0.023	0.143
				β_2	-0.038	0.320	0.324	0.934	0.315	0.931	-0.021	0.385	-0.041	0.252
		0.4	3.38	β_1	-0.013	0.182	0.187	0.948	0.183	0.942	0.012	0.216	-0.030	0.163
				β_2	-0.009	0.325	0.325	0.940	0.317	0.933	-0.002	0.387	-0.034	0.282
	0.50	0.2	1.85	β_1	0.012	0.231	0.233	0.955	0.227	0.947	0.013	0.244	0.018	0.210
				β_2	0.007	0.388	0.402	0.958	0.388	0.951	0.004	0.415	0.017	0.336
		0.4	1.92	β_1	0.017	0.242	0.232	0.935	0.226	0.922	0.028	0.257	0.030	0.236
				β_2	0.007	0.395	0.402	0.943	0.389	0.935	0.003	0.423	0.021	0.366
200	0.25	0.2	3.23	β_1	-0.010	0.128	0.128	0.941	0.127	0.939	0.002	0.153	-0.025	0.101
				β_2	-0.028	0.221	0.223	0.950	0.220	0.946	-0.022	0.269	-0.038	0.174
		0.4	3.38	β_1	-0.012	0.126	0.131	0.946	0.129	0.946	0.012	0.151	-0.032	0.111
				β_2	-0.032	0.225	0.225	0.947	0.223	0.948	-0.012	0.268	-0.053	0.192
	0.50	0.2	1.85	β_1	-0.011	0.155	0.161	0.964	0.159	0.959	-0.009	0.167	-0.002	0.140
				β_2	0.001	0.273	0.278	0.950	0.274	0.946	0.002	0.293	0.020	0.244
		0.4	1.92	β_1	-0.013	0.162	0.161	0.943	0.159	0.939	-0.004	0.172	0.012	0.160
				β_2	0.006	0.266	0.278	0.964	0.275	0.957	0.009	0.285	0.027	0.256

the CP being close to its nominal level (95%). The Monte-Carlo SD of the estimates from the proposed smooth estimating function method and the non-smooth method with Gehan's weight are close; and both are smaller than that of the non-smooth method with log-rank weight [17] for the simulated data. It should be noted that since the bootstrap method and the perturbation method [27] for the non-smooth method with log-rank weight need solving the non-smooth estimating equations for numerous times, the variance estimation suffers from the same non-convergence problem as in the point

Table 4. Regression results for schizophrenia data. SE is standard error estimate; CI is confidence interval.

	Non-smooth						Smooth		
	Log-rank weight			Gehan's weight					
	Estimate	SE	95% CI	Estimate	SE	95% CI	Estimate	SE	95% CI
Log(onset age)	1.444	0.298	(0.860, 2.028)	1.295	0.242	(0.822,1.769)	1.295	0.241	(0.822,1.768)
Gender	0.095	0.276	(-0.445, 0.636)	0.123	0.233	(-0.334,0.580)	0.125	0.235	(-0.335,0.585)

estimation (the ASE and CP in the tables are based on the converged bootstrap samples or perturbed samples only). The computing time of the asymptotic variance estimator based on the proposed smooth estimating function method was substantially shorter than that of the bootstrap or perturbation method of the non-smooth methods as expected.

For comparison, we also applied two existing methods to recurrent gap time data: (1) analyzing the time to first event only data with the induced smoothing method for univariate survival data [29], and (2) applying the induced smoothing method for clustered survival data [31] to the recurrent gap time data, by ignoring their sequential structure. The results are shown in the lower-right panel of Tables 1-3. Whereas the point estimates of the univariate method are satisfactory, this method is obviously less efficient (i.e., larger SDs) than the proposed method. As expected, the point estimates of the clustered survival data method are biased, and the biases increase with the within-subject correlation, which demonstrates that naively applying methods for clustered survival data in the analysis of recurrent gap times can yield substantial bias.

5. Data analysis

We applied the proposed method to the hospitalization data from the Danish Psychiatric Central Register [39] which computerized all admissions to psychiatric hospitals and psychiatric wards in general hospitals in Denmark since 1969. In this paper, we only considered a subset of the published data, which was composed of a cohort of 286 individuals who were first admitted to or contacted with Danish psychiatric services between April 1 and December 31, 1970. The maximum follow-up time was set to be 3 years to avoid any potential change in the distributional pattern of recurrent gap times. The details about this cohort have been described elsewhere [37, 21]. Briefly, among the 286 subjects, 106 (37%) were females, 230 (80%) had schizophrenia onset after 20 years old, 115 (40%) were censored after the initial hospitalization or contact with no records of rehospitalization, 56 (20%) had one rehospitalization, and 115 (40%) had two or more rehospitalization records. The average number of rehospitalization was 1.7. The median disease onset age was 26 with a range of 14 to 88 years old. Note that 9 of the 286 patients died before the end of the follow-up time, hence, the independent censoring assumption was not expected to be seriously violated.

Our main interest was to estimate the effect of the disease onset age on the gap time between two successive hospitalizations. We fitted the AFT model to the data with two covariates, the logarithm-transformed onset age and gender. We applied both the proposed smooth method and non-smooth methods with log-rank or Gehan's weight. The variance for the non-smooth and smooth methods was estimated by the bootstrap and the asymptotic methods, respectively.

As shown in Table 4, the point estimates of the effects of log onset age and gender from the non-smooth and smooth estimating functions are similar, while the CIs from the proposed method and the non-smooth method with Gehan's weight are narrower than the non-smooth method with log-rank weight [17], similar to the findings from the simulation study. All methods show that the effect of onset age was significantly associated with gap times between recurrent hospitalization while gender did not have a significant effect, which is in line with the previous findings in literature [37, 21].

6. Discussion

Despite its appealing direct interpretation, the AFT model [17] has not been widely used in recurrent event data analysis possibly due to the lack of reliable and efficient computing programs. In this paper, we have introduced an induced smoothing technique to improve the performance of the rank-based AFT model for recurrent gap time data. With simulations and a real data analysis, we have shown that the proposed smooth estimating function method provides similar but more computational stable point and variance estimates as compared to the existing non-smooth estimating function method in [17]. The proposed induced smoothing method also has been shown to be more computationally efficient than the non-smooth methods. Hence we recommend to use the proposed induced smoothing method with the asymptotic variance estimator for data analysis.

In this paper, we adopted a Gehan-type weight for the induced smoothing method in order to achieve a more tractable objective function. However, the induced smoothing method is applicable to other weight functions such as the log-rank weight or a general weight function. Note that estimating functions with general weights may not be monotonic. In that case, by following similar techniques in [30], one can use an iterative procedure and within each iteration, reweight a monotonic estimating function in the same form as (6) to approximate the estimating function with a general weight. We note that, like many correlated-data methods, the proposed induced smoothing method for the recurrent gap time AFT model is robust in the sense that its validity does not depend on the correct specification of the correlation structure. A possible future research direction is to improve the efficiency of estimation by incorporating the correlation structure in the estimating function, such as using the generalized method of moments estimation studied by [33] for clustered survival data.

Appendix

We provide a brief proof of consistency and asymptotic normality of $\hat{\boldsymbol{\beta}}_{\mathrm{G}}^{(s)}$ by following the proofs for Theorem 1 and 2 in [31]. We assume the following regularity conditions:

Condition A1. The parameter space $\mathbb{B}$ containing $\boldsymbol{\beta}_0$ is a compact subset of $\mathbb{R}^p$.

Condition A2. $\|Z_i\| + m_i^*$ is uniformly bounded almost surely by a nonrandom constant ($i = 1, \ldots, n$).

Condition A3. Var$(\epsilon_{11}) < \infty$.

Condition A4. The matrix A and V defined in Theorem 1 exist and A is not singular.

Condition A5. Let $f_0(\cdot)$ denote the marginal density associated with model error term ϵ_{11}. Assume $f_0(\cdot)$ and $f_0'(\cdot)$ are bounded functions on $\mathbb{R}$ with

$$\int_{\mathbb{R}} \left\{ \frac{f_0'(t)}{f_0(t)} \right\}^2 f_0(t)dt < \infty.$$

Condition A6. The marginal distribution of C_i is absolutely continuous and has a uniformly bounded density $g_i(\cdot)$ on $\mathbb{R}$ for $i = 1, \ldots, n$.

Among the above conditions, A1, A2, A4, A5, A6 are standard conditions to ensure consistency and the asymptotic normality of the estimator from Equation (6) according to [31]. Since $|\mathrm{Cov}(\epsilon_{ij}, \epsilon_{ik})| \le \mathrm{Var}(\epsilon_{11}), i = 1, \ldots, n, j, k = 1, \ldots, m_i^*$, Condition A3 ensures that the covariances between the error terms of recurrent events of the same person are bounded.

Proof of consistency. We know that the estimating function in Equation (6) is the gradient of convex objective function $L_{\mathrm{G}}(\boldsymbol{\beta}) = \frac{1}{n}\sum_{i=1}^n \sum_{j=1}^{m_i^*} \sum_{l=1}^n \sum_{k=1}^{m_l^*} (m_i^* m_l^*)^{-1} \delta_{ij} \{e_{lk}(\boldsymbol{\beta}) - e_{ij}(\boldsymbol{\beta})\} I\{e_{lk}(\boldsymbol{\beta}) \ge e_{ij}(\boldsymbol{\beta})\}$ which is continuous almost everywhere. Using a similar approach as the proofs for Lemmas 1 and 2 in [31], we can prove that $\sup_{\boldsymbol{\beta} \in \mathbb{B}} |\frac{1}{n} L_{\mathrm{G}}(\boldsymbol{\beta}) - L_0(\boldsymbol{\beta})| \to 0$ almost surely where $L_0(\boldsymbol{\beta})$ is convex for $\boldsymbol{\beta} \in \mathbb{B}$ and $\sup_{\boldsymbol{\beta} \in \mathbb{B}} |\frac{1}{n} L_{\mathrm{G}}^{(s)}(\boldsymbol{\beta}) - L_0(\boldsymbol{\beta})| \to 0$ almost surely. Condition A4 implies that $L_0(\boldsymbol{\beta})$ is strictly convex at $\boldsymbol{\beta}_0$ and thus $\boldsymbol{\beta}_0$ is a unique minimizer of $L_0(\boldsymbol{\beta})$.

Let $\hat{\beta}_{\mathrm{G}}$ be the minimizer of $L_{\mathrm{G}}(\beta)$ and $\hat{\beta}_{\mathrm{G}}^{(s)}$ be the minimizer of $L_{\mathrm{G}}^{(s)}(\beta)$. According to Theorem II.1 and Corollary II.2 in [6], we can conclude that both $\hat{\beta}_{\mathrm{G}}$ and $\hat{\beta}_{\mathrm{G}}^{(s)}$ converge almost surely to β_0.

Proof of asymptotic normality. First we prove the asymptotic normality of $n^{1/2}(\hat{\beta}_{\mathrm{G}}-\beta_0)$. Using similar arguments as in Theorem 2 in [40], we can show that

$$n^{1/2}(\hat{\beta}_{\mathrm{G}}-\beta_0)=-A^{-1}n^{-1/2}U_{\mathrm{G}}(\beta_0)+o_p(1+\sqrt{n}\|\hat{\beta}_{\mathrm{G}}-\beta_0\|).$$

We define

$$\begin{aligned} M_i^*(\beta,t) &= \frac{1}{m_i^*}\sum_{j=1}^{m_i^*}M_{ij}(\beta,t), \\ M_{ij}(\beta,t) &= N_{ij}(\beta,t)-\int_{-\infty}^{t}R_{ij}(\beta,u)\lambda_0(u)du, \text{ and} \\ \bar{z}(\beta,t) &= \frac{\mathrm{E}\{S_1^*(\beta,t)\}}{\mathrm{E}\{S_0^*(\beta,t)\}}, \end{aligned}$$

where $\lambda_0(\cdot)$ is the common hazard function of $\epsilon_{ij}, i=1,\ldots,n, j=1,\ldots,m_i^*$. Let $s_0(\beta,x)=\mathrm{E}\{S_0(\beta,x)\}$ and $s_1(\beta,x)=\mathrm{E}\{S_1(\beta,x)\}$. Following similar argument as in [37], we can show that $\mathrm{E}\{S_0^*(\beta,t)\}=\mathrm{E}\left[n^{-1}\sum_{l=1}^n \mathrm{I}\{e_{l1}(\beta)\geq t\}\right]=s_0(\beta,t)$ and $\mathrm{E}\{S_1^*(\beta,t)\}=\mathrm{E}\left[n^{-1}\sum_{l=1}^n Z_l\mathrm{I}\{e_{l1}(\beta)\geq t\}\right]=s_1(\beta,t)$, and hence we prove that $\bar{z}(\beta,t)=s_1(\beta,t)/s_0(\beta,t)$. Then, following [41], we have

$$\frac{1}{n}U_{\mathrm{G}}(\beta_0)=\frac{1}{n}\sum_{i=1}^n u_i+o_p(n^{-1/2}),$$

where

$$u_i=\int_{-\infty}^{\tau}s_0(\beta_0,t)\{Z_i-\bar{z}(\beta,t)\}dM_i^*(\beta,t).$$

According to central limit theorem, we have $\sqrt{n}\{n^{-1}U_{\mathrm{G}}(\beta_0)\}$ converge in distribution to $N(0,V)$, thus, $\sqrt{n}(\hat{\beta}_{\mathrm{G}}-\beta_0)$ converges in distribution to $N(0,A^{-1}V(A^{-1})^{\intercal})$.

Next we prove the asymptotic normality of $\sqrt{n}(\hat{\beta}_{\mathrm{G}}^{(s)}-\beta_0)$ and show that $\sqrt{n}(\hat{\beta}_{\mathrm{G}}^{(s)}-\beta_0)$ and $\sqrt{n}(\hat{\beta}_{\mathrm{G}}-\beta_0)$ converge to the same limiting distribution. First, following a similar approach as in [31], Lemma 3, we can prove that $\|\dot{U}_{\mathrm{G}}^{(s)}(\beta_0)-A\|\to 0$. Second, since we know that $A^{-1}\left\{n^{-1/2}U_{\mathrm{G}}(\beta_0)\right\}$ is asymptotically normal with mean zero and variance $A^{-1}VA^{-1}$, then if we can prove

$$\sqrt{n}(\hat{\beta}_{\mathrm{G}}^{(s)}-\beta_0)+A^{-1}n^{-1/2}U_{\mathrm{G}}(\beta_0)\to 0 \tag{A.1}$$

in probability, it will imply that $\sqrt{n}(\hat{\beta}_{\mathrm{G}}^{(s)}-\beta_0)$ converge in distribution to $N(0,A^{-1}VA^{-1})$.

Following [42], let $G_n(\beta)=L_{\mathrm{G}}^{(s)}(\beta)$, $\eta_n=n^{-1/2}U_{\mathrm{G}}(\beta_0)$, $M_n=n^{1/2}I_p$, $V_n=(1/2)A$. Then (A.1) can be written as

$$M_n\left(\hat{\beta}_{\mathrm{G}}^{(s)}-\beta_0\right)+\frac{1}{2}V_n^{-1}\eta_n\to 0 \tag{A.2}$$

in probability. According to Theorem 3 in [42], (A.2) holds if the following conditions are met:

Condition B1. $G_n(\beta)$ is convex and $\hat{\beta}_{\mathrm{G}}^{(s)}$ is a sequence satisfying $G_n(\hat{\beta}_{\mathrm{G}}^{(s)})\leq\inf_{\beta\in\mathbb{B}}G_n(\beta)+o_p(1)$.

Condition B2. $\eta_n=O_p(1)$, $\liminf_{n\to\infty}\inf_{|\beta|=1}\beta'V_n\beta>0$ and $\limsup_{n\to\infty}\sup_{|\beta|=1}\beta'V_n\beta<\infty$.

Condition B3. For each $\beta\in\mathbb{R}^P$, $G_n(\beta_0+M_n^{-1}\beta)-G_n(\beta_0)-\beta'\eta_n-\beta'V_n\beta=o_p(1)$.

It is easy to show that Conditions B1 and B2 hold when Conditions A1-A6 hold. We need to prove that Condition B3 holds. By Taylor expansion, we have

$$G_n(\beta_0+M_n^{-1}\beta)=G_n(\beta_0)+(M_n^{-1}\beta)^{\intercal}\left\{\frac{\partial}{\partial\beta}G_n(\beta_0)\right\}+\frac{1}{2}(M_n^{-1}\beta)^{\intercal}\left\{\frac{\partial^2}{\partial\beta^2}G_n(\beta_n^*)\right\}(M_n^{-1}\beta)+o_p(1),$$

then

$$G_n\left(\beta_0 + M_n^{-1}\beta\right) - G_n(\beta_0) - \beta'\left\{n^{-1/2}U_{\mathrm{G}}^{(s)}(\beta_0)\right\} - \frac{1}{2}\beta'\left\{\dot{U}_{\mathrm{G}}^{(s)}(\beta_n^*)\right\}\beta = o_p(1), \tag{A.3}$$

where $\|\beta_n^* - \beta_0\| \leq \|M_n^{-1}\beta\|$. Since $\{n^{-1}U_{\mathrm{G}}^{(s)}(\beta)\}$ is a sequence of bounded, continuously differentiable functions and $\|\dot{U}_{\mathrm{G}}^{(s)}(\beta_0) - A\| \to 0$, $\dot{U}_G^{(s)}(\beta_n^*)$ in (A.3) can be replaced by A. Thus we have

$$G_n(\beta_0 + M_n^{-1}\beta) - G_n(\beta_0) - \beta'\left\{n^{-1/2}U_{\mathrm{G}}^{(s)}(\beta_0)\right\} - \beta' V_n \beta = o_p(1). \tag{A.4}$$

Then Condition B3 holds if we can prove

$$n^{-1/2}\|U_{\mathrm{G}}^{(s)}(\beta_0) - U_{\mathrm{G}}(\beta_0)\| \to 0 \tag{A.5}$$

in probability.

By the definition of $U_{\mathrm{G}}^{(s)}(\beta_0)$, we have

$$U_{\mathrm{G}}^{(s)}(\beta_0) - U_{\mathrm{G}}(\beta_0) = \int_{\mathbb{R}^P} \{U_{\mathrm{G}}(\beta_0 + n^{-1/2}u) - U_{\mathrm{G}}(\beta_0)\}\phi(u)du, \tag{A.6}$$

where $\phi(u)$ is the pdf of W. Define $K_n(u;\beta_0,\Theta) = \|\frac{1}{\sqrt{n}}\{U_{\mathrm{G}}(\beta_0 + n^{-1/2}u) - U_{\mathrm{G}}(\beta_0)\} - \Theta u\|$ where Θ is a fixed matrix that satisfies $\|\Theta\| \leq M$ and $M < \infty$. We know that $\mathrm{E}(W) = \int_{\mathbb{R}^P} u\phi(u) = 0$, so we can derive

$$\begin{aligned}
n^{-1/2}\|U_{\mathrm{G}}^{(s)}(\beta_0) - U_{\mathrm{G}}(\beta_0)\| &= \Bigg\| \int_{\mathbb{R}^P} \left[\frac{1}{\sqrt{n}}\left\{U_{\mathrm{G}}(\beta_0 + n^{-1/2}u) - U_{\mathrm{G}}(\beta_0)\right\} - \Theta u\right]\phi(u)du \\
&\quad + \int_{\mathbb{R}^P} \Theta u\phi(u)du\Bigg\| \\
&\leq \left\| \int_{\mathbb{R}^P} \left[\frac{1}{\sqrt{n}}\left\{U_{\mathrm{G}}(\beta_0 + n^{-1/2}u) - U_{\mathrm{G}}(\beta_0)\right\} - \Theta u\right]\phi(u)du\right\| \\
&\quad + \left\|\int_{\mathbb{R}^P} \Theta u\phi(u)du\right\| \\
&= \int_{\mathbb{R}^P} K_n(u;\beta_0,\Theta)\phi(u)du \\
&= I_1 + I_2,
\end{aligned}$$

where $I_1 = \int_{\|u\|\leq\epsilon_n} K_n(u;\beta_0,\Theta)\phi(u)du$ and $I_2 = \int_{\|u\|>\epsilon_n} K_n(u;\beta_0,\Theta)\phi(u)du$ for any $\epsilon_n > 0$. Following a similar approach as in Theorem 2 in [40], we have

$$\sup_{\|b-\beta_0\|\leq d_n} \frac{\left\|\frac{1}{\sqrt{n}}\{U_{\mathrm{G}}(b) - U_{\mathrm{G}}(\beta_0)\} - A\sqrt{n}(b-\beta_0)\right\|}{1 + n^{1/2}\|b - \beta_0\|} = o_p(1) \tag{A.7}$$

for any positive sequence $d_n \to 0$. Let $b = \beta_0 + n^{-1/2}u$, $d_n = n^{-1/2}\epsilon_n$, $\Theta = A$ and suppose that $\epsilon_n = o(\sqrt{n})$, then it follows Equation (A.7) that

$$\sup_{\|u\|\leq\epsilon_n} \frac{K_n(u;\beta_0,\Theta)}{1 + \|u\|} = o_p(1),$$

which implies $I_1 \to 0$ in probability.

Let $\Theta = A$. Because of the triangle inequality, we have

$$
\begin{aligned}
I_2 = & \int_{\|u\|>\epsilon_n} \left\| \frac{1}{\sqrt{n}} \left\{ U_{\mathrm{G}}\left(\boldsymbol{\beta}_0 + n^{-1/2}u\right) - U_{\mathrm{G}}(\boldsymbol{\beta}_0) \right\} - Au \right\| \phi(u)du \\
\leq & \sqrt{n} \int_{\|u\|>\epsilon_n} \left\| \frac{1}{n} \left\{ U_{\mathrm{G}}\left(\boldsymbol{\beta}_0 + n^{-1/2}u\right) - U_{\mathrm{G}}(\boldsymbol{\beta}_0) \right\} \right\| \phi(u)du \\
& + \int_{\|u\|>\epsilon_n} \|Au\| \phi(u)du \\
\leq & \sup_{\|u\|>\epsilon_n} \left\| \frac{1}{n} \left\{ U_{\mathrm{G}}(\boldsymbol{\beta}_0 + n^{-1/2}u) - U_{\mathrm{G}}(\boldsymbol{\beta}_0) \right\} \right\| \sqrt{n} \int_{\|u\|>\epsilon_n} \phi(u)du \\
& + \|A\| \int_{\|u\|>\epsilon_n} \|u\| \phi(u)du. \qquad \text{(A.8)}
\end{aligned}
$$

Since there is a constant $Q < \infty$ such that $n^{-1}U_{\mathrm{G}}(\boldsymbol{\beta}) < Q$ based on Condition A2, we can derive that the first component in (A.8) is $\leq 2Q\sqrt{n}P(\|W\| > \epsilon_n)$. It is easy to show that a sequence of ϵ_n can be selected so that $\epsilon_n = o(\sqrt{n})$, $\epsilon_n \to \infty$ as $n \to \infty$ and $2Q\sqrt{n}P(\|W\| > \epsilon_n) \to 0$, $\int_{\|u\|>\epsilon_n} \|u\|\phi(u)du \to 0$ as $n \to \infty$. Thus, we have shown that (A.8) $\to 0$ in probability, which implies $I_2 \to 0$ in probability, then (A.5) holds. Therefore, the asymptotic normality of $\hat{\boldsymbol{\beta}}_{\mathrm{G}}^{(s)}$ is proved.

Acknowledgements

The authors thank the University of Minnesota Supercomputing Institute for providing computing resources. This research was supported by NIH/NCI R03CA187991 to Luo, NSF SES-1659328 and NSA H98230-17-1-0308 to Xu, and R01CA193888 to Huang.

References

[1] Heslin KC, Weiss AJ. 2015. Hospital readmissions involving psychiatric disorders, 2012. https://www.hcup-us.ahrq.gov/reports/statbriefs/sb189-Hospital-Readmissions-Psychiatric-Disorders-2012.jsp.

[2] Barker JN, Hough RE, van Burik J-A, DeFor TE, MacMillan ML, OBrien MR, Wagner JE. Serious infections after unrelated donor transplantation in 136 children: Impact of stem cell source. *Biology of Blood and Marrow Transplantation* 2005; **11**:362–370.

[3] Lawless JF, Nadeau C. Some simple robust methods for the analysis of recurrent events. *Technometrics* 1995; **37**(2):158–168.

[4] Nelson W. Confidence limits for recurrence dataapplied to cost or number of product repairs. *Technometrics* 1995; **37**(2):147–157.

[5] Chiang C-T, Wang M-C, Huang C-Y. Kernel estimation of rate function for recurrent event data. *Scandinavian journal of statistics* 2005; **32**(1):77–91.

[6] Andersen PK, Gill RD. Cox's regression model for counting processes: a large sample study. *The Annals of Statistics* 1982; **10**(4):1100–1120.

[7] Pepe MS, Cai J. Some graphical displays and marginal regression analyses for recurrent failure times and time dependent covariates. *Journal of the American Statistical Association* 1993; **88**(423):811–820.

[8] Prentice RL, Williams BJ, Peterson AV. On the regression analysis of multivariate failure time data. *Biometrika* 1981; **68**(2):373–379.

[9] Lin DY, Wei LJ, Yang I, Ying Z. Semiparametric regression for the mean and rate functions of recurrent events. *Journal of the Royal Statistical Society. Series B (Statistical Methodology)* 2000; **62**(4):711–730.

[10] Liu YY, Wu YS. Semiparametric additive intensity model with frailty for recurrent events. *Acta Mathematica Sinica, English Series* 2011; **27**(9):1831–1842.

[11] Schaubel DE, Zeng D, Cai J. A semiparametric additive rates model for recurrent event data. *Lifetime Data Analysis* 2006; **12**(4):389–406.

[12] Wang M-C, Chang S-H. Nonparametric estimation of a recurrent survival function. *Journal of the American Statistical Association* 1999; **94**(445):146–153.

[13] Peña EA, Strawderman RL, Hollander M. Nonparametric estimation with recurrent event data. *Journal of the American Statistical Association* 2001; **96**(456):1299–1315.

[14] Du P. Nonparametric modeling of the gap time in recurrent event data. *Lifetime Data Analysis* 2009; **15**(2):256277.

[15] Du P, Jiang Y, Wang Y. Smoothing spline ANOVA frailty model for recurrent event data. *Biometrics* 2011; **67**(4):1330-1339.

[16] Huang Y, Chen YQ. Marginal regression of gaps between recurrent events. *Lifetime Data Analysis* 2003; **9**(3):293–303.

[17] Chang S-H. Estimating marginal effects in accelerated failure time models for serial sojourn times among repeated events. *Lifetime Data Analysis* 2004; **10**(2):175–190.

[18] Strawderman RL. The accelerated gap times model. *Biometrika* 2005; **92**(3):647–666.

[19] Lu W. Marginal regression of multivariate event times based on linear transformation models. *Lifetime Data Analysis* 2005; **11**(3):389–404.

[20] Sun LQ, Park DH, Sun JG. The additive hazards model for recurrent gap times. *Statistica Sinica* 2006; **16**:919–932.

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

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

[23] Prentice RL. Linear rank tests with right censored data. *Biometrika* 1978; **65**(1):167–179.

[24] Tsiatis AA. Estimating regression parameters using linear rank tests for censored data. *The Annals of Statistics* 1990; **18**(1):354–372.

[25] Kalbfleisch JD, Prentice RL. 2011. *The Statistical Analysis of Failure Time Data*. John Wiley & Sons.

[26] Efron B, Tibshirani RJ. 1993. *An Introduction to the Bootstrap*. Chapman & Hall.

[27] Parzen MI, Wei LJ, Ying Z. A resampling method based on pivotal estimating functions. *Biometrika* 1994; **81**(2):341–350.

[28] Zeng D, Lin DY. Efficient resampling methods for nonsmooth estimating functions. *Biostatistics* 2008; **9**(2):355–363.

[29] Brown BM, Wang Y-G. Induced smoothing for rank regression with censored survival times. *Statistics in Medicine* 2007; **26**(4):828–836.

[30] Chiou S, Kang S, Yan J. Rank-based estimating equations with general weight for accelerated failure time models: an induced smoothing approach. *Statistics in Medicine* 2015; **34**(9):1495–1510.

[31] Johnson LM, Strawderman RL. Induced smoothing for the semiparametric accelerated failure time model: asymptotics and extensions to clustered data. *Biometrika* 2009; **96**(3):577–590.

[32] Chiou S, Kang S, Yan J. Semiparametric accelerated failure time modeling for clustered failure times from stratified sampling. *Journal of the American Statistical Association* 2015; **110**(510):621–629.

[33] Li H, Yin G. Generalized method of moments estimation for linear regression with clustered failure time data. *Biometrika* 2009; **96**(2):293–306.

[34] Wei LJ, Ying Z, Lin DY. Linear regression analysis of censored survival data based on rank tests. *Biometrika* 1990; **77**(4):845–851.

[35] Jin Z, Lin DY, Wei LJ, Ying Z Rank-based inference for the accelerated failure time model. *Biometrika* 2003; **90**(2):341–353.

[36] Gehan EA. A generalized Wilcoxon test for comparing arbitrarily singly-censored samples. *Biometrika* 1965; **52**(1/2):203–223.

[37] Luo X, Huang C-Y. Analysis of recurrent gap time data using the weighted risk-set method and the modified within-cluster resampling method. *Statistics in Medicine* 2011; **30**(4):301–311.

[38] Fygenson M, Ritov Y. Monotone estimating equations for censored data. *The Annals of Statistics* 1994; **22**(2):732–746.

[39] Munk-Jørgensen P, Mortensen PB. The Danish Psychiatric Central Register. *Danish Medical Bulletin* 1997; **44**(1):82–84.

[40] Ying Z. A large sample study of rank estimation for censored regression data. *The Annals of Statistics* 1993; **21**(1):76–99.

[41] Jin Z, Lin DY, Ying Z. Rank regression analysis of multivariate failure time data based on marginal linear models. *Scandinavian Journal of Statistics* 2006; **33**(1):1–23.

[42] Arcones MA. Asymptotic theory for M-estimators over a convex kernel. *Econometric Theory* 1998; **14**(4):387–422.

