University of California, Berkeley

U.C. Berkeley Division of Biostatistics Working Paper Series

Year 2001 *Paper* 101

Marginal Regression of Gaps Between Recurrent Events

Yijian Huang* Ying Qing Chen†

*Division of Public Health Sciences, Fred Hutchinson Cancer Research Center , yhuang5@sph.emory.edu

†Division of Biostatistics, School of Public Health, University of California, Berkeley, yqchen@stat.berkeley.edu

http://biostats.bepress.com/ucbbiostat/paper101

Marginal Regression of Gaps Between Recurrent Events

Yijian Huang and Ying Qing Chen

Abstract

Recurrent event data typically exhibit the phenomenon of intra-individual correlation, owing to not only observed covariates but also random effects. In many applications, the population can be reasonably postulated as a heterogeneous mixture of individual renewal processes, and the inference of interest is the effect of individual-level covariates. In this article, we suggest and investigate a marginal proportional hazards model for gaps between recurrent events. A connection is established between observed gap times and clustered survival data, however, with informative cluster size. We then derive a novel and general inference procedure for the latter, based on a functional formulation of standard Cox regression. Large-sample theory is established for the proposed estimators of the regression coefficients and the baseline cumulative hazard function. Numerical studies demonstrate that the procedure performs well under practical sample sizes. Application to the well-known bladder tumor data is given as illustration

1. Introduction

Recurrent event data arise in longitudinal studies where each individual may experience multiple episodes of the same event. Examples include machine breakdowns, hospitalizations, and tumor recurrences. Various extensions of the original proportional hazards model (Cox, 1972) have been proposed for semiparametric regression analysis of recurrent events. Many of them formulate the covariate effect on the event intensity function as conditioned on the process history; see Andersen and Gill (1982), Prentice, Williams and Peterson (1981), Chang and Hsiung (1994), and Chang and Wang (1999) among others. As a requirement, these baseline processes are either Markov or semi-Markov. It is presumed that the intra-individual correlation, as often evident from data, is due fully to the observed covariates. For some of these models, Pepe and Cai (1993), Lawless and Nadeau (1995), and Lin et al. (2000) relaxed the Markovian requirement, with the covariate effect formulated on the marginal event rate instead (i.e., without conditioning on the complete process history). Nevertheless, such a marginal rate function is necessarily a function of time from the study origin. However, in many applications, gaps between recurrent events is a natural outcome of interest. This article is focused on developing a marginal proportional hazards model for time between recurrent events, which recognizes that intra-individual correlation may not be fully explained by observed covariates.

Our investigation is motivated by the well-known bladder tumor study (Byar, 1980); see also Wei, Lin and Weissfeld (1989). At the study enrollment, all participants had superficial bladder tumors. After these tumors were removed transurethrally, the participants were randomized to three treatment groups: placebo, pyridoxine, and thiotepa. During the study follow-up, many participants had multiple tumor recurrences, and new tumors were removed at each visit. One question of interest is the treatment effect on gaps between recurrent tumors, as well as the effect of other baseline covariates. To be explicit on the data structure,

let $j = 0, 1, 2, \cdots$ index the sequence of recurrent events for an individual, with the follow-up initiating at event $j = 0$. Denote the gap time between events $j-1$ and j by $T_{(j)}$ for $j \geq 1$. The recurrent event process can be represented by the collection $\mathbf{T} \equiv \{T_{(j)} : j = 1, 2, \cdots\}$. With follow-up time C, define the event index M such that $\sum_{j=1}^{M-1} T_{(j)} \leq C$ and $\sum_{j=1}^{M} T_{(j)} > C$, where $\sum_1^0 \cdot \equiv 0$. Thus, M is the number of gaps observed, with the first $M-1$ complete and the last one censored at $T^+_{(M)} \equiv C - \sum_{j=1}^{M-1} T_{(j)}$. Let $\mathbf{Z}$ be a vector of time-independent covariates associated with the individual. The observed data consist of $\{T_{i(j)} : j = 1, \cdots, M_i - 1; T^+_{i(M_i)}; \mathbf{Z}_i\}$, $i = 1, \cdots, n$, which are n iid replicates of $\{T_{(j)} : j = 1, \cdots, M-1; T^+_{(M)}; \mathbf{Z}\}$.

As a characteristic, a recurrent event process typically exhibits certain homogeneity among gaps within an individual. Formally, we adopt the following assumption.

Assumption 1. Each individual recurrent event process is a renewal process. That is, for given i, $T_{i(j)}$, $j = 1, 2, \cdots$, are iid replicates of, say, T_i.

Nevertheless, heterogeneity across individuals may exist, owing to individual-level random effects in addition to observed covariates. We shall model the gap times marginally with the observed covariates. Specifically, formulate a marginal proportional hazards model for the cumulative hazard function of T_i given $\mathbf{Z}_i$:

Assumption 2. $\Lambda(\cdot|\mathbf{Z}_i) = \Lambda(\cdot)\exp(\boldsymbol{\beta}^T\mathbf{Z}_i)$, where $\Lambda(\cdot)$ is a completely unspecified baseline cumulative hazard function and $\boldsymbol{\beta}$ is the vector of regression coefficients.

The conventional independent censorship mechanism is assumed:

Assumption 3. Given $\mathbf{Z}$, $\mathbf{T}$ is independent of C.

The model specifies the recurrent event processes as a possibly heterogeneous mixture of renewal processes, even when $\boldsymbol{\beta} = \mathbf{0}$; note that this baseline setting was studied by Wang

and Chang (1999) for one-sample nonparametric estimation. Section 5 will discuss possible generalization of our model.

The underlying gap times show the structure of clustered survival times. But the two have critical distinctions in observed data; see Chang and Wang (1999). The former is subject to *serial* censoring, whereas the latter is subject to *parallel* censoring in the sense that each survival time has its own censoring time. Therefore, observed gap times cannot be naively treated as clustered survival data in analysis. Nevertheless, in Section 2 we establish a connection between a subset of the observed gap times and clustered survival times with informative cluster size. We propose in Section 3 a novel and general inference procedure for clustered survival data based on a functional formulation of Cox regression, after noting that existing methods (e.g., Lee, Wei and Amato, 1992; Liang, Self and Chang, 1993) accommodate only non-informative cluster size. Simulation studies and the application to the bladder tumor data set are presented in Section 4. Section 5 concludes with discussion.

2. Connection with clustered survival data

Aalen and Husebye (1991) and Wang and Chang (1999) pointed out that recurrent event data are frequently analyzed with inappropriate methods in medical research. To understand the complexity of recurrent event data, we address two important statistical issues. First, due to the between-individual heterogeneity from random effects, for any $j \geq 2$ the censoring on $T_{(j)}$, by $C - T_{(1)} - \cdots - T_{(j-1)}$, is induced to be *dependent*. Second, within an individual, say, i, neither uncensored intervals $T_{i(j)}$, $j = 1, \cdots, M_i - 1$ nor the censored one $T_{i(M_i)}$ are representative of T_i. As Aalen and Husebye (1991) explained, given C_i a longer gap is more likely to contain the censoring time. Thus, $T_{i(M_i)}$ tends to be longer and in turn $T_{i(j)}$, $j = 1, \cdots, M_i - 1$, tend to be shorter. These issues challenge the statistical analysis.

Our proposed inference is based on the establishment of a connection between a subset of the observed gap times and clustered survival data. Observe that standard Cox regression

can be applied to the time-to-the-first-event data for the inference of $\boldsymbol{\beta}$ and $\Lambda(\cdot)$. On the other hand, one finds that for individual i the observed complete gaps, $T_{i(j)}$, $j = 1, \cdots, M_i - 1$, are identically distributed conditional on C_i, M_i, and $T^+_{i(M_i)}$; indeed, $T_{i(j)}$, $j = 1, \cdots, M_i - 1$, have the same probability to take any permutation of the set of observed values. Essentially this result was given in Wang and Chang (1999), as a direct consequence of Assumption 1. Thus, given the exchangeability of the observed complete gaps it becomes intuitively clear that a subset of the observed data can be treated as clustered survival data; however, apparently the cluster size is informative. Specifically, for $M_i > 1$ we remove the censored gap. Write $\Delta_i \equiv I(M_i > 1)$, $S_i \equiv \max(M_i - 1, 1)$, and

$$X_{i(j)} \equiv \begin{cases} T_{i(j)} & \text{if } \Delta_i = 1 \\ T^+_{i(j)} & \text{if } \Delta_i = 0 \end{cases}, \qquad j = 1, \cdots, S_i,$$

where $I(\cdot)$ is the indicator function. The subset consists of $\{X_{i(j)} : j = 1, \cdots, S_i; \Delta_i; \mathbf{Z}_i\}$, $i = 1, \cdots, n$, which are n iid replicates of, say, $\{X_{(j)} : j = 1, \cdots, S; \Delta; \mathbf{Z}\}$.

3. The proposed inference

The inference of interest lies on $\boldsymbol{\beta}$ and $\Lambda(\cdot)$. Our proposal is developed from the functional formulation of Cox regression by Huang and Wang (2000). To focus on the main ideas, the large-sample arguments herewith are purposefully sketchy. Interested reader is referred to Huang and Wang (2000) for additional technical details.

As mentioned earlier, the standard Cox regression procedure can be applied to data $\{X_{i(1)}, \Delta_i, \mathbf{Z}_i\}$, $i = 1, \cdots, n$. Specifically, the (normalized) partial score function is given by

$$\widehat{\mathbf{U}}_{(1)}(\mathbf{b}) = n^{-1} \sum_{i=1}^{n} \Delta_i \left\{ \mathbf{Z}_i - \frac{\sum_{i'=1}^{n} \mathbf{Z}_{i'} \exp(\mathbf{b}^T \mathbf{Z}_{i'}) I(X_{i'(1)} \geq X_{i(1)})}{\sum_{i'=1}^{n} \exp(\mathbf{b}^T \mathbf{Z}_{i'}) I(X_{i'(1)} \geq X_{i(1)})} \right\} I(X_{i(1)} \leq \tau),$$

with time limit τ as a constant. For technical reasons, τ satisfies $\Pr(X \geq \tau) > 0$, where X is a random select from $\{X_{(j)} : j = 1, \cdots, S\}$ and indeed it has the same distribution as $X_{(1)}$. The zero-crossing to $\widehat{\mathbf{U}}_{(1)}(\cdot)$, say, $\widehat{\boldsymbol{\beta}}_{(1)}$, is known to be consistent. However, of concern

is its efficiency, given that only the first gap time from each individual is used. In fact, the structure of clustered survival data established in Section 2 suggests that the first gap time may be replaced by a random choice from the same cluster. To be systematic, one may generate $\prod_{i=1}^{n} S_i$ different sets to replace $\{X_{i(1)} : i = 1, \cdots, n\}$. Naturally, it is expected that the sum of these estimating functions might yield more efficient estimation. However, this approach is too computationally overwhelming to be feasible. In the following, we propose a different method.

With the functional representation of Huang and Wang (2000), reformulate

$$\widehat{\mathbf{U}}_{(1)}(\mathbf{b}) = \widehat{\mathcal{E}}_i\{\mathbf{Z}_i\Delta_i I(X_{i(1)} \le \tau)\} - \int_0^\tau \frac{\widehat{\mathcal{E}}_i\{\mathbf{Z}_i \exp(\mathbf{b}^T\mathbf{Z}_i) I(X_{i(1)} \ge s)\}}{\widehat{\mathcal{E}}_i\{\exp(\mathbf{b}^T\mathbf{Z}_i) I(X_{i(1)} \ge s)\}}\, d\widehat{\mathcal{E}}_i\{\Delta_i I(X_{i(1)} \le s)\},$$

where $\widehat{\mathcal{E}}_i$ represents empirical average over $i = 1, \cdots, n$. As seen, $\widehat{\mathbf{U}}_{(1)}(\cdot)$ is a functional of four empirical processes. Thus, its limit follows the same mapping, from the four corresponding limits. Given that the consistency of $\widehat{\boldsymbol{\beta}}_{(1)}$ is determined by the limit of $\widehat{\mathbf{U}}_{(1)}(\cdot)$, we are motivated to reconstruct the four empirical processes based on the clustered survival data, while maintaining their limits. The resulting estimating function is given as follows,

$$\widehat{\mathbf{U}}(\mathbf{b}) = \widehat{\mathcal{E}}_{ij}\{\mathbf{Z}_i\Delta_i I(X_{i(j)} \le \tau)\} - \int_0^\tau \frac{\widehat{\mathcal{E}}_{ij}\{\mathbf{Z}_i \exp(\mathbf{b}^T\mathbf{Z}_i) I(X_{i(j)} \ge s)\}}{\widehat{\mathcal{E}}_{ij}\{\exp(\mathbf{b}^T\mathbf{Z}_i) I(X_{i(j)} \ge s)\}}\, d\widehat{\mathcal{E}}_{ij}\{\Delta_i I(X_{i(j)} \le s)\}, \tag{1}$$

where $\widehat{\mathcal{E}}_{ij} \equiv \widehat{\mathcal{E}}_i\widehat{\mathcal{E}}_j$ and $\widehat{\mathcal{E}}_j$ averages over $j = 1, \cdots, S_i$. Like $\widehat{\mathbf{U}}_{(1)}(\mathbf{b})$, $\widehat{\mathbf{U}}(\mathbf{b})$ is a monotone function of $\mathbf{b}$. Thus, if attainable, the zero-crossing of $\widehat{\mathbf{U}}(\cdot)$, say, $\widehat{\boldsymbol{\beta}}$, is unique.

From the foregoing discussion, $\widehat{\mathbf{U}}(\cdot)$ has the same limit as $\widehat{\mathbf{U}}_{(1)}(\cdot)$, under mild regularity conditions. Thus, it is implied that $\widehat{\boldsymbol{\beta}}$ is consistent for $\boldsymbol{\beta}$. Now, denote expectation by $\mathcal{E}$. Write $K(s) = \mathcal{E}\{\Delta I(X \le s)\}$, $G_0(s, \mathbf{b}) = \mathcal{E}\{\exp(\mathbf{b}^T\mathbf{Z}) I(X \ge s)\}$, and $\mathbf{G}_1(s, \mathbf{b}) = \mathcal{E}\{\mathbf{Z}\exp(\mathbf{b}^T\mathbf{Z}) I(X \ge s)\}$, which are the limits of $\widehat{K}(s) = \widehat{\mathcal{E}}_{ij}\{\Delta_i I(X_{i(j)} \le s)\}$, $\widehat{G}_0(s, \mathbf{b}) = \widehat{\mathcal{E}}_{ij}\{\exp(\mathbf{b}^T\mathbf{Z}_i) I(X_{i(j)} \ge s)\}$, and $\widehat{\mathbf{G}}_1(s, \mathbf{b}) = \widehat{\mathcal{E}}_{ij}\{\mathbf{Z}_i \exp(\mathbf{b}^T\mathbf{Z}_i) I(X_{i(j)} \ge s)\}$, respectively. The functional version of the Taylor expansion gives

$$\widehat{\mathbf{U}}(\boldsymbol{\beta}) = \widehat{\mathcal{E}}_{ij}\mathbf{w}(X_{i(j)}, \Delta_i, \mathbf{Z}_i) + o_p(n^{-1/2}),$$

where

$$\mathbf{w}(X,\Delta,\mathbf{Z}) = \int_0^\tau \left\{\mathbf{Z} - \frac{\mathbf{G}_1(s,\boldsymbol{\beta})}{G_0(s,\boldsymbol{\beta})}\right\}\left[d\{\Delta I(X\le s)\} - \frac{\exp(\boldsymbol{\beta}^T\mathbf{Z})}{G_0(s,\boldsymbol{\beta})} I(X\ge s)\, dK(s)\right].$$

Therefore, $n^{1/2}\widehat{\mathbf{U}}(\boldsymbol{\beta})$ is asymptotically normal with mean $\mathbf{0}$ and variance $n\boldsymbol{\Sigma} = \mathcal{E}[\{\widehat{\mathcal{E}}_j \mathbf{w}(X_{(j)},\Delta,\mathbf{Z})\}^2]$. This variance can be consistently estimated by $n\widehat{\boldsymbol{\Sigma}} = \widehat{\mathcal{E}}_i[\{\widehat{\mathcal{E}}_j \widehat{\mathbf{w}}(X_{i(j)},\Delta_i,\mathbf{Z}_i)\}^2]$, where

$$\widehat{\mathbf{w}}(X_{i(j)},\Delta_i,\mathbf{Z}_i\} = \int_0^\tau \left\{\mathbf{Z}_i - \frac{\widehat{\mathbf{G}}_1(s,\widehat{\boldsymbol{\beta}})}{\widehat{G}_0(s,\widehat{\boldsymbol{\beta}})}\right\}\left[d\{\Delta_i I(X_{i(j)}\le s)\} - \frac{\exp(\widehat{\boldsymbol{\beta}}^T\mathbf{Z}_i)}{\widehat{G}_0(s,\widehat{\boldsymbol{\beta}})} I(X_{i(j)}\ge s)\, d\widehat{K}(s)\right].$$

We define $\widehat{\boldsymbol{\Gamma}}(\mathbf{b}) = -d\widehat{\mathbf{U}}(\mathbf{b})/d\mathbf{b}^T$, which converges to, say, $\boldsymbol{\Gamma}(\mathbf{b})$. Given that $\widehat{\mathbf{U}}(\mathbf{b})$ is asymptotically linear at $\mathbf{b}=\boldsymbol{\beta}$, it follows that $n^{1/2}(\widehat{\boldsymbol{\beta}}-\boldsymbol{\beta})$ is asymptotically normal with mean $\mathbf{0}$ and variance $n\boldsymbol{\Gamma}(\boldsymbol{\beta})^{-1}\boldsymbol{\Sigma}\boldsymbol{\Gamma}(\boldsymbol{\beta})^{-1}$. Furthermore, the variance can be consistently estimated by $n\widehat{\boldsymbol{\Omega}}_{MF}$, where

$$\widehat{\boldsymbol{\Omega}}_{MF} = \widehat{\boldsymbol{\Gamma}}(\widehat{\boldsymbol{\beta}})^{-1}\widehat{\boldsymbol{\Sigma}}\widehat{\boldsymbol{\Gamma}}(\widehat{\boldsymbol{\beta}})^{-1}. \tag{2}$$

Notice that $\widehat{\boldsymbol{\Omega}}_{MF}$ is a model-free variance estimator. That is, $\widehat{\boldsymbol{\Omega}}_{MF}$ is a consistent variance estimate of the zero-crossing to $\widehat{\mathbf{U}}(\cdot)$, even when Assumptions 1, 2, and 3 are violated.

We now examine the efficiency improvement of $\widehat{\boldsymbol{\beta}}$ over $\widehat{\boldsymbol{\beta}}_{(1)}$. Note that the asymptotic variance of $n^{1/2}\widehat{\mathbf{U}}(\boldsymbol{\beta})$,

$$n\boldsymbol{\Sigma} = \mathcal{E}\{\mathbf{w}(X,\Delta,\mathbf{Z})^2\} - \mathcal{E}[\widehat{\mathcal{E}}_j\{\mathbf{w}(X_{(j)},\Delta,\mathbf{Z}) - \widehat{\mathcal{E}}_j\mathbf{w}(X_{(j)},\Delta,\mathbf{Z})\}^2].$$

Further, the asymptotic variance of $n^{1/2}(\widehat{\boldsymbol{\beta}}_{(1)}-\boldsymbol{\beta})$ is $\boldsymbol{\Gamma}(\boldsymbol{\beta})^{-1}\mathcal{E}\{\mathbf{w}(X,\Delta,\mathbf{Z})^2\}\boldsymbol{\Gamma}(\boldsymbol{\beta})^{-1}$. Thus, the asymptotic variance of $\widehat{\boldsymbol{\beta}}$ is no larger than $\widehat{\boldsymbol{\beta}}_{(1)}$. Meanwhile, this relationship gives rise to an alternative model-based variance estimate for $\widehat{\boldsymbol{\beta}}$. As well known, under the proportional hazards model,

$$\boldsymbol{\Gamma}(\boldsymbol{\beta}) = \mathcal{E}\{\mathbf{w}(X,\Delta,\mathbf{Z})^2\}.$$

Therefore,

$$n\boldsymbol{\Gamma}(\boldsymbol{\beta})^{-1}\boldsymbol{\Sigma}\boldsymbol{\Gamma}(\boldsymbol{\beta})^{-1} = \boldsymbol{\Gamma}(\boldsymbol{\beta})^{-1} - \boldsymbol{\Gamma}(\boldsymbol{\beta})^{-1}\mathcal{E}[\widehat{\mathcal{E}}_j\{\mathbf{w}(X_{(j)},\Delta,\mathbf{Z}) - \widehat{\mathcal{E}}_j\mathbf{w}(X_{(j)},\Delta,\mathbf{Z})\}^2]\boldsymbol{\Gamma}(\boldsymbol{\beta})^{-1}.$$

As a result, the asymptotic variance of $n^{1/2}(\widehat{\boldsymbol{\beta}}-\boldsymbol{\beta})$ can be alternatively estimated by $n\widehat{\boldsymbol{\Omega}}_{MB}$, where

$$\widehat{\boldsymbol{\Omega}}_{MB} = n^{-1}\left[\widehat{\boldsymbol{\Gamma}}(\widehat{\boldsymbol{\beta}})^{-1} - \widehat{\boldsymbol{\Gamma}}(\widehat{\boldsymbol{\beta}})^{-1}\widehat{\mathcal{E}}_i[\widehat{\mathcal{E}}_j\{\widehat{\mathbf{w}}(X_{i(j)},\Delta_i,\mathbf{Z}_i) - \widehat{\mathcal{E}}_j\mathbf{w}(X_{i(j)},\Delta_i,\mathbf{Z}_i)\}^2]\widehat{\boldsymbol{\Gamma}}(\widehat{\boldsymbol{\beta}})^{-1}\right]. \quad (3)$$

In contrast to the model-free variance estimate $\widehat{\boldsymbol{\Omega}}_{MF}$, the validity of $\widehat{\boldsymbol{\Omega}}_{MB}$ depends on Assumptions 1, 2, and 3.

Next, we consider the estimation of the baseline cumulative hazard function $\Lambda(\cdot)$. Again, we start with the standard Breslow estimator with the time-to-the-first-event data, along with its functional representation given in Huang and Wang (2000):

$$\begin{aligned}\widehat{\Lambda}_{(1)}(t;\widehat{\boldsymbol{\beta}}_{(1)}) &= \sum_{i=1}^{n}\frac{\Delta_i I(X_{i(1)}\le t)}{\sum_{i'=1}^{n}\exp(\widehat{\boldsymbol{\beta}}_{(1)}^T\mathbf{Z}_{i'})I(X_{i'(1)}\ge X_{i(1)})}\\ &= \int_0^t \frac{d\widehat{\mathcal{E}}_i\{\Delta_i I(X_{i(1)}\le s)\}}{\widehat{\mathcal{E}}_i\{\exp(\widehat{\boldsymbol{\beta}}_{(1)}^T\mathbf{Z}_i)I(X_{i(1)}\ge s)\}}.\end{aligned}$$

With the very same motivation for $\widehat{\mathbf{U}}(\cdot)$, we suggest using the following estimator to take advantage of the clustered survival data,

$$\widehat{\Lambda}(t;\widehat{\boldsymbol{\beta}}) = \int_0^t \frac{d\widehat{\mathcal{E}}_{ij}\{\Delta_i I(X_{i(j)}\le s)\}}{\widehat{\mathcal{E}}_{ij}\{\exp(\widehat{\boldsymbol{\beta}}^T\mathbf{Z}_i)I(X_{i(j)}\ge s)\}}. \quad (4)$$

Note that, for the one-sample problem, i.e., in the absence of covariates $\mathbf{Z}$, $\widehat{\Lambda}(t;\mathbf{0})$ is an estimator proposed by Wang and Chang (1999). It can be shown that $\widehat{\Lambda}(\cdot;\widehat{\boldsymbol{\beta}})$ is consistent for $\Lambda(\cdot)$ on $[0,\tau]$. Further, $n^{1/2}\{\widehat{\Lambda}(\cdot;\widehat{\boldsymbol{\beta}})-\Lambda(\cdot)\}$ on $[0,\tau]$ converges weakly to a zero mean Gaussian process. In addition, as expected, $\widehat{\Lambda}(\cdot;\widehat{\boldsymbol{\beta}})$ is more efficient than $\widehat{\Lambda}_{(1)}(\cdot;\widehat{\boldsymbol{\beta}}_{(1)})$ in general.

The focus of this article is on recurrent event data. Thus, the clustered survival data under consideration are special in that the members in each cluster share the same censoring indicator and the same covariates. Nevertheless, the inference procedure developed is by no means limited to this special structure, and indeed can be applied generally. In the case of uniform cluster size, our estimators reduce to those of Lee et al. (1992).

4. Numerical studies

We have developed an inference procedure for recurrent event data. In this section, we investigate its performance under practical sample sizes via Monte Carlo simulations and illustrate the procedure through an application to the bladder tumor study. As a convention, in our estimation the time limit τ was set to be large enough to cover all the follow-up times.

Besides estimators $\widehat{\boldsymbol{\beta}}$ and $\widehat{\boldsymbol{\beta}}_{(1)}$, for the purpose of comparison we also naively applied the procedure of Lee et al. (1992) to the complete data set, $\{T_{i(j)} : j = 1, \cdots, M_i - 1; T^{+}_{i(M_i)}; \mathbf{Z}_i\}$, $i = 1, \cdots, n$, and to the subset, $\{X_{i(j)} : j = 1, \cdots, S_i; \Delta_i; \mathbf{Z}_i\}$, $i = 1, \cdots, n$. Denote the two corresponding estimators by $\widetilde{\boldsymbol{\beta}}_a$ and $\widetilde{\boldsymbol{\beta}}_b$. With respect to variance estimation, both model-free and model-based methods are available for $\widehat{\boldsymbol{\beta}}$ and $\widehat{\boldsymbol{\beta}}_{(1)}$. For $\widetilde{\boldsymbol{\beta}}_a$ and $\widetilde{\boldsymbol{\beta}}_b$, only the model-free variance estimate was considered.

4.1 *Simulations*

The following algorithm was adopted to generate a heterogeneous mixture of individual renewal processes, such that the baseline gap time has the standard exponential distribution marginally. Let individual-specific A and episode-specific B be independent mean-zero normal random variables with variances ρ and $1 - \rho$, respectively, for $\rho \in [0, 1]$. Within an individual, a baseline gap time is set to $-\ln\{1 - \Phi(A + B)\}$, where $\Phi(\cdot)$ is the cumulative density function of the standard normal distribution. Thus, marginally the gap time has the standard exponential distribution. Meanwhile, the parameter ρ determines the level of between-individual heterogeneity: $\rho = 0$ indicates the absence of such heterogeneity, i.e., semi-Markov baseline process; $\rho = 1$ corresponds to the situation that all episodes within each individual are equal. Once the baseline recurrence processes are generated, covariate effect under the marginal proportional hazards model can be easily mounted.

Table 1 about here

We considered a single covariate with the standard normal distribution and $\beta = 1$. The follow-up time C was chosen to be uniformly distributed between 0 and an upper limit. Various follow-up limits, values of heterogeneity parameter ρ, and sample sizes were investigated. Note that ρ would not impact the performance of $\widehat{\boldsymbol{\beta}}_{(1)}$, which only takes advantage of the first gap times. Table 1 reports the summary statistics based on 1,000 iterations for each scenario. As shown, estimators $\widetilde{\boldsymbol{\beta}}_a$ and $\widetilde{\boldsymbol{\beta}}_b$ are in general biased, especially when the between-individual heterogeneity becomes substantial. Somewhat unexpectedly, their model-free standard errors also underestimate to a large extend, which might be due to the large variation of cluster sizes in some situations. In contrast, both $\widehat{\boldsymbol{\beta}}_{(1)}$ and $\widehat{\boldsymbol{\beta}}$ exhibit little bias under all these scenarios considered. Meanwhile, $\widehat{\boldsymbol{\beta}}$ is of better precision and its improvement over $\widehat{\boldsymbol{\beta}}_{(1)}$ increases as the between-individual heterogeneity diminishes. Their standard error estimates, both model-free and model-based, perform well, and the 95% confidence intervals achieve accurate coverage probability. Comparatively, our simulations suggest that the model-based estimate is a better choice over the model-free one, when the model assumptions hold.

Simulations were also conducted for models with multiple covariates. The results are similar and, therefore, omitted.

4.2 *Application to the bladder tumor study*

Now we return to the bladder tumor study discussed in Section 1. There are 118 individuals in total: 48, 32, and 38 were randomized to placebo, pyridoxine, and thiotepa, respectively. Overall, 189 tumor recurrences were observed in 62 participants: Among them, 23, 11, 8, 4, 8, 1, 1, 3, and 3 participants experienced from 1 to the maximum 9 tumor recurrences, respectively. Initial examination of the data suggests that a renewable process may be a reasonable assumption for each individual, with obvious between-individual heterogeneity.

Table 2 about here

Figure 1 about here

We adopted the marginal proportional hazards model with four covariates considered: pyridoxine and thiotepa indicators, initial tumor number, and initial tumor size. The results on the estimation of regression coefficients are shown in Table 2. Not surprisingly, both $\widetilde{\boldsymbol{\beta}}_a$ and $\widetilde{\boldsymbol{\beta}}_b$ deviate from $\widehat{\boldsymbol{\beta}}_{(1)}$ and $\widehat{\boldsymbol{\beta}}$ substantially. Nevertheless, $\widehat{\boldsymbol{\beta}}_{(1)}$ and $\widehat{\boldsymbol{\beta}}$ have very similar performance for this data set, in both point and standard error estimation. This suggests a strong intra-individual correlation and so the efficiency gain of $\widehat{\boldsymbol{\beta}}$ over $\widehat{\boldsymbol{\beta}}_{(1)}$ is not apparent. Also, we estimated the baseline cumulative hazard function using $\widehat{\Lambda}_{(1)}$ and $\widehat{\Lambda}$, as shown in Figure 1. Again, the two estimates are similar to each other for this data set.

This bladder tumor data have been analyzed in the literature with other extensions of the proportional hazards model. See, for example, Wei et al. (1989), who analyzed a subset of the data, with the first four episodes in the placebo and thiotepa arms only. Taking gaps between recurrent events as the outcome of interest, our approach complements the existing methods.

5. Remarks

Recurrent event data are often encountered in practice. The analysis of gaps between recurrent events, however, has received relatively little development. In this article, we have suggested and developed a marginal proportional hazards model which allows between-individual heterogeneity arising from random effects. Our proposed inference is based on an established connection between observed gap times and clustered survival data with informative cluster size. Further, a novel and general inference procedure procedure for clustered survival data has been proposed. The procedure is numerically stable and reliable for practical use.

In our model, we have only considered individual-specific covariates. However, some applications might also involve episode-specific covariates. For instance, Wang and Chen

(2000) studied the trend effect over the number of episodes. As such, each individual recurrent process is no longer a renewal process. Furthermore, they showed that the observed complete gaps from an individual may not be *comparable* in general (after the adjustment of covariate effect), given that the observation is dictated by the follow-up time. Thus, the structure of clustered survival times given in Section 2 no longer holds, unless additional criteria are imposed in selecting cluster members. This is one of our current research topics.

Another feature of our covariates under consideration is their time-independence. In fact, this may be relaxed to some extent. Specifically, if one considers covariates that depend on time from the earlier episode and are uniform across all gaps, the model and inference proposed in this article still apply. However, difficulties arise with more complicated time-varying covariates when they are episode-specific.

Acknowledgment

The authors thank Professor Mei-Cheng Wang for stimulating discussion during the research, and the reviewers for helpful comments on earlier versions of this article. Partial support was provided by the National Institutes of Health funds CA90747 and AI29168 and by the Hellman Family Faculty Fund.

References

O. O. Aalen and E. Husebye, "Statistical analysis of repeated events forming renewal processes," *Statistics in Medicine* vol. 10, pp. 1227–1240, 1991.

P. K. Andersen and R. D. Gill, "Cox's regression model for counting processes: A large sample study," *The Annals of Statistics* vol. 10, pp. 1100–1120, 1982.

D. P. Byar, "The veterans administration study of chemoprophylaxis for recurrent stage I bladder tumors: Comparisons of placebo, pyridoxine, and topical thiotepa," in *Bladder Tumors and Other Topics in Urological Oncology*, M. Pavone-Macaluso, P. H. Smith and F. Edsmyn (eds), pp. 363–370, New York: Plenum, 1980.

I.-S. Chang and C. A. Hsiung, "Information and asymptotic efficiency in some generalized proportional hazards models for counting processes," *The Annals of Statistics* vol. 22, pp. 1275–1298, 1994.

S.-H. Chang and M.-C. Wang, "Conditional regression analysis for recurrence time data," *Journal of the American Statistical Association* vol. 94, pp. 1221–1230, 1999.

D. R. Cox, "Regression Models and Life Tables (with Discussion)," *Journal of the Royal Statistical Society (series B)* vol. 34, pp. 187–220, 1972.

Y. Huang and C. Y. Wang, "Cox regression with accurate covariates unascertainable: A nonparametric-correction approach," *Journal of the American Statistical Association* vol. 95, pp. 1209–1219, 2000.

J. F. Lawless and C. Nadeau, "Some simple robust methods for the analysis of recurrent events," *Technometrics* vol. 37, pp. 158–168, 1995.

E. W. Lee, L. J. Wei and D. A. Amato, "Cox-type regression analysis for large numbers of small groups of correlated failure time observations," in *Survival Analysis: State of the Art*, J. P. Klein and P. K. Goel (eds), pp. 237–247, Dordrecht: Kluwer Academic Publishers, 1992.

K.-Y. Liang, S. G. Self and Y.-C. Chang, "Modelling marginal hazards in multivariate failure time data," *Journal of the Royal Statistical Society (series B)* vol. 55, pp. 441–453, 1993.

D. Y. Lin, L. J. Wei, I. Yang and Z. Ying, "Semiparametric regression for the mean and rate functions of recurrent events," *Journal of the Royal Statistical Society (series B)* vol. 62, pp. 711–730, 2000.

M. S. Pepe and J. Cai, "Some graphical displays and marginal regression analyses for recurrent failure times and time dependent covariates," *Journal of the American Statistical Association* vol. 88, pp. 811–820, 1993.

R. L. Prentice, B. J. Williams and A. V. Peterson, "On the regression analysis of multivariate failure time data," *Biometrika* vol. 68, pp. 373–379, 1981.

M.-C. Wang and S.-H. Chang, "Nonparametric estimation of a recurrent survival function," *Journal of the American Statistical Association* vol. 94, pp. 146–153, 1999.

M.-C. Wang and Y. Q. Chen, "Nonparametric and semiparametric trend analysis for stratified recurrence time data," *Biometrics* vol. 56, pp. 789–794, 2000.

L. J. Wei, D. Y. Lin and L. Weissfeld, "Regression analysis of multivariate incomplete failure time data by modeling marginal distributions," *Journal of the American Statistical Association* vol. 84, pp. 1065–1073, 1989.

Figure caption

Figure 1: Estimators of the baseline cumulative hazard function.

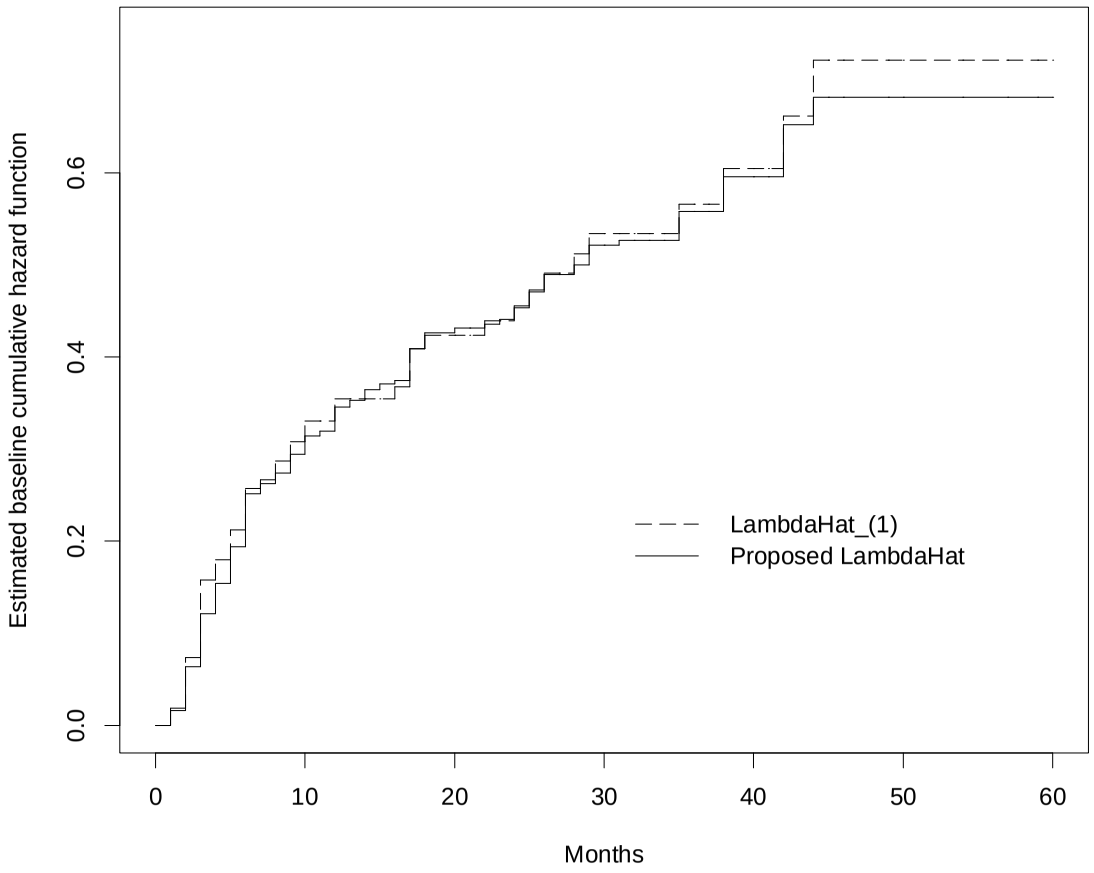

Table 1: Simulation Summary Statistics.

	$\widehat{\boldsymbol{\beta}}_{(1)}$	$\widetilde{\boldsymbol{\beta}}_a$	$\widetilde{\boldsymbol{\beta}}_b$	$\widehat{\boldsymbol{\beta}}$	$\widetilde{\boldsymbol{\beta}}_a$	$\widetilde{\boldsymbol{\beta}}_b$	$\widehat{\boldsymbol{\beta}}$	$\widetilde{\boldsymbol{\beta}}_a$	$\widetilde{\boldsymbol{\beta}}_b$	$\widehat{\boldsymbol{\beta}}$
$\rho =$		.25			.5			.75		
	censoring time: Unif[0,1]									
$\overline{M}$		2.00			2.33			3.39		
	size=50									
B	58.2	4.1	15.3	48.0	-35.7	-37.9	54.3	-142.7	-146.2	64.8
D	313.7	245.1	247.4	279.3	313.0	303.3	291.0	420.9	402.2	303.6
$\widehat{\mathrm{D}}_{MF}$	272.8	196.3	202.0	251.5	219.3	216.9	259.4	244.2	235.1	261.1
C_{MF}	93.2	88.6	90.5	92.2	84.2	83.4	93.1	69.5	70.0	91.2
$\widehat{\mathrm{D}}_{MB}$	295.6	—	—	262.6	—	—	273.5	—	—	279.5
C_{MB}	95.4	—	—	93.0	—	—	94.7	—	—	94.3
	size=100									
B	37.0	-9.3	-8.3	24.7	-49.5	-58.4	33.1	-198.2	-207.7	35.4
D	203.8	160.7	159.9	186.9	212.7	205.1	193.2	308.0	294.6	199.9
$\widehat{\mathrm{D}}_{MF}$	187.6	139.2	139.8	173.0	164.3	158.9	177.8	195.7	187.6	181.8
C_{MF}	92.7	89.6	90.6	93.1	83.9	83.8	92.2	66.8	65.5	91.6
$\widehat{\mathrm{D}}_{MB}$	196.4	—	—	178.8	—	—	184.1	—	—	188.9
C_{MB}	94.9	—	—	93.2	—	—	93.7	—	—	94.0
	censoring time: Unif[0,2]									
$\overline{M}$		3.00			3.74			5.72		
	size=50									
B	41.1	-25.4	-71.0	23.3	-91.8	-134.9	36.7	-232.7	-268.2	29.3
D	255.0	195.5	191.6	219.5	259.2	248.3	224.3	362.6	349.1	246.9
$\widehat{\mathrm{D}}_{MF}$	236.4	159.6	159.3	206.1	191.0	184.7	214.9	219.2	209.2	221.3
C_{MF}	93.0	87.0	84.1	93.3	78.8	74.6	93.5	65.0	59.6	92.4
$\widehat{\mathrm{D}}_{MB}$	252.4	—	—	215.9	—	—	224.0	—	—	234.1
C_{MB}	95.2	—	—	93.9	—	—	93.4	—	—	93.7
	size=100									
B	18.2	-34.9	-86.3	12.9	-103.8	-152.9	20.4	-253.4	-292.8	20.2
D	174.2	135.1	131.6	149.5	195.4	186.6	160.3	295.1	284.6	163.8
$\widehat{\mathrm{D}}_{MF}$	162.9	116.8	115.1	145.3	147.8	141.8	149.5	187.6	179.4	155.4
C_{MF}	93.9	87.9	81.8	93.3	77.4	68.7	92.8	57.7	50.5	93.7
$\widehat{\mathrm{D}}_{MB}$	170.6	—	—	148.4	—	—	154.8	—	—	161.2
C_{MB}	94.5	—	—	93.3	—	—	93.2	—	—	95.0

ρ determines between-individual heterogeneity. $\overline{M}$: average number of observed gaps M;
B: Empirical bias (×1000); D: Empirical standard deviation (×1000); $\widehat{\mathrm{D}}_{MF}$ / $\widehat{\mathrm{D}}_{MB}$: Empirical average of the estimated model-free / model-based standard error (×1000); C_{MF} / C_{MB}: Empirical coverage (%) of the Wald-type 95% confidence interval based on $\widehat{\mathrm{D}}_{MF}$ / $\widehat{\mathrm{D}}_{MB}$.

Table 2: Analysis Results of the Bladder Tumor Study.

	Estimate (SE_{MF},SE_{MB})			
	Pyridoxine	Thiotepa	initial number	initial size
$\widetilde{\beta}_a$	0.080 (0.230, —)	-0.360 (0.205, —)	0.149 (0.046, —)	0.002 (0.048, —)
$\widetilde{\beta}_b$	0.037 (0.219, —)	-0.480 (0.236, —)	0.197 (0.049, —)	0.019 (0.054, —)
$\widehat{\beta}_{(1)}$	-0.343 (0.297,0.322)	-0.540 (0.306,0.313)	0.250 (0.057,0.065)	0.055 (0.075,0.074)
$\widehat{\beta}$	-0.297 (0.298,0.318)	-0.531 (0.306,0.309)	0.258 (0.056,0.064)	0.056 (0.077,0.073)

Pyridoxine and Thiotepa are two treatment indicators, with Placebo as the reference.

Initial size is measured in centimeters.