Semiparametric regression analysis on longitudinal pattern of recurrent gap times

YING QING CHEN†
Division of Biostatistics, School of Public Health, University of California, Berkeley, CA 94720-7360, USA and Department of Statistics, College of Medicine, University of Florida, Gainesville, FL 32610-0212, USA
yqchen@biostat.ufl.edu

MEI-CHENG WANG
Department of Biostatistics, School of Hygiene and Public Health, Johns Hopkins University, Baltimore, MD 21205, USA

YIJIAN HUANG
Division of Public Health Sciences, Fred Hutchinson Cancer Research Center, Seattle, WA 98109, USA

SUMMARY
In longitudinal studies, individual subject may experience recurrent events of the same type over a relatively long period of time. The longitudinal pattern of gaps between successive recurrent events is often of great research interest. In this article, the probability structure of the recurrent gap times is first explored in the presence of censoring. According to the discovered structure, we introduce the stratified proportional reverse-time hazards models with unspecified baseline functions to accommodate individual heterogeneity, when the longitudinal pattern parameter is of main interest. Inference procedures are proposed and studied by way of proper riskset construction. The proposed methodology is demonstrated by the Monte Carlo simulations and an application to a well-known Denmark schizophrenia cohort study data set.

Keywords: Induced dependent censorship; Longitudinal studies; Reverse-time hazard function; Right truncation; Riskset.

1. INTRODUCTION
Since 1938, systematic registration has been conducted in Denmark for mental health patients admitted to hospitals for treatment. The registration includes all the cases from 86 psychiatric institutions in the entire nation of Denmark (Eaton et al., 1992). For patients in this longitudinal follow-up study, individual subject may experience events of the same type, such as recurrent hospitalizations, over a relatively long period of time. When the events are considered as points occurring along the time progression, they form point processes. More examples of this type of recurrent events are recurrent seizures of pediatric cerebral malaria patients or recurrent superficial tumors of cancer patients.

†To whom correspondence should be addressed: Department of Statistics, College of Medicine, University of Florida, P.O. Box 100212, Gainesville, FL 32610-0212, USA.
As noted in Cox and Isham (1980, p. 11), there are generally three equivalent perspectives to study the point processes: (1) the intensity perspective (the complete intensity function of occurrences), (2) the counting perspective (the joint distribution of the occurrence counts in any arbitrary sets), and (3) the gap perspective (the joint distribution of gaps between successive events). Perspectives (1) and (2) are relatively convenient to study in general theory development. The gap perspective, however, is often of more important scientific interest in practice.

For example, in Eaton et al. (1992), there were 8811 patients admitted to the hospitals due to schizophrenic symptoms for the first time in their lives between April 1, 1970 and March 25, 1988. The gap times between two consecutive hospitalizations were selected as endpoint to study the longitudinal pattern of schizophrenia progression, since the distributional tendency of the gaps between the successive hospitalizations serves as an important index of the schizophrenic disease progression: do the gaps progress longer and longer (progressive amelioration), or shorter and shorter (progressive deterioration), over time? If there seems to be such a progressive pattern, can it be tested and the magnitude be estimated?

Suppose that $T$ is the gap time and $Z$ is its episode indicator, taking the values of 1, 2, 3, ..., For all the gap times, we first employ a parametric accelerated failure time model of

$$\log T = \beta Z + e,$$  (1.1)

where $\exp(e)$ follows the Exponential distribution. The estimate of $\beta$ is $-0.0604$ (s.e. = 0.0004). Other distributions of the Weibull, Gamma and Log-normal are also employed, and the estimates are $-0.0640$ (s.e. = 0.0007), $-0.0593$ (s.e. = 0.0010) and $-0.0618$ (s.e. = 0.0009), respectively. In addition, the Cox proportional hazards model of

$$\lambda(t|Z) = \lambda_0(t) \exp(\beta Z)$$  (1.2)

is further employed, and the estimate of $\beta$ is 0.0424 (s.e. = 0.0005). In all the models, the associated $p$-values are highly significant. Although the significance results may be due to the extremely large sample size, these estimates show a seemly similar pattern of progressive deterioration.

In reality, however, study subjects are often heterogeneous in their underlying distributions. The accelerated failure time model (1.1) and the Cox proportional hazards model (1.2) are also fitted for each subject individually. Histograms of the individual estimates for the episode indicators in both models are plotted in Figure 1. As shown in the histograms, the estimates are quite different from subject to subject. Specifically, the estimates in model (1.1) range between $-6.6593$ and $7.4346$ with a median of 0.2512, while those in model (1.2) range between $-1.5260$ and 1.3764 with a median of $-0.0841$.

To accommodate such heterogeneity among individuals, one straightforward approach is by way of the frailty models assuming different frailties for individuals that follow a certain parametric distribution. These models often heavily rely on their parametric assumptions and depend on computer-intensive methods in estimation. The marginalized model often does not maintain the same features such as the proportionality in the Cox model any longer.

Our specific approach is to define and estimate the longitudinal pattern parameters through appropriate regression models, when each individual underlying distribution is considered as nuisance. Toward this end, we risk ourselves with a fast growing number of nuisance parameters when the sample size increases, and therefore encounter the classical problem of Neyman–Scott type. More seriously, when censoring is present, the well-known ‘induced dependent censorship’ further complicates the statistical modeling of recurrent event data and may lead to bias with naive application of traditional survival analysis techniques (Gelber et al., 1989; Huang, 1999; Lin et al., 1999).

Because of the longitudinal nature and the induced dependent censorship, the recurrent event data have distinctive features of their own. In Section 2.1, we first explore some of the features in probability structure of the observed gap times. In Section 2.2, the semiparametric regression models are introduced

*Y. Q. Chen et al.*
2. SEMIPARAMETRIC REGRESSION MODELS

2.1 Probability structure

Suppose there are \( n \) independent subjects. Denote \( i = 1, 2, \ldots, n \) the subject index, and \( j = 0, 1, 2, \ldots \) the recurrent event index. For the \( i \)th subject, \( T_{ij} \) is the gap time between the \((j-1)\)st and the \( j \)th recurrent events, \( j = 1, 2, \ldots \), where \( T_{i,0} \equiv 0 \). Therefore, \((T_{i1}, T_{i2}, \ldots, T_{ij}, \ldots)\) form the collection of random variables of potential gap times. To simplify our discussion, we first assume that there are no covariates according to the discovered structure. In Section 3, we develop model estimation procedures by way of proper riskset construction. Numerical analyses including some Monte Carlo simulations and an analysis of the Denmark psychiatric registry data are in Section 4. Some remaining issues are discussed in Section 5. The technical proofs are collected in the Appendix.
involved, i.e. the underlying \((T_{11}, T_{12}, \ldots)\) are identically distributed with density function of \(f(\cdot)\). Thus,
\[
E(T_{11}) = E(T_{12}) = \cdots = E(T_{ij}) = \cdots.
\] (2.1)

Suppose \(C_i\) is the censoring time starting at \(T_{i0}\) until the occurrence of a censoring event such as lost-to-followup, and \(M_{ij}\) is the event index of stopping time such that
\[
\sum_{k=1}^{M_i-1} T_{ij} \leq C_i \quad \text{and} \quad \sum_{k=1}^{M_i} T_{ij} > C_i.
\]

For individual \(T_{ij}, W_{ij} = C_i - \sum_{k=1}^{j-1} T_{ik}\) denotes its associated censoring time, and \(\Delta_{ij} = I(T_{ij} \leq W_{ij})\) the censoring indicator. Naturally, \(W_{i,1} \geq W_{i,2} \geq \cdots \geq W_{i,M_i}\). Let the observed data be \((t_{i1}, \ldots, t_{im_i-1}, t_{im_i}), (w_{i1,j}, w_{i2,j}, \ldots, w_{im_i,j-1}, w_{im_i,j})\) and \((\delta_{i1,j}, \ldots, \delta_{im_i-1,j}, \delta_{im_i,j}), i = 1, 2, \ldots, n\).

As defined as above, the first \(m_i - 1\) gap times are considered as ‘complete’ gap times, while the last gap time \(t_{im_i}\) is always ‘censored’. For any fixed index \(j \geq 1\), every observed complete \(t_{ij}\) must satisfy that \(\delta_{ij} = 1\). That is, they are observed according to the sampling distribution, \(f_{T_{ij}}(t|\Delta_{ij} = 1)\), which is actually the conditional distribution of \(T_{ij}\) given \(T_{ij} \leq W_{ij}\), and hence right-truncated. Naturally, \(\Pr(T_{ij} > w|\Delta_{ij} = 1, W_{ij} = w) = 0\). Given the monotonicity of \(W_{ij}\) decreasing in \(j\), the observed complete gap times tend to be shorter:
\[
E(T_{11}|\Delta_{11} = 1) \geq E(T_{12}|\Delta_{12} = 1) \geq \cdots \geq E(T_{M_i,M_i-1}|\Delta_{M_i,M_i-1} = 1).
\] (2.2)

Thus, the equalities in (2.1) do not hold for the complete gap times any more. The serial ordering of gap times may simply make the earlier occurrences appear longer than the later ones, even though the underlying distributions are identical.

For the last gap time of \(T_{i,M_i}\), it is always observed subject to intercept sampling (Vardi, 1982). That is, given \(W_{i,M_i} = w\), the sampling distribution of \(T_{i,M_i}\) is the conditional distribution of \(T_{i1}\) given \(T_{i1} \geq w\), and hence left-truncated. Although \(T_{i,M_i}\) is subject to the truncation of opposite direction from the observed complete gaps, it is not counter-balanced by simply pooling together the gap times of all the subjects (Chang and Wang, 1999).

Furthermore, due to unobservable heterogeneity among individuals, \((T_{11}, T_{12}, \ldots)\) tend to correlate with each other within an individual, i.e. they are not necessarily independent. If this is the case, then \(T_{ij}\) and \(W_{ij}\) are no longer independent for any \(j \geq 2\), even though \(C_i\) may be independent of \((T_{i1}, T_{i2}, \ldots)\). So any naive applications of the conventional regression methods such as the Cox model are not appropriate any more for the pooled data \(\{(t_{ij}, \delta_{ij}), i = 1, 2, \ldots, n, j = 1, 2, \ldots, m_i\}\).

### 2.2 Proportional reverse-time hazards models

Since the last censored gap times are of different structure from the complete ones, it is challenging to simply pool all the gap times together for analysis. As previously discussed, the complete gaps are always subject to right-truncation. In fact, researchers have extended the traditional survival techniques in reverse time to the right-truncated failure times (Lagakos et al., 1988; Kalbfleisch and Lawless, 1991; Gross and Huber-Carol, 1992). Denote the cumulative distribution function \(F_{ij}(t) = \Pr(T_{ij} \leq t)\) and its corresponding reverse-time hazard function
\[
\kappa_{ij}(t) = \lim_{\Delta t \to 0^+} \frac{\Pr(t - \Delta t \leq T_{ij} < t|T_{ij} \leq t)}{\Delta t} = \frac{d \log F_{ij}(t)}{dt},
\]
for the \((i, j)\)th gap. Then the natural extension of the Cox proportional hazards model to the right-truncated failure times is the proportional reverse-time hazards model, as recommended in Kalbfleisch...
and Lawless (1991) and Gross and Huber-Carol (1992):

$$
\kappa(t|Z_{ij}) = \kappa_{i0}(t) \exp(\beta^T Z_{ij}),
$$

(2.3)

where $Z_{ij}$ is $p$-dimensional covariate and $\beta \in \mathcal{B} \subset \mathbb{R}^p$ is parameter for $i = 1, \ldots, n$ and $j = 1, 2, \ldots$.

The parameter $\beta$ in (2.3) serves as the longitudinal pattern parameter of recurrent gap times. Its interpretation is better reflected in an equivalent form of (2.3):

$$
F(t|Z_{ij}) = F_{i0}(t) \exp(\beta^T Z_{ij}),
$$

(2.4)

For example, when $Z_{ij}$ is univariate and increases with $j$, $\beta$ represents an assigned trend measure over the longitudinal course of the gap times (Abelson and Tukey, 1963).

In model (2.3), every subject is assumed to have its own baseline reverse-time hazard function. All the $\kappa_{i0}(\cdot)$, $i = 1, 2, \ldots, n$ are treated as nuisance parameters. This allows the model to have much flexibility for the extremely high heterogeneity in the study population. The trade-off, however, is that the gap-independent information contained in individual $\kappa_{i0}(\cdot)$ may not be identifiable in model (2.3). That is, the identifiability of $\beta$ is in doubt if the longitudinal pattern is of little interest: for an extreme example, $Z_{ij}$ are always constant from gap to gap within every subject. This is not the situation concerning the longitudinal pattern as the center topic in this article, though, because certain gap-dependent covariates will have to be included into the models to measure their longitudinal effect.

In (2.3), $\{\kappa_{i0}(t); t \geq 0, i = 1, 2, \ldots, n\}$ are unspecified and hence the models are semiparametric. This is similar to the proportional hazards models or the log-linear models proposed for the paired failure times in Kalbfleisch and Prentice (1980, p. 190). When all the subjects are believed to share similar distributions of baseline characteristics, as a special case of model (2.3), we can include frailties to model $\{\kappa_{i0}(t); t \geq 0, i = 1, 2, \ldots, n\}$,

$$
\kappa_{i0}(t) = \alpha_i \kappa_0(t),
$$

for $i = 1, 2, \ldots, n$, where $\kappa_0(t)$ is unspecified and $\alpha_1, \alpha_2, \ldots, \alpha_n$ are frailties following some parametric distribution function, as in Aalen and Husebye (1991). Thus, the heterogeneity of gap times, i.e. the correlation between gap times of subject $i$, is explicitly explained by the common $\alpha_i$.

3. Inference procedures

3.1 Biased risksets

The concept of riskset was used to develop proper inference procedures in analysis of left-truncated failure times (Woodroofe, 1985; Wang et al., 1986). Brookmeyer and Gail (1994, p. 89) extended the same concept in reverse-time to right-truncated failure times. A proper riskset is supposed to contain a random sample at risk in order to construct the parameter estimators with sound statistical properties, otherwise it is called ‘biased’. We first consider the usual way of riskset construction for the complete gap times as right-truncated observations.

According to the definition in Brookmeyer and Gail (1994), for subjects sharing the same baseline reverse-time hazard function, the individuals in an appropriate riskset at $t_{ij}$, are those ‘whose truncation times [$w_{ik}$] are greater than or equal to’ $t_{ij}$ and ‘whose incubation periods [i.e. observed failure times, $t_{ik}$] are less than or equal to’ $t_{ij}$. As assumed in model (2.3), since only the gap times of same individual share identical $\kappa_0(\cdot)$, the proper riskset at $t_{ij}$ for the $i$th subject is

$$
R_{ij} = \{k : t_{ik} \leq t_{ij} \leq w_{ik}, k = 1, \ldots, m_i - 1\},
$$

(3.1)
where \( w_{ik} = c_i - \sum_{j=1}^{k-1} t_{ij} \) as defined in Section 2.1. This leads to the partial likelihood function in reverse-time as

\[
PL_i = \prod_{j=1}^{m_i-1} \frac{\exp(\beta^T Z_{ij})}{\sum_{k \in R_{ij}} \exp(\beta^T Z_{ik})}
\]

assuming that \( t_{ij} \) are distinct complete gap times. If \( R_{ij} \) were proper, the members in \( R_{ij} \) would form a random sample, and each one of them would have fair probability to fail at \( t_{ij} \). Therefore, the score function that is the derivative of \( \log(PL_i) \),

\[
S_i(\beta) = \sum_{j=1}^{m_i-1} \left\{ Z_{ij} - \frac{\sum_{k \in R_{ij}} Z_{ik} \exp(\beta^T Z_{ik})}{\sum_{k \in R_{ij}} \exp(\beta^T Z_{ik})} \right\},
\]

would be zero unbiased.

However, complication arises with \( R_{ij} \). For a specific \( k \) in \( R_{ij} \) and any \( j < k \), \( t_{ij} \) has influence on \( t_{ik} \) censoring time \( w_{ik} = c_i - \sum_{l<k} t_{il} = c_i - (t_{i1} + \cdots + t_{ij} + \cdots + t_{ik-1}) \). That is, increment (decrement) in \( t_{ij} \) causes decrement (increment) in \( w_{ik} \). As a result, the occurrence ordering of \( t_{ik} \) relative to \( t_{ij} \) alone may determine its membership in \( R_{ij} \), regardless of model (2.3). Because of this embedded longitudinal nature, \( R_{ij} \) are not random samples at risk and hence biased. The estimators constructed by \( S_i(\beta) \) of the biased risksets are no longer guaranteed with the sound statistical properties as basic as consistency.

### 3.2 Unbiased reduced risksets

As discussed in Section 3.1, the cause of biased risksets is clear, i.e. the gaps in \( R_{ij} \) do not have fair probabilities to fail at \( t_{ij} \). An ideal correction is to allow fair probability for \( t_{ik} \in R_{ij} \) to fail at \( t_{ij} \). That is, if we replace every gap \( t_{ik} \) with \( t_{ij} \) in \( R_{ij} \),

\[
\frac{t_{ij} + t_{ij} + \cdots + t_{ij} + \sum_{l \in R_{ij}^c} t_{il}}{|R_{ij}|} = |R_{ij}| t_{ij} + \sum_{l \in R_{ij}^c} t_{il} \leq c_i \tag{3.2}
\]

would still hold, where \( |R_{ij}| \) is the size of \( R_{ij} \) and \( R_{ij}^c = \{1, 2, \ldots, m_i-1\}\setminus R_{ij} \). Then unbiased estimating functions can be constructed based on the risksets satisfying (3.2). However, this way of construction is expected to be cumbersome because (3.2) needs to be verified for at most \( 2^{m_i-1} \) times to obtain the maximal \( R_{ij} \).

More feasible approaches can be considered by limiting the number of replacements of \( t_{ik} \) in \( R_{ij} \). This can be achieved by reducing \( |R_{ij}| \). The most aggressive reduction is to include only one gap \( t_{ik} \in R_{ij} \), say, at a time, in addition to \( t_{ij} \) itself. That is, a meaningful reduced riskset \( R_{ij} \) at \( t_{ij} \) would always have two gaps, \( t_{ij} \) and \( t_{ik} \). To explore the eligibility of \( t_{ik} \in R_{ij} \), we plot two plausible situations in Figure 2 that may arise in reality:

1. \( k > j \), \( R_{ij} \) is to include a later gap and therefore \( w_{ij} > w_{ik} \). In this case, the usual condition for riskset construction of right-truncated observations applies. That is, \( t_{ik} \leq t_{ij} \leq w_{ik} \leq w_{ij} \).
2. \( k < j \), \( R_{ij} \) is to include an earlier gap and therefore \( w_{ik} > w_{ij} \). Since \( w_{ik} \) is \( t_{ik} \)'s right-truncation time, it is possible that \( w_{ik} \geq t_{ik} > w_{ij} > t_{ij} \). If this is the case, it is impossible for \( t_{ij} \) to be as big as \( t_{ik} \). Thus, we need to curtail \( w_{ik} \). The largest room for \( t_{ij} \) to grow is \( w_{ij} - t_{ij} \). It is also supposed to be the largest room left for \( t_{ik} \). Therefore, the curtailed right-truncation time for \( t_{ik} \) should be \( t_{ik} + (w_{ij} - t_{ij}) \).

In summary, two gaps must satisfy one of the following two conditions in the unbiased \( R_{ij} \): (a) for \( k > j \), \( t_{ik} \leq t_{ij} \leq w_{ik} \), (b) for \( k < j \), \( t_{ik} \leq t_{ij} \leq w_{ij} - t_{ij} + t_{ik} \). Therefore, \( |R_{ij}| \) is 2 if either condition holds, and degenerates to 1 otherwise.
3.3 Inferences based on reduced risksets

The unbiased reduced risksets constructed in Section 3.2 are neither necessarily existing, nor necessarily unique when existing. Denote \( \tilde{R}_{ijk}, k = 1, 2, \ldots, m_i - 1 \) the entire reduced risksets at \( t_{ij} \), and \( \delta_{ijk} = I(|\tilde{R}_{ijk}| > 1) \) the unbiased \( \tilde{R}_{ijk} \) indicator. Given \( \tilde{R}_{ijk} \) and \( \delta_{ijk} = 1 \), the conditional likelihood contribution of \( \tilde{R}_{ijk} \) is then

\[
\exp(\beta^T Z_{ij}) \exp(\beta^T Z_{ik}) + \exp(\beta^T Z_{ik}),
\]

which resembles the ones from the Cox proportional hazards model for paired failure times as in Kalbfleisch and Prentice (1980, p. 191). Its corresponding score function is

\[
\tilde{S}_{ijk}(\beta) = Z_{ij} - \tilde{Z}_{ijk}(\beta),
\]

where

\[
\tilde{Z}_{ijk}(\beta) = \frac{Z_{ij} \exp(\beta^T Z_{ij}) + Z_{ik} \exp(\beta^T Z_{ik})}{\exp(\beta^T Z_{ij}) + \exp(\beta^T Z_{ik})}.
\]
It is true that \( E(\tilde{S}_{ij}(\beta)\tilde{R}_{ijk}, \delta_{ijk} = 1; \beta) = 0 \). In addition, \( E(\tilde{S}_{ij}(\beta)\tilde{R}_{ijk}, \delta_{ijk} = 0; \beta) = 0 \). Therefore, we can use \( \tilde{S}_{ij}(\beta) \) as ‘building blocks’ to construct the estimating function for subject \( i \) as

\[
\tilde{S}_i(\beta) = \frac{\sum_{j=1}^{n_i} \sum_{k=1}^{m_{ij}} \delta_{ijk} \tilde{S}_{ijk}(\beta)}{\sum_{j=1}^{n_i} \sum_{k=1}^{m_{ij}} \delta_{ijk}}.
\]

If we let \( \delta_{ij} = \sum_{k=1}^{m_{ij}} \delta_{ijk}, g_{ij} = \delta_{ij}/ \sum_{k=1}^{m_{ij}} \delta_{ij} \) and \( \tilde{Z}_{ij}(\beta) = \delta_{ij}^{-1} \sum_{k=1}^{m_{ij}} \delta_{ijk} \tilde{Z}_{ijk}(\beta) \), straightforward algebraic manipulation shows that the ultimate set of estimating functions using all the subjects are

\[
\tilde{S}(\beta) = n^{-1} \sum_{i=1}^{n} \tilde{S}_i(\beta) = n^{-1} \sum_{i=1}^{n} \sum_{j=1}^{m_i} g_{ij} \{ \tilde{Z}_{ij} - \tilde{Z}_{ij}(\beta) \}.
\]

It is straightforward that \( \tilde{S}(\beta) \) is unbiased and thus the estimators of \( \beta \) can be obtained by solving \( \tilde{S}(\beta) = 0 \).

With the special way of the reduced riskset construction, the martingale theory for counting processes of the usual proportional reverse-time hazards model may not be applied in any straightforward sense. However, standard asymptotic likelihood methods will be able to show the existence of \( \hat{\beta} \), its uniqueness and consistency under the assumed regularity conditions in the Appendix.

In addition, since \( \tilde{S}(\beta) \) is the sum of \( \{ \tilde{S}_i(0) \}_{i=1}^{n} \) as iid unbiased estimating functions, it is true that \( n^{-1/2} \tilde{S}(\beta) \) is asymptotically normal with mean zero and variance \( \Sigma(\beta) \) by the Central Limit Theorem. Following the consistency of \( \hat{\beta} \) and a Taylor series expansion, we are able to establish the asymptotic normality of \( \hat{\beta} \) under the assumed regularity conditions. That is,

\[
n^{-1/2}(\hat{\beta} - \beta_0) \overset{D}{=} N(0, D^{-1}(\beta_0) \Sigma(\beta_0) [D^{-1}(\beta_0)]^T).
\]

Therefore, a consistent estimator of \( D^{-1}(\beta_0) \Sigma(\beta_0) [D^{-1}(\beta_0)]^T \) can be obtained by replacing \( \beta_0 \) with \( \hat{\beta} \), that is, \( n^{-1} \hat{\Sigma}(\beta) = D^{-1}(\hat{\beta}) \Sigma(\hat{\beta}) [D^{-1}(\hat{\beta})]^T \), where \( \hat{\Sigma}(\beta) = E([\partial/\partial \beta] \tilde{S}_i(\beta)) \). Technical details of the proof are given in the Appendix.

In practice, to solve the estimating equation, a Newton–Raphson iteration algorithm can be adapted. That is, at the \( k \)th step of iteration, let the \((k + 1)\)st solution to the equation to be

\[
\hat{\beta}(k+1) = \hat{\beta}(k) + \hat{D}^{-1}(\hat{\beta}(k)) \tilde{S}(\hat{\beta}(k)).
\]

In our experience, this algorithm is reasonably efficient and the computation burden is not demanding. The variance estimation of sandwich-type is also straightforward.

Since the estimating equations are constructed from the conditional score functions based on the eligible reduced risksets with equal weight, it is not expected that the proposed estimating equations would be fully efficient in general. However, if one prefers, deterministic weights can be added to the components in \( \tilde{S}(\beta) \) to enable potentially more efficient estimating equations. For example, let

\[
\tilde{S}_G(\beta) = n^{-1} \sum_{i=1}^{n} \sum_{j=1}^{m_i-1} G_{ij}^{-1/2} g_{ij} \{ \tilde{Z}_{ij} - \tilde{Z}_{ij}(\beta) \},
\]

where \( G_{ij} \) is the diagonal matrix with identical diagonal elements in \( E [g_{ij} | Z_{ij} - \tilde{Z}_{ij}(\beta)] \).

Similar to the Cox proportional hazards model, the assumption of multiplicative form is critical to the proportional reverse-time hazards models. To assess the model adequacy, the rationale in Gill and Schumacher (1987) for the Cox proportional hazards model can be adopted. Denote \( \hat{\beta}_G \) the solution to
\( \tilde{S}_G(\hat{\beta}) = 0 \). Then we can compare in difference the two corresponding estimators of \( \hat{\beta}_G \) and \( \hat{\beta} \), using the quadratic form

\[
T_{GS} = (\hat{\beta}_G - \hat{\beta})^\top \hat{V}^{-1} (\hat{\beta}_G - \hat{\beta}),
\]

where \( \hat{V} \) is an appropriate estimator of variance-covariance matrix of \( \hat{\beta}_G - \hat{\beta} \). The statistic \( T_{GS} \) is asymptotically \( \chi^2_p \) if the proposed models are true.

4. Numerical Studies

4.1 Simulations

Monte Carlo simulation studies are conducted to evaluate the validity of the proposed inference procedures, when the underlying models are the reverse-time hazards models as specified in (2.3). The following frailty models are used in our simulation studies:

\[
\kappa_{ij}(t) = \alpha_i \kappa_0(t) \exp(\beta_1 Z_{ij1} + \beta_2 Z_{ij2}), \tag{4.1}
\]

where \( Z_{ij} = (Z_{ij1}, Z_{ij2}) \) and \( (\beta_1, \beta_2) \) are two-dimensional vectors, for \( j = 1, 2, \ldots, i = 1, \ldots, n \). Here, \( (\alpha_1, \alpha_2, \ldots, \alpha_n) \) are frailties generated according to the Gamma distribution

\[
f(\alpha) = c_0 \Gamma(c_1)^{-1} \alpha^{c_1-1} \exp(-c_0 \alpha),
\]

with \( c_0 = c_1 = 2 \) to mimic the heterogeneity and hence the correlation among the gap times of an individual subject. The baseline reverse-time hazard function \( \kappa_0(t) \) is of the Weibull density

\[
f(t) = c_2 t^{c_2-1} \exp(-t^{c_2})
\]

with \( c_2 \) being 0.8, 1 and 2.5, representing decreasing, constant and increasing hazard functions, respectively. The censoring times \( C_i, i = 1, 2, \ldots, n \), are generated by the exponential distribution with mean \( \mu \), which is selected to be 10 or 15 to represent relatively short or long follow-up period, respectively.

Sample sizes are selected to be 100 and 250 for relatively small and large sample sizes, respectively. Two covariates are used: \( Z_{ij1} = j \) for trend measure, while \( Z_{ij2} = e_{ij} \) simulated from uniform distribution \( U[0, 1] \) to represent some time-dependent confounding variable needed to be adjusted. True parameters \( \beta_0 = (\beta_{10}, \beta_{20}) \) are selected to be \( (0, 0) \), \( (1, 0) \), \( (0, 1) \) and \( (1, 1) \), respectively. For each configuration, 10000 simulations are conducted. Its empirical bias, defined as the difference between empirical mean and the true parameter, and coverage probabilities are computed. Details of results are listed in Table 1. As shown in the table, the proposed estimators are virtually unbiased and the corresponding confidence intervals have proper confidence levels.

4.2 Application to Danish registry data

As mentioned in Section 1, there were 8811 patients in the Danish registry data by Eaton et al. (1992). Of all the patients, 5493 patients were male and 3318 were female, and about 67.8% of the patients (5972) were married. Their ages of onset differed with 1065 patients of onset ages less than or equal to 20, and 7746 more than 20. About 4324 patients in the registry data never had any hospitalization before, 3938 patients have 1–5 prior hospitalizations and 549 patients with more than 5 prior hospitalizations.

A testing procedure was proposed and applied to this data set in Wang and Chen (2000) to detect whether or not there was a progressive trend in the gap times of hospitalizations. It was found that there
was a deterioration patterns of the disease among the patients. A regression model based on the accelerated failure time model was also used to estimate the magnitude of deterioration.

To contrast with their findings, we also choose the same index of trend measure of $Z_{ij}$ = $j$ in model (2.3). The estimate of parameter $\beta$ is $-0.0033$ (s.e. = 0.0009, $p = 0.0002$). Their negative sign suggests deterioration pattern, which is consistent with the reported results in Wang and Chen (2000). For both male and female patients, they share similar deterioration patterns with estimates of $-0.0015$ (s.e. = 0.0010, $p > 0.2$) and $-0.0107$ (s.e. = 0.0019, $p < 0.0001$), respectively, although it was not significant in male patients. For married patients, the deterioration pattern was significant with an estimate of $-0.0131$ (s.e. = 0.0022, $p < 0.001$), while it was not as significant among the unmarried patients ($\hat{\beta} = -0.0015$, s.e. = 0.0010, $p > 0.1$). There was deterioration pattern significantly shown in the late onset age group ($\hat{\beta} = -0.0057$, s.e. = 0.0010, $p < 0.0001$), while there was seemingly amelioration pattern in the early onset age group ($\hat{\beta} = 0.0025$, s.e. = 0.0016, $p > 0.1$). The direction of deterioration pattern was persistent no matter whether there was prior hospitalizations. But the pattern is significantly deteriorating for those with at least one prior hospitalization, while it is not for those without any prior hospitalizations.

Since the admission age may confound the hospitalization trend, model (2.3) is further fitted to

---

**Table 1. Summary of simulation studies**

<table>
<thead>
<tr>
<th>$n$</th>
<th>$\mu$</th>
<th>$c^2$</th>
<th>$(0,0)$</th>
<th>$(1,0)$</th>
<th>$(\beta_{01}, \beta_{02})$</th>
<th>$(1,1)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>100</td>
<td>10</td>
<td>0.8</td>
<td>$-0.0060$</td>
<td>$-0.0033$</td>
<td>$0.0059$</td>
<td>$0.0061$</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Cov. Pr.</td>
<td>0.9526</td>
<td>0.9472</td>
<td>0.9475</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>1.0 Bias</td>
<td>0.0023</td>
<td>0.0026</td>
<td>0.0042</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Cov. Pr.</td>
<td>0.9511</td>
<td>0.9473</td>
<td>0.9508</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>2.5 Bias</td>
<td>0.0047</td>
<td>$-0.0100$</td>
<td>0.0039</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Cov. Pr.</td>
<td>0.9498</td>
<td>0.9511</td>
<td>0.9483</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>15 Bias</td>
<td>$-0.0028$</td>
<td>$-0.0014$</td>
<td>0.0025</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Cov. Pr.</td>
<td>0.9522</td>
<td>0.9509</td>
<td>0.9534</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>1.0 Bias</td>
<td>0.0090</td>
<td>$-0.0018$</td>
<td>0.0031</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Cov. Pr.</td>
<td>0.9505</td>
<td>0.9496</td>
<td>0.9476</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>2.5 Bias</td>
<td>$-0.0015$</td>
<td>0.1011</td>
<td>$-0.0080$</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Cov. Pr.</td>
<td>0.9516</td>
<td>0.9485</td>
<td>0.9482</td>
</tr>
<tr>
<td>250</td>
<td>10</td>
<td>0.8</td>
<td>$-0.0012$</td>
<td>0.0005</td>
<td>$-0.0003$</td>
<td>$-0.0043$</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Cov. Pr.</td>
<td>0.9527</td>
<td>0.9494</td>
<td>0.9455</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>1.0 Bias</td>
<td>0.0061</td>
<td>0.0028</td>
<td>$-0.0021$</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Cov. Pr.</td>
<td>0.9477</td>
<td>0.9496</td>
<td>0.9523</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>2.5 Bias</td>
<td>0.0034</td>
<td>0.0029</td>
<td>0.0011</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Cov. Pr.</td>
<td>0.9522</td>
<td>0.9491</td>
<td>0.9493</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>15 Bias</td>
<td>0.0013</td>
<td>0.0016</td>
<td>$-0.0011$</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Cov. Pr.</td>
<td>0.9488</td>
<td>0.9502</td>
<td>0.9507</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>1.0 Bias</td>
<td>$-0.0023$</td>
<td>$-0.0016$</td>
<td>$-0.0045$</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Cov. Pr.</td>
<td>0.9505</td>
<td>0.9524</td>
<td>0.9491</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>2.5 Bias</td>
<td>0.0048</td>
<td>0.0019</td>
<td>0.0018</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Cov. Pr.</td>
<td>0.9507</td>
<td>0.9466</td>
<td>0.9482</td>
</tr>
</tbody>
</table>

$\mu$ is mean censoring time. $c^2$ is the shape parameter of the baseline hazard function. Bias is the average $\beta$ minus $\beta_0$. Cov. Pr. is the coverage probability of the 95% confidence intervals. All the entries are computed from 10 000 simulations.
measure the adjusted trend. Overall, the adjusted pattern is still significantly deteriorating ($\hat{\beta} = -0.0082$, s.e. = 0.0013, $p < 0.0001$). This is also true in either male ($\hat{\beta} = -0.0039$, s.e. = 0.0015, $p < 0.01$) or female patients ($\hat{\beta} = -0.0248$, s.e. = 0.0029, $p < 0.0001$), married ($\hat{\beta} = -0.0216$, s.e. = 0.0033, $p < 0.001$) or unmarried patients ($\hat{\beta} = -0.0063$, s.e. = 0.0015, $p < 0.0001$), and patients with more than five prior hospitalizations ($\hat{\beta} = -0.0254$, s.e. = 0.0058, $p < 0.0001$), one to five hospitalizations ($\hat{\beta} = -0.0111$, s.e. = 0.0021, $p < 0.0001$) or no hospitalizations ($\hat{\beta} = -0.0036$, s.e. = 0.0019, $p < 0.05$). The early onset age group again shows the amelioration pattern ($\hat{\beta} = 0.0016$, s.e. = 0.0023, $p > 0.5$), while it is significantly deteriorating in the late onset age group ($\hat{\beta} = -0.0130$, s.e. = 0.0016, $p < 0.0001$).

The grouping effect, i.e. the effect of the time-independent covariates, are not estimated because of the ‘matching’ nature of the proposed model and associated inference procedures. However, similar to the conditional logistic regression models for the matched case-control study, we are able to estimate the interactions between the time-independent and -dependent covariates. For the unadjusted trend measure, the interaction term between gender and the gap indicator is $-0.0095$ (s.e. = 0.0022, $p < 0.0001$), which implies that the deterioration pattern is more severe in the female patients. Similarly, the interaction term between the onset age and gap indicator is 0.0082 (s.e. = 0.0020, $p < 0.0001$), which implies that the earlier onset age group has milder progression pattern that the later group. As to the interaction between the prior history of hospitalization and the gap indicator, it is estimated as $-0.0050$ (s.e. = 0.0014, $p < 0.0004$) and implies that the patients with more hospitalizations in history tend to have more severe deteriorating pattern. But there is no significant difference in deterioration pattern between married and unmarried patients.

For the trend measure adjusted by the age at admission, the interaction terms between the gap indicator and gender, onset age and prior history are $-0.0101$ (s.e. = 0.0022, $p < 0.0001$), 0.0097 (s.e. = 0.0020, $p < 0.0001$) and $-0.0049$ (s.e. = 0.0014, $p < 0.0006$), respectively, while the interaction term between the gap indicator and marital status is not significant. These results shows similarity with those unadjusted as well.

We also select the trend indicator to be $Z_{ij} = \sqrt{J}$. Then the adjusted and unadjusted progression pattern is estimated by $-0.0526$ (s.e. = 0.0062, $p < 0.0001$) and $-0.1891$ (s.e. = 0.0114, $p < 0.0001$), respectively, which implies a deterioration pattern similar to that discovered above. More stratified analyses by gender, onset age, marital status and prior history are conducted and yielded similar results.

5. DISCUSSION

Because of the longitudinal nature, it is well known that the recurrence times as a type of serial multivariate survival times have different statistical structure from those of parallel multivariate survival times, such as collected in the family studies. For example, previous works such as Wei et al. (1989) may be more applicable to the data sets of the latter type, as noted in Pepe and Cai (1993). Other works such as Prentice et al. (1981) and Chang and Wang (1999) focus more on the conditional analysis of recurrence times. More recently, marginal approaches such as in Huang (2000) are also explored to model the recurrence times, although maybe under different contexts. The focus of our article is to model and estimate the longitudinal pattern parameter of the recurrence times, when the study population is considered highly heterogeneous and under censoring.

To accommodate censoring and heterogeneity, this paper utilizes the comparability concept to construct appropriate risksets of gap times as truncated observations. Similar to the usual univariate right-truncated data, the proportional hazards model does not serve as a natural model, but instead, the proportional reverse-time hazards model is proven to be a more proper candidate. The comparability condition for the reduced risksets identified in this paper subsequently fits the model and overcomes all
the heterogeneous baseline distribution functions as nuisance.

However, the comparability condition does have limitations to certain degree. One major limitation is that the complete recurrence times are only considered as ‘comparable’ pairwise. So it is of greater interest but non-trivial to extend to the comparability condition to more than paired recurrence times, which will allow us to gain more efficiency in estimation. In addition, similar to the conditional inference procedures of the fixed-effect logistic regression models for matched case-control studies, the proposed inference procedure does not aim to estimate the subject-specific covariate effects, if the population heterogeneity is related to such subject-specific covariates. A straightforward approach is to use first gap times only, although more effective and more efficient approaches are needed.

ACKNOWLEDGEMENTS

This research was partially supported by National Institute of Health grants MH56639, CA90747 and AI29168. The authors would like to thank Professor Scott Zeger, an Associate Editor and a reviewer for their helpful comments that lead to great improvement of the original manuscript. They would also like to thank Dr. Bill Eaton for his generous assistance with the Danish Registry Data.

APPENDIX

Asymptotics

Martingale theory has been useful in developing asymptotic theory for the inference procedures of the Cox proportional hazards models (Andersen and Gill, 1982). However, martingales are concerned with future events conditioning on the entire history up to the time points at which risksets are constructed. Within the current framework, however, the usual martingale theory is not able to be used in straightforward terms and alternative techniques are applied in developing asymptotic properties in this article. In the following development, without loss of generality, we further assume that \( \beta \) is a scalar. It should not be difficult to extend all the results to the multivariate situation.

The following regularity conditions are assumed:

1. There exist an \( l \in \{1, 2, \ldots, n\} \) and enough big constant \( C_0 > 0 \) such that \( \int_0^{C_0} \kappa_0(s) \, ds < \infty \). In addition, \( \Pr(\sum_{(i,j,k)} \delta_{ijk} > 0) = 1 \).

2. There exists a finite \( M > 0 \) for a neighborhood \( U_0 \) at \( \beta_0 \) such that

\[
\sup_{(i,j), \beta \in U_0} \left[ E[Z_{ij} \exp(\beta' Z_{ij})]\right] < M.
\]

3. There exist \( \Sigma(\beta_0) \) and positive-definite \( D(\beta_0) \) such that

\[
\| \hat{\Sigma}(\beta_0) - \Sigma(\beta_0) \| \to 0
\]

and

\[
\| \hat{D}(\beta_0) - D(\beta_0) \| \to 0,
\]

respectively, where

\[
\hat{\Sigma}(\beta) = \frac{1}{n} \sum_{i=1}^{n} \left\{ \sum_{j} \sum_{k} \delta_{ijk} \sum_{j} \sum_{k} \frac{\delta_{ijk} \exp(\beta Z_{ik}) (Z_{ij} - Z_{ik})}{\exp(\beta Z_{ij}) + \exp(\beta Z_{ik})} \right\} \otimes 2
\]
and
\[
\hat{D}(\beta) = \frac{1}{n} \sum_{i=1}^{n} \frac{1}{\sum_{j} \sum_{k} \delta_{ijk}} \sum_{j} \sum_{k} \frac{\delta_{ijk} \exp(\beta Z_{ij}) \exp(\beta Z_{ik})(Z_{ij} - Z_{ik})^2}{[\exp(\beta Z_{ij}) + \exp(\beta Z_{ik})]^2}.
\]

Here, \(v^{\otimes 0} = 1, v^{\otimes 1} = v\) and \(v^{\otimes 2} = vv^t\), and \(\| \cdot \|\) defines the Euclidean norm.

As shown in Section 3.3, the estimating function \(\tilde{S}(\beta)\) is unbiased. According to the conditions in Foutz (1977) and later used in Pepe and Cai (1993), if the following conditions are satisfied:

1. **Condition F.1.** The partial derivatives of \(\tilde{S}(\beta)\) with respect to \(\beta\) exist and are continuous,
2. **Condition F.2.** The matrix \(n^{-1}(\partial \tilde{S}(\beta_0))\) is non-singular with probability converging to 1 as \(n \to \infty\),
3. **Condition F.3.** The matrix \(n^{-1}(\partial \tilde{S}(\beta))\) converges in probability to the function
\[
A(\beta) = \lim_{n \to \infty} E[n^{-1}(\partial \tilde{S}(\beta))]
\]
uniformly in \(\beta\),
then there exists a neighborhood such that a unique consistent solution to \(\tilde{S}(\beta) = 0\) exist with probability converging to 1.

By the Taylor series expansion, we know that in the neighborhood of \(\beta_0\)
\[
\tilde{S}(\hat{\beta}) - \tilde{S}(\beta_0) = \frac{\partial \tilde{S}(\beta_0)}{\partial \beta} (\hat{\beta} - \beta_0) + \frac{1}{2} \frac{\partial^2 \tilde{S}(\beta^*)}{\partial \beta^2} (\hat{\beta} - \beta_0)^2,
\]
where \(\beta^*\) lies between \(\beta_0\) and \(\hat{\beta}\). Straightforward algebraic manipulation shows that
\[
n^{1/2}(\hat{\beta} - \beta_0) = \left\{ n^{-1} \frac{\partial \tilde{S}(\beta_0)}{\partial \beta} + n^{-1} \frac{1}{2} \frac{\partial^2 \tilde{S}(\beta^*)}{\partial \beta^2} (\hat{\beta} - \beta_0) \right\}^{-1} \left\{ -n^{-1/2} \tilde{S}(\beta_0) \right\}.
\]

By above regularity conditions, \(n^{-1}((\partial^2 / \partial \beta^2) \tilde{S}(\beta))\) is uniformly bounded in the neighborhood of \(\beta_0\). Therefore, \(n^{-1}((\partial^2 / \partial \beta^2) \tilde{S}(\beta^*))\) converges to 0 in probability.

Because of the way of constructing \(\tilde{S}(\beta)\), \(n^{-1}(\partial \tilde{S}(\beta))\) is an average of \(n\) iid random variables with finite variance. Therefore, by the Weak Law of Large Numbers, it converges in probability to \(D(\beta_0) = E\{(\partial / \partial \beta) \tilde{S}_i(\beta_0)\}, i = 1, 2, \ldots, n\). In addition, all the \(\tilde{S}_i(\beta_0)\) are iid zero-mean random variables, so by the central limit theorem, \(n^{-1/2} \tilde{S}(\beta)\) converges in distribution to a normal with mean zero and variance of \(\Sigma(\beta_0)\). Because of the positive-definiteness of \(D(\beta_0)\), it is straightforward to establish the asymptotic normality of \(\hat{\beta}\) as specified in Theorem 1. Using the result in Andersen and Gill (1982) and the consistency of \(\tilde{\beta}\), the consistency of the variance estimators in Theorem 1 is also implied.

**References**


[Received March 3, 2003; revised October 16, 2003; accepted for publication October 30, 2003]