Communications for Statistical Applications and Methods
2013, Vol. 20, No. 4, 283–290

DOI: http://dx.doi.org/10.5351/CSAM.2013.20.4.283

Statistical Analysis of Bivariate Recurrent Event Data with Incomplete Observation Gaps

Yang-Jin Kim[1,a]

[a]Department of Statistics, Sookmyung Women's University

Abstract

Subjects can experience two types of recurrent events in a longitudinal study. In addition, there may exist intermittent dropouts that results in repeated observation gaps during which no recurrent events are observed. Therefore, theses periods are regarded as non-risk status. In this paper, we consider a special case where information on the observation gap is incomplete, that is, the termination time of observation gap is not available while the starting time is known. For a statistical inference, incomplete termination time is incorporated in terms of interval-censored data and estimated with two approaches. A shared frailty effect is also employed for the association between two recurrent events. An EM algorithm is applied to recover unknown termination times as well as frailty effect. We apply the suggested method to young drivers' convictions data with several suspensions.

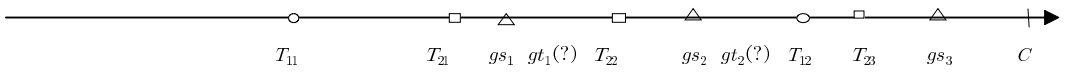
Keywords: Bivariate recurrent event data, frailty effect, observation gap, piecewise constant.

1. Introduction

A bivariate recurrent event arises when two different types of events repeatedly occur. Like other types of multivariate failure times, a frailty effect approach (Cook *et al.*, 2010) and a marginal approach (Cai and Schaubel, 2004) are applied to analyze bivariate recurrent event data. In this paper, we consider a special situation occurring in a longitudinal study where a subject is assumed to be followed until the termination of study. However, some subjects can leave the study and then return to the study at a later time. Such an intermittent missing is denoted as the observation gaps (Zhao and Sun, 2006) and is regarded as a non-risk status at statistical analysis. Furthermore, risk status becomes unclear when such observation gaps result in an incomplete form. As a motivation of this paper, Young Traffic Offence Program(YTOP) data included 40 subjects who experienced at least one suspension and the suspended subjects were banned from driving cars during certain periods. Once the suspension began, subjects dropped out of the risk state and came back to the risk state after completing the suspension. Thus, the observation gap was determined by both start and termination of suspensions and the duration of observation gap depends on the terminating time of suspension. With a complete information about observation gap, Therneau and Hamilton (1997) remarked on discontinuous intervals at risk in recurrent event data and utilized a counting process technique. For similar problems, Duchateau *et al.* (2003) applied a gap time scale to analyze asthma data with the non-risk period. However, the YTOP data provides only the starting time of the suspension while the terminating time is not available. Instead, the terminating time was placed somewhere between the starting time and the recurrent event time occurring subsequently after suspension. To estimate covariate effects on bivariate recurrent event, a frailty is incorporated and an EM algorithm is applied to recover two kinds of incomplete

This Research was supported by the Sookmyung Women's University Research Grants 2012.

[1] Assistant Professor, Department of Statistics, Sookmyung Women's University, Chungpa-Dong, Yonnsan-Gu, Seoul 140-742, Korea. E-mail: yjin@sookmyung.ac.kr



○ : recurrent event of type1
□ : recurrent event of type2
△ : suspension starting time

Figure 1: *Interval censored termination time of suspension with bivariate recurrent events*

data: frailty effects and termination times. The reminder of this paper is organized as follows. Section 2 describes models and notations and Section 3 develops an estimation procedure. In Section 4, the suggested method is applied to the YTOP dataset. Concluding remarks appear in Section 5.

2. Model

Consider n subjects with two different types of recurrent events. Denote T_{1ij} and $T_{2ij'}$ as the j^{th} and j'^{th} recurrent event time of two types of events of subject i, where $i = 1, 2, \ldots, n$, $j = 1, \ldots, n_{1i}$ and $j' = 1, \ldots, n_{2i}$. Define $N_{1i}(t) = \sum_j I(T_{1ij} \le t)$ as a counting process recording the cumulative number of the first event of subject i with $dN_{1i}(t) = 1$ indicating subject i has experienced the first event at time t. Similar definitions are applied to $N_{2i}(t)$ and $dN_{2i}(t)$ for the second recurrent event. Assume that two events follow proportional hazard rates with covariates vectors, $x_{i1}(t)$ and $x_{i2}(t)$, respectively. Given shared frailty ν_i and covariates vectors, the intensity functions are

$$\lambda_{1i}(t|x_i, \nu_i) = Y_i(t)\, \lambda_{01}(t)\exp\left(\beta_1' x_{i1}(t) + \nu_i\right),$$
$$\lambda_{2i}(t|x_i, \nu_i) = Y_i(t)\, \lambda_{02}(t)\exp\left(\beta_2' x_{i2}(t) + \nu_i\right),$$

where $Y_i(t) = I(t \le C_i)$ with a censoring variable C_i indicates at risk process and is a predictable process. $\lambda_{01}(t)$ and $\lambda_{02}(t)$ are arbitrary baseline intensity functions with cumulative functions $\Lambda_{01}(t)$ and $\Lambda_{02}(t)$, respectively. In addition, β_1 and β_2 are regression coefficients of covariate vectors for the first and second recurrent events, respectively. For modeling a correlation between recurrent events, a shared frailty, ν_i is assigned and assumed to be independent and identically distributed with $N(0, \sigma^2)$. In order to incorporate observation gaps, let $G_i = \{g_{i1}, \ldots, g_{im_i}\}$ be the set of suspensions for subject i and every subject has a different suspension number, $(0 \le m_i < \infty)$. Each g_{ik} consists of two time points under suspension, $g_{ik} = [gs_{ik}, gt_{ik}]$, where gs_{ik} and gt_{ik} are the starting time and terminating time of the k^{th} suspension of subject i, respectively. Then the risk indicator is redefined as

$$Y_i^*(t) = I\left(t \le C_i, \text{ and } t \notin g_{ik} = [gs_{ik}, gt_{ik}],\, k = 1, \ldots, m_i\right).$$

An incompleteness of G_i makes it impossible to use Y_i^* directly because of unknown terminating times. Figure 1 shows two types of censoring for termination times of observation gaps. gt_1 and gt_2 are the first and second terminating times and are interval-censored with $[gs_1, T_{22})$ and $[gs_2, T_{12})$, respectively. For the third observation gap, gt_3 is right-censored, $gt_3 \in [gs_3, \infty)$.

For the easy implementation of computation, a piecewise constant baseline rate function is applied. Lawless and Zhan (1998), Liu *et al.* (2004) showed that the pieces defined with appropriately selected cutpoints that perform well at several situations. Let $0 = a_{r0} < a_{r1} < a_{r2} < \cdots < a_{rq_r} = \tau_r$ be the cutpoints of the r^{th} $(r = 1, 2)$ recurrent event, where τ_r is the largest event time of the r^{th} recurrent

event. Then the baseline intensity functions for bivariate recurrent events are redefined as

$$\rho_{10}(t_1;\lambda_{01}) = \sum_{l_1=1}^{q_1} \lambda_{01l_1} I\left(a_{1l_1-1} < t_1 \le a_{1l_1}\right),$$

$$\rho_{20}(t_2;\lambda_{02}) = \sum_{l_2=1}^{q_2} \lambda_{02l_2} I\left(a_{2l_2-1} < t_2 \le a_{2l_2}\right),$$

where $\lambda_{01} = (\lambda_{011},\ldots,\lambda_{01q_1})$ and $\lambda_{02} = (\lambda_{021},\ldots,\lambda_{02q_2})$ are parameters of the baseline rates.

Without observation gaps, the cumulative baseline intensities are defined as

$$\Lambda_{10}\left(t_1;\lambda_{01}\right) = \sum_{l_1=1}^{q_1} \lambda_{01l_1} \max\left(0, \min\left(a_{1l_1} - a_{1l_1-1}, t_1 - a_{1l_1-1}\right)\right),$$

$$\Lambda_{20}\left(t_2;\lambda_{02}\right) = \sum_{l_2=1}^{q_2} \lambda_{02l_2} \max\left(0, \min\left(a_{2l_2} - a_{2l_2-1}, t_2 - a_{2l_2-1}\right)\right).$$

With completely observed observation gaps, the cumulative intensities are adjusted by the duration of observation gaps as follows,

$$\tilde{\Lambda}_{01}(t_1;\lambda_{01}) = \sum_{l_1=1}^{q_1} \lambda_{01l_1} \max\left(0, \min\left(a_{1l_1} - a_{1l_1-1} - \mathrm{dur}_{il_1}^{(1)}, t_1 - a_{1l_1-1} - \mathrm{dur}_{il_1}^{(1)}\right)\right) = \sum_{l_1=1}^{q_1} \lambda_{01l_1} \tilde{s}_{il_1},$$

$$\tilde{\Lambda}_{02}(t_2;\lambda_{02}) = \sum_{l_2=1}^{q_2} \lambda_{02l_2} \max\left(0, \min\left(a_{2l_2} - a_{2l_2-1} - \mathrm{dur}_{il_2}^{(2)}, t_2 - a_{2l_2-1} - \mathrm{dur}_{il_2}^{(2)}\right)\right) = \sum_{l_2=1}^{q_2} \lambda_{02l_2} \tilde{s}_{il_2},$$

where $\mathrm{dur}_{il_k}^{(r)}$ is a duration time by observation gap at the l_k^{th} interval of the $r\,(=1,2)^{th}$ event for the i^{th} subject and is calculated according to the locations of observation gaps $\{(gs_{ik}, gt_{ik}; k = 1,\ldots,m_i\}$ and cutpoints $(a_{rl_r}; l_r = 1,\ldots,q_r)$,

(i) $a_{1l-1} < gs_{ik} < gt_{ik} < a_{1l}$:

$$\mathrm{dur}_{il}^{(1)} = gt_{ik} - gs_{ik}, \tag{2.1}$$

(ii) $a_{1l-1} < gs_{ik} < a_{1l} < gt_{ik} < a_{1l+1}$:

$$\mathrm{dur}_{i,l}^{(1)} = a_{1l} - gs_{ik}, \quad \mathrm{dur}_{i,l+1}^{(1)} = gt_{ik} - a_{1l}, \tag{2.2}$$

(iii) $a_{1l-1} < gs_{ik} < a_{1l} < a_{1l+1} < gt_{ik}$:

$$\mathrm{dur}_{il}^{(1)} = a_{1l} - gs_{ik}, \quad \mathrm{dur}_{i,l+1}^{(1)} = a_{1l+1} - a_{1l}, \quad \mathrm{dur}_{i,l+2}^{(1)} = gt_{ik} - a_{1l+1}. \tag{2.3}$$

Denote $e_{1il}(t_{1ij}) = I(a_{1l-1} < t_{1ij} \le a_{1l})$, $l = 1,\ldots,q$ that the j^{th} occurrence of the first-type of event is observed at the l^{th} interval and $m_{1l} = \sum_{ij} e_{1il}(t_{1ij})$ denote a total number of the first-type of event occurring at the l^{th} interval. For the second-type of event, e_{2il} and m_{2l} are similarly defined. Define $\theta = (\beta_1, \beta_2, \lambda_{01}, \lambda_{02}, \sigma)$. Then given a frailty ν_i, a conditional likelihood is

$$L_c(\theta|\nu) = \prod_{l_1=1}^{q_1} L_{1l_1}(\theta|\nu) \prod_{l_2=1}^{q_2} L_{2l_2}(\theta|\nu), \tag{2.4}$$

where

$$L_{1l_1}(\theta|\nu) = \lambda_{01l_1}^{m_{1l_1}} \prod_{i=1}^{n} \left\{ \exp\left(\sum_{j=1}^{n_{1i}} (x_i(t_{1ij})\beta_1 + \nu_i) e_{1il}(t_{1ij}) - \tilde{\Lambda}_{01l_1}(t_{1ij}) \int_{a_{1l_1-1}}^{a_{1l_1}} \exp\left(x_i'(s)\beta_1 + \nu_i\right) ds \right) \right\},$$

$$L_{2l_2}(\theta|\nu) = \lambda_{02l_2}^{m_{2l_2}} \prod_{i=1}^{n} \left\{ \exp\left(\sum_{j=1}^{n_{2i}} (x_i(t_{2ij})\beta_2 + \nu_i) e_{2il}(t_{2ij}) - \tilde{\Lambda}_{02l_2}(t_{2ij}) \int_{a_{2l_2-1}}^{a_{2l_2}} \exp\left(x_i'(u)\beta_2 + \nu_i\right) du \right) \right\}.$$

However, unknown duration times cause incomplete cumulative intensities, $\tilde{\Lambda}_{01}$ and $\tilde{\Lambda}_{02}$.

3. Estimation

An EM algorithm is applied to recover unknown quantities. In the E-step, a two-stage procedure is applied; The first stage is to estimate unknown terminating times of observation gaps and the complete likelihood is integrated with respect to a frailty effect at the second stage,

Stage1:

(i) *Termination time is independent of covariates*

Assume that termination of observation gap is independent of recurrent events and covariates. For an interval-censored terminating times, $gt_{ik} \in (gs_{ik}, \tilde{t}_{ik}) = (tl_{ik}, tr_{ik})$, A self-consistent algorithm (Turnbull, 1976)is applied to estimate survival distributions of gt_{ik}. Then using estimated $\hat{S}$ and suitably selected time points, $\{w_r, r = 1, \ldots, s\}$, $\hat{f}_r = \hat{S}_{r-1} - \hat{S}_r$ is calculated. Then, $\mathrm{dur}_{il}^{(1)}$s defined at (2.1)–(2.3) for the first recurrent event are redefined as follows,

(case i) $a_{1l_1} < gs_{ik} < \tilde{t}_{ik} < a_{1l_1+1}$:

$$\mathrm{dur}_{il_1}^{(1)} = \frac{\sum_{r=1}^{s} I(gs_{ik} \le w_r < \tilde{t}_{ik})\, (w_r - gs_{ik})\hat{f}_r}{\sum_{r=1}^{s} I(gs_{ik} \le w_r < \tilde{t}_{ik})\hat{f}_r}. \tag{3.1}$$

(case ii) $a_{1l_1} < gs_{ik} < a_{1l_1+1} < \tilde{t}_{ik} < a_{1l_1+2}$:

$$\mathrm{dur}_{il_1}^{(1)} = \frac{\sum_{r=1}^{s} I(gs_{ik} \le w_r < a_{1l_1+1})(w_r - gs_{ik})\hat{f}_r}{\sum_{r=1}^{s} I(gs_{ik} \le w_r < a_{1l_1+1})\hat{f}_r}, \tag{3.2}$$

$$\mathrm{dur}_{il_1+1}^{(1)} = \sum_{k=1}^{m_i} \frac{\sum_{r=1}^{s} I(a_{1l_1+1} \le w_r < \tilde{t}_{ik})(w_r - a_{1l_1+1})\hat{f}_r}{\sum_{r=1}^{s} I(a_{1l_1+1} \le w_r < \tilde{t}_{ik})\hat{f}_r}. \tag{3.3}$$

(case iii) $a_{1l_1} < gs_{ik} < a_{1l_1+1} < a_{1l_1+2} < \tilde{t}_{ik}$:

$$\mathrm{dur}_{il_1}^{(1)} : \text{same with (case ii).}$$

$$\mathrm{dur}_{il_1+1}^{(1)} = \frac{\sum_{r=1}^{s} I(a_{1l_1+1} \le w_r < a_{l_1+2})(w_r - a_{1l_1+1})\hat{f}_r}{\sum_{r=1}^{s} I(a_{1l_1+1} \le w_r < a_{l_1+2})\hat{f}_r}, \tag{3.4}$$

$$\mathrm{dur}_{il_1+2}^{(1)} = \frac{\sum_{r=1}^{s} I(a_{1l_1+2} \le w_r < \tilde{t}_{ik})(w_r - a_{1l_1+2})\hat{f}_r}{\sum_{r=1}^{s} I(a_{1l_1+2} \le w_r < \tilde{t}_{ik})\hat{f}_r}. \tag{3.5}$$

(ii) *Terminating times are related with covariates*

For modeling the relation between terminating times and covariates, a following proportional hazard model is applied,

$$\gamma(gt_{ik};\eta) = \gamma_0(gt_{ik})\exp\left(\eta' z_{ik}\right). \tag{3.6}$$

Several approaches are considered to estimate regression coefficients of model (6) for interval censored data (Pan, 2000; Finkelstein, 1986). In this study, a counting process approach by Goetghebeur and Ryan (2000) is adopted to estimate a probability mass function,

$$\hat{f}_{ikr} = \hat{f}_r(z_{ik};\eta,\gamma) = \frac{p_{ikr}}{\sum_{tl_{ik}\le w_l\le tr_{ik}} p_{ikl}}, \quad tl_{ik} \le w_r \le tr_{ik}, \tag{3.7}$$

where $p_{ikl} = \gamma_l \exp[\eta' z_{ik} - \exp(\eta' z_{ik}) \sum_{r=1}^{l} \gamma_r]$.

To estimate $(\eta, \gamma = (\gamma_1, \ldots, \gamma_s))$, a following log pseudo-likelihood is considered (Goetghebeur and Ryan, 2000),

$$l_g = \log L_g = \sum_{i=1}^{n}\sum_{k=1}^{m_i}\sum_{l=1}^{s}\left\{\log(\gamma_l)dN_{ikl} + \eta' z_{ik} dN_{ikl} - \gamma_l \exp\left(\eta' z_{ik}\right)\tilde{Y}_{ikl}\right\}$$

a Newton-Raphson algorithm is adopted to estimate $\hat{\eta}, \hat{\gamma}$ which update $\hat{f}_{ikr}$ and (3.1)–(3.5).

Stage 2: Integration a conditional likelihood (2.4) with respect to a frailty effect includes the calculation of conditional expectations of the functions of frailties,

$$E\{h(\nu_i|O_i,\theta)\} = E^*[h(\nu_i)] = \int h(\nu_i)L_i(\theta|\nu_i)g\left(\nu_i|\sigma^2\right)d\nu_i,$$

where $g \sim N(0,\sigma^2)$. Since this integration has no closed form, a 13-point Gauss-Hermite integration is applied. All computations can be performed using *Proc NLMIXED* and *%macro EMICM* in SAS which is implemented to estimate a survival distribution of interval censored termination time.

Once completing E-step, the vector of parameters is updated by applying a Newton-Raphson algorithm at the M-step using the score function and hessian matrix derived from likelihood. The standard error is estimated by the observed fisher information updated with final estimates of θ.

4. Data Analysis

In this section, we applied the suggested method to the YTOP dataset (Sun *et al.*, 2001). The program was a 1-day educational intervention for adolescents and young people. The original data set includes the record of several other causes of conviction. For example, convictions can happen at a subject with traffic signal violations and alcohol related violations as well as speed violation. In this paper, (i) speed rule violation and (ii) other traffic rule violations(including traffic signal and alcohol related violation) are regarded as a bivariate recurrent event data. A total of 441 convictions were related with speed rule violations and 165 ones resulted from other traffic rule violation. Also, among 193 young drivers, 40 subjects experienced at least one suspension and the maximum number of suspensions was seven. Kim and Jhun (2008) did not distinguish types of conviction and also considered only the first suspension. Table 1 shows the frequencies and mean of convictions by gender groups. The female

Table 1: Conviction records by gender

	Speed related conviction (441)		Other-rule related conviction (165)	
	Male	Female	Male	Female
Frequency	326	115	144	21
Mean	2.345	2.129	1.036	0.389

Table 2: YTOP data analysis

	Model 0	Model 1	Model 2
		Speed rule violation	
YTOP	−1.923(0.176)	−1.509(0.185)	−1.437(0.193)
Gender	0.012(0.186)	−0.054(0.153)	−0.034(0.143)
		Other rule Traffic rule violation	
YTOP	−1.932(0.237)	−1.297(0.239)	−1.242(0.237)
Gender	0.648(0.285)	0.798(0.279)	0.674(0.262)
$\hat{\sigma}$	0.739(0.164)	0.346(0.127)	0.501(0.121)

group had a smaller mean at the second type of conviction while both genders have similar mean conviction number at the speed rule violation. The cutpoints were determined according to quantiles of recurrent event times and ten pieces are used.

The effects of two covariates(YTOP: participant group = 1, non-participant group = 0; gender: Male(= 1) and Female(= 0))are investigated with adjusting correlation using a shared frailty effect. Table 2 contains the results of three different models. The observation gap was ignored at the Model 0,and Model 1 estimated a termination time with nonparametric method under covariate independence assumption. A semiparametric model (3.6) is incorporated for a terminating time at Model 2 and estimated coefficients are ($\hat{\eta} = (-0.515, -0.706)$) which means YTOP participants and male drivers have long duration time of observation gap. Comparing three models, all coefficients are almost the same: however, Model 0 showed different sign for gender effect at speed rule convictions.

Based on the results, YTOP group has significant effect on both types of conviction. Program participants have less events than the non-participant group for both types of recurrent. However, two gender groups have no significant difference on speed conviction while the male group has more other traffic rule convictions than the female group. Standard deviations of frailty effect in the three models show significant results which means subjects have diverse patterns.

Figure 3 showed the estimated intensity functions and the cumulative baseline intensity functions of two type of recurrent events with $h = 10$ under Model 2. For speed convictions, cutpoints are ($a_{1,0} = 40, 368, 630, 875, 1091, 1256, 1514, 1797, 2008, 2283, 3504 = a_{1,10}$) and the eighth interval(at about 5-7 years since the driver got the license)has the highest intensity and the intensity abruptly decrease after that interval. For the second type of conviction, cutpoints are ($a_{2,0} = 23, 272, 693, 889.5, 1147, 1365.5, 1597.5, 1768, 2110, 2634, 38026 = a_{2,10}$) and their intensities have very similar values over all pieces.

5. Discussions

This paper considers a bivariate recurrent event data with incomplete observation gaps. To estimate unknown duration times of observation gaps, interval censored data is incorporated and a self consistent algorithm and pseudo likelihood is applied to estimate the distribution of terminating times.

In this study, there are some assumptions related with the observation gap process and recurrent events process. Occurrence times and terminating times are assumed to be independent with recurrent events. A multi-state model approach would be considered to implement these process simultaneously,(Foucher *et al.*, 2007; Cook *et al.*, 2008). A shared frailty effect is incorporated for

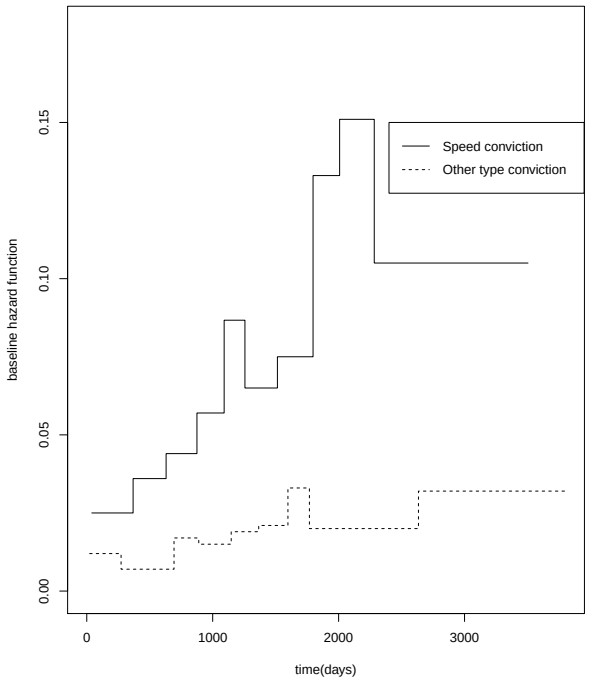


Figure 2: *Estimated baseline hazard functions for 10 pieces*

the correlation between two recurrent events. A bivariate frailty would be applied for event-specific variations in future research.

References

Cai, J. and Schaubel, D. E. (2004). Marginal means/rates models for multiple type recurrent event types, *Lifetime Data Analysis*, **10**, 121–138.

Cook, R. J., Lawless, J. F. and Lee, K. A. (2010). A copula-based mixed Poisson model for bivariate recurrent events under event-dependent censoring, *Statistics in Medicine*, **29**, 694–707.

Cook, R., Zeng, L. and Lee, K. (2008). A multistate model for bivariate interval-censored failure time data, *Biometrics*, **64**, 1100–1109.

Duchateau, L., Jassen, P., Kezic, I. and Fortpied, C. (2003). Evolution of recurrent asthma event rate over time in frailty models, *Journal of the Royal Statistical Society, Series C (Applied Statistics)*, **52**, 355–363.

Finkelstein, D. M. (1986). A proportional hazards model for interval-censored failure time data, *Biometrics*, **42**, 845–854.

Foucher, Y., Giral, M., Soulillou, J.-F. and Daures, J.-P. (2007). A semi-Markov model for multistate and interval-censored data with multiple terminal events. Application in renal transplantation, *Statistics in Medicine*, **26**, 5381–5393.

Goetghebeur, E. and Ryan, L. (2000). Semiparametric regression analysis of interval-censored data, *Biometrics*, **56**, 1139–1144.

Kim, Y. and Jhun, M. (2008). Analysis of recurrent event data with incomplete observation gaps, *Statistics in Medicine*, **27**, 1075–1085.

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

Lawless, J. F. and Zhan, M. (1998). Analysis of interval-grouped recurrent event data using piecewise constant rate functions, *Canadian Journal of Statistics*, **26**, 549–565.

Lindsey, J. and Ryan, L. (1998). Methods for interval censored data. Tutorial in biostatistics, *Statistics in Medicine*, **17**, 219–138.

Liu, L., Wolfe, R. A. and Huang, X. (2004). Shared frailty models for recurrent events and a terminal event, *Biometrics*, **60**, 747–756.

Pan, W. (2000). Multiple imputation approach to Cox regression with interval censored data, *Biometrics*, **56**, 199–203.

Sun, J., Kim, Y., Hewett, J., Johnson, J. C., Farmer, J. and Gibler, M. (2001). Evaluation of traffic injury prevention programs using counting process approaches, *Statistics in Medicine*, **96**, 469–475.

Therneau, T. M. and Hamilton, S. C. (1997). rhDNase as an example of recurrent event analysis, *Statistics in Medicine*, **16**, 2029–2047.

Turnbull, B. W. (1976). The empirical distribution function with arbitrarily grouped censored and truncated data, *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, **38**, 290–295.

Zhao, Q. and Sun, J. (2006). Semiparametric and nonparametric estimation of recurrent event with observation gaps, *Computational Statistics & Data Analysis*, **51**, 1924–1933.

Received March 5, 2013; Revised May 8, 2013; Accepted June 13, 2013