STATISTICA, anno LXXXIII, n. 1, 2023

NONPARAMETRIC ESTIMATION OF CUMULATIVE INCIDENCE FUNCTIONS OF RECURRENT EVENTS

Sisuma Mandakathingal Sivadasan [1]
Department of Statistics, Cochin University of Science and Technology, Kochi-22, Kerala
Paduthol Godan Sankaran
Department of Statistics, Cochin University of Science and Technology, Kochi-22, Kerala

Summary

The present paper discusses modeling and analysis of recurrent event data with competing risks. We propose non-parametric estimation of cumulative incidence functions of recurrent event competing risks model. Asymptotic properties of the proposed estimators are established. Simulation procedures are carried out to assess the finite sample properties of the proposed estimators. The proposed method is applied to real-life data.

Keywords: Recurrent event; Competing risks; Cause-specific hazard function; Cumulative incidence function; Empirical process.

1. Introduction

In biomedical studies, subjects or individuals can experience the event of interest more than once. Such events are termed recurrent events. The recurrent event data arise in diverse fields such as public health, medicine, insurance, social science, economics, manufacturing and reliability. Recurrent hospitalisation of patients with chronic diseases, episodes of hypoglycemia in diabetics, the breakdown of mechanical or electronic systems, computer software crashes, stoppages of nuclear power plants, warranty claims for manufactured products, serious disagreements in marriage, the onset of labour strikes and auto insurance claims are some examples of recurrent phenomena. There are several statistical procedures proposed in the survival analysis literature for analysing the single type of recurrent events. Two types of time scales, viz. time since entering the study and time since the last event (gap time), are commonly considered in the literature. Conditional and marginal models have been proposed to analyse data with a single type of

[1] Corresponding Author. E-mail: sisumasivadasan@gmal.com

recurrent event. One can refer to Prentice *et al.* (1981) for conditional models and Wei *et al.* (1989) for marginal models.

The standard method for the analysis of recurrent event data is focused either on the mean function or the rate function of the underlying recurrent event process. The non-parametric estimation of mean function and rate function is extensively studied by different researchers. One may refer to Cook and Lawless (2007) for a comprehensive review on this topic. Pepe and Cai (1993) proposed semi parametric procedure for the analysis of recurrent event data using the mean function. Lawless and Nadeau (1995) have proposed a class of marginal models for recurrent event data, and they focused on the mean number of events experienced by a person up to the time considered. This approach enables a comparison between treatments via robust tests (Cook *et al.*, 1996). Wang and Chang (1999) proposed the non-parametric estimation of univariate recurrent survival function using the independence assumption on gap times. Semi-parametric inferences have been proposed by Lin *et al.* (1998) and Ghosh (2004) for accelerated failure time models and Lin *et al.* (2000), Maitra *et al.* (2020) for multiplicative rate or mean models. Zeng and Cai (2010) and Stocker and Adekpedjou (2020) have developed a semi-parametric additive model for recurrent event data.

A comprehensive review of models and methods for analysing recurrent event data is given by Geffray (2013). Non-parametric estimation of intensity function of recurrent event data is developed by Bouaziz *et al.* (2013). Che and Angus (2016), Qu *et al.* (2017) and Han *et al.* (2018) have developed joint models of recurrent events and terminal event based on the proportional intensity model and additive hazard models. Andersen *et al.* (2019) focused on the marginal analysis of the expected number of events of recurrent events with the terminal event.

Competing risks data emerge naturally in medical research when subjects under study are at risk of more than one cause. Two frameworks, viz. cause-specific hazards and cumulative incidence functions, are commonly employed for the analysis of such competing risks data. The cause-specific hazards and cumulative incidence curves capture different aspects of event histories in competing risks data, and inference on these measures may yield different results. Crowder (2001) and Lawless (2011) have provided a comprehensive review on this topic.

The recurrent events with competing risks set up arise in many medical studies. The development of new stochastic models for the analysis of recurrent event data with competing risks is a topic of interest. Dauxois and Sencey (2009) proposed a non-parametric test for mean-specific functions in the contexts of recurrent events with competing risks. Sankaran and Anisha (2011) and Sankaran and Anisha (2012) discussed semi-parametric inference for gap time distributions of recurrent event data with multiple causes. Taylor and Pena (2014) developed non-parametric estimators of component and system life distributions of recurrent competing risks data. Dong *et al.* (2015) have used the method of the mean cumulative count for estimating the burden of recurrent events in the presence of competing risks. Gouskova *et al.* (2017) studied the non-parametric analysis of such data type with the missing event type. Ning *et al.* (2017) derived semi-parametric rate models for multivariate recurrent event data to estimate blood product ratios. Ma *et al.*

(2018) developed the accelerated failure time for multivariate recurrent event data with the missing event type. The association between multiple recurrent events with multivariate modelling studied by Osmani *et al.* (2018). Yu *et al.* (2018) proposed the non-parametric correction approaches to simultaneously correct for the informative censoring and measurement errors in the analysis of multivariate recurrent event data.

Most of the models developed earlier were based either on the mean number of events or the gap time between successive events. However, one of the basic questions for the analysis of recurrent competing risk data is to evaluate the occurrence rate of the recurrent event of interest. Motivated by this, we propose an alternate approach by introducing recurrent cumulative incidence functions. We propose non-parametric estimators of recurrent cumulative incidence functions.

The paper is organized as follows. Section 2 presents the data and model. Section 3 derives recurrent non-parametric estimators of recurrent cumulative incidence functions (RCIF). The asymptotic properties of the proposed estimators are given in Section 4. In Section 5, we present the simulation results to assess the efficiency of the proposed estimator. We illustrate the proposed methods with kidney catheter data in Section 6. Finally, Section 7 presents the major conclusions of the study.

2. The data and model

Suppose there are n independent individuals having a sequence of events experiencing k competing risks to be observed in a longitudinal study. Let T_{ij}, $j = 1,2,\ldots,m_i$ and $i = 1,2,\ldots,n$ denote the time from $(j-1)^{\text{th}}$ to the j^{th} event for the i^{th} individual. Let C_i be the censoring time of i^{th} individual, $i = 1,2,\ldots,n$. Let J_{ij} be the cause of failure for jth recurrence of i^{th} individual, $j = 1,2,\ldots,m_i, i = 1,2,\ldots,n$.

The proposed model requires important assumptions : There exists a baseline random variable Z_i for each individual, which may be unobservable or partially observable, $i = 1,2,\ldots,n$. The recurrence times $T_{i1}, T_{i2},\ldots,T_{im_i}$ are independent and identically distributed given Z_i and for each individual one could observe cause of failure along with $T_{ij}, j = 1,2,\ldots,m_i, i = 1,2,\ldots,n$.

The frailty assumption ensures the correlation of recurrence time from the same subject. This assumption is stronger than necessary in our proposed method.

Let m_{il} be the number of recurrent events that occurred due to the l^{th} competing cause for the i^{th} individual, $l = 1,2,\ldots,k, i = 1,2,\ldots,n$. Let $m_i = \sum_{l=1}^{k} m_{il}$ be the number of recurrent events experienced by the i^{th} individual. Note that, $i = 1,2,\ldots,n$,

$$\sum_{j=1}^{m_i-1} T_{ij} \leq C_i, \tag{1}$$

$$\sum_{j=1}^{m_i} T_{ij} > C_i. \tag{2}$$

Clearly both m_i and m_{il} are random variables.

We define the number of recurrence m_i^* as

$$m_i^* = \begin{cases} 1, & \text{if } m_i = 1, \\ m_i - 1, & \text{if } m_i \geq 2. \end{cases} \tag{3}$$

The observed recurrence times are given by,

$$y_{ij} = \begin{cases} t_{ij}, & \text{if } j = 1, 2, .., m_i - 1, \\ t^+_{i,m_i}, & \text{if } j = m_i. \end{cases}$$

where t^+_{i,m_i} is the censored recurrence time. It is the time from event $m_i - 1$ to the end of the follow-up.

Let $F_Z(.)$ denote the distribution function of the baseline random vector Z_i. Let $f(t|z_i)$ be the conditional density of T_{ij} given z_i. The recurrent survival function of T_{ij} is expressed as

$$S(t) = P(T_{ij} > t) = \int \int_t^\infty f(u|z) d(u) dF_Z(z), \tag{4}$$

In a competing risks framework, each individual is exposed to k distinct types of risks, and then the cause-specific hazard function for an individual is defined by

$$\lambda_l(t) = \lim_{\Delta t \to 0} \frac{1}{\Delta t} P(t \leq T_{ij} < t + \Delta t, J_{ij} = l | T_{ij} \geq t), \quad l = 1, 2, \ldots, k \tag{5}$$

The cumulative incidence function of T_{ij} is given by

$$F_l(t) = P(T_{ij} \leq t, J_{ij} = l), \quad l = 1, 2, \ldots, k. \tag{6}$$

Our objective is to derive non-parametric estimators of recurrent cumulative incidence functions $F_l(t), l = 1, 2, \ldots, k$.

3. NON-PARAMETRIC ESTIMATION OF RECURRENT CUMULATIVE INCIDENCE FUNCTION.

Consider $a_i = a(C_i)$ a non negative function of C_i subject to the constraint $E(a_i^2) < \infty$. To estimate $F_l(t)$, we consider the functions given in Wang and Chang (1999),

$$H_a(t) = E[a_i I(T_{i1} \geq t) I(C_i \geq t)], \tag{7}$$

and

$$G_a(t) = E[a_i I(T_{i1} \leq t) I(T_{i1} \leq C_i)], \tag{8}$$

where $I(.)$ represents indicator function. We now define functions

$$G_{al}(t) = E[a_i I(T_{i1} \leq t, J_{ij} = l) I(T_{i1} \leq C_i)], \quad , l = 1, 2, \ldots, k. \tag{9}$$

The cumulative hazard function, given in Wang and Chang (1999), is

$$\Lambda(t) = \int_0^t \frac{E[a_i I(C_i \geq u)]d\{1 - S(u)\}}{E[a_i I(C_i \geq u)]S(u)} = \int_0^t \frac{dG_a(u)}{H_a(u)}. \tag{10}$$

The cause-specific cumulative hazard function is defined by

$$\Lambda_l(t) = \int_0^t \frac{E[a_i I(C_i \geq u, J_{ij} = l)]d\{F_l(u)\}}{E[a_i I(C_i \geq u)]S(u)} = \int_0^t \frac{dG_{al}(u)}{H_a(u)}, \quad l = 1, 2, \ldots, k. \tag{11}$$

Equations (10) and (11) have been useful in the development of asymptotic properties of the proposed estimators. Estimators of $H_a(t)$ and $G_a(t)$ are given by

$$\hat{H}_a(t) = n^{-1} R^*(t) = n^{-1} \sum_{i=1}^{n} \left[\frac{a_i}{m_i^*} \sum_{j=1}^{m_i^*} I(y_{ij} \geq t) \right] \tag{12}$$

and

$$\hat{G}_a(t) = n^{-1} \sum_{i=1}^{n} \left[\frac{a_i I(m_i \geq 2)}{m_i^*} \sum_{j=1}^{m_i^*} I(y_{ij} \leq t) \right], \tag{13}$$

where $R^*(t)$ is the total mass of risk set at t given by

$$R^*(t) = \sum_{i=1}^{n} \left[\frac{a_i}{m_i^*} \sum_{j=1}^{m^*} I(y_{ij} \geq t) \right]. \tag{14}$$

Non-parametric estimator of $G_{al}(t)$ is given by

$$\hat{G}_{al}(t) = n^{-1} \sum_{i=1}^{n} \left[\frac{a_i I(m_{il} \geq 2)}{m_{il}} \sum_{j=1}^{m_{il}} I(y_{ij} \leq t, J_{ij} = l) \right], \quad l = 1, 2, \ldots, k. \tag{15}$$

Note that $\hat{H}_a(t)$, $\hat{G}_a(t)$ and $\hat{G}_{al}(t)$ are moment type estimators of $H_a(t)$, $G_a(t)$ and $G_{al}(t)$ respectively. The estimators of the cause-specific cumulative hazard functions $\Lambda_l(t)$ is obtained as

$$\hat{\Lambda}_l(t) = \int_0^t \frac{d\hat{G}_{al}(u)}{\hat{H}_a(u)}, \quad l = 1, 2, \ldots, k. \tag{16}$$

From Wang and Chang (1999), the estimator of the cumulative hazard function $\Lambda(t)$ is

$$\hat{\Lambda}(t) = \int_0^t \frac{d\hat{G}_a(u)}{\hat{H}_a(u)}. \tag{17}$$

Then, the estimator of the survival function $\hat{S}(t)$ can be derived from the relation $S(t) = \exp\{-\Lambda(t)\}$.

The recurrent cumulative incidence function can be estimated by the non-parametric plug-in estimator as

$$\hat{F}_l(t) = \int_0^t \hat{S}(u) d\hat{\Lambda}_l(u), \quad l = 1, 2, \ldots, k. \tag{18}$$

Note that $\hat{H}_a(t)$, $\hat{G}_a(t)$ and $\hat{G}_{al}(t)$ are unbiased estimators of $H_a(t)$, $G_a(t)$ and $G_{al}(t)$, respectively.

4. Asymptotic properties of estimators

We use an empirical process approach for the development of asymptotic properties of estimators. Theorem 1 gives the asymptotic normality of proposed estimators. Let t^* be a non negative constant satisfying $t^* < \sup\{t : S(t)G(t) > 0\}$.

THEOREM 1. *Assume that a_i is a bounded function. As $n \to \infty$, the random process $\sqrt{n}\{\hat{F}_l(t) - F_l(t)\}$ for $0 < t < t^*$ converges weakly to a zero mean Gaussian process with variance-covariance function $\sigma_l(t_1, t_2) = E[\epsilon_{li}(t_1)\epsilon_{li}(t_2)], l = 1, 2, \ldots, k$.*

PROOF. See the Appendix. □

REMARK 2. *Since the analytic expression for the variance-covariance function is not in a closed form, we use the bootstrap procedure for estimating the variance-covariance function using the re-sampling method.*

THEOREM 3. *Let $T_{i1}, T_{i2}, .., T_{ij}$ be independent and identically distributed random variables given Z_i. Denote $\hat{F}_l(t)$ be the non-parametric estimator $F_l(t)$ for $0 < t < t^*$. Then $\hat{F}_l(t)$ is strongly consistent for $F_l(t)$.*

PROOF. See the Appendix. □

REMARK 4. *Selection of a_i with minimum asymptotic variance of $\hat{F}_l(t)$ is always a topic of interest. The optimal weight, however, does not have a closed analytical expression and could vary at different recurrent time points. The bootstrap procedure is employed to obtain the optimal weight.*

5. Simulation Studies

We carry out an extensive simulation study to evaluate the performance of the proposed estimators.

The following algorithm is used for simulating data with recurrent events.

1. Generate frailty values $(z_i, i = 1, 2, \ldots, n)$ from the exponential distribution with parameter one.
2. The i.i.d. recurrence times for the given z_i are generated from the cause specific Weibull hazard function $\lambda_l(t|z_i)$, where $\lambda_l(t|z_i) = z_i \lambda_l \alpha_l t^{\alpha_l - 1}$ with parameters $\lambda_l, \alpha_l \geq 0, l = 1, 2, .., k$.
3. For simplicity, we take the number of causes as two.
4. We choose the parameters $\lambda_1 = 0.1, \lambda_2 = 0.12$ and $\alpha_1 = \alpha_2 = \alpha = 2$.
5. Two causes are chosen using a binomial experiment with probability of success $\frac{\lambda_l(t)}{\lambda_1(t)+\lambda_2(t)}$, $l = 1, 2$.

The recurrence times with competing risks data are generated by the method given in Beyersmann *et al.* (2009). The simulation is carried out with sample of sizes 50, 100, 200 and 500. The observation of the recurrence time process is terminated by the censoring time $C_i, i = 1, 2, \ldots, n$. We consider both random censoring and fixed censoring. In the case of random censoring, censoring time C_i are generated from U(0,3) and U(0,5) distributions, $i = 1, 2, \ldots, n$. In the case of fixed censoring, we consider $C_i = 3$ and $C_i = 5$ for $i = 1, 2, \ldots, n$. To study the effect of weight function $a(C_i)$, we consider the weight function $a(C_i) = 1$ and $a(C_i) = C_i$, where $C_i \sim U(0, 3)$ and $U(0, 5)$ distributions. We calculate the estimator of $F_l(t)$ from Eq. (18) and calculate absolute bias (Abs.bias) and mean squared errors (MSE) of the estimators based on 1000 replications.

The absolute bias and mean squared error of estimators with $a(C_i) = C_i \sim U(0, 3)$ and $a(C_i) = C_i \sim U(0, 5)$ are presented in Table 1 and Table 2, respectively. Similarly, absolute bias and mean squared error of estimators with $a(C_i) = 1$ in the case of random censoring are presented in Table 3 and Table 4, respectively. Simulation results for fixed censoring times $C_i = 3$ and $C_i = 5$ are presented in Table 5 and Table 6, respectively.

Mean squared errors of the estimators decrease as the sample size increases. The absolute bias of the estimators also decreases as the sample size increases.

TABLE 1
Estimates of recurrent cumulative incidence functions with $a(C_i) = C_i \sim U(0,3)$.

n	t	Abs.bias $\hat{F}_1(t)$	MSE $\hat{F}_1(t)$	Abs.bias $\hat{F}_2(t)$	MSE $\hat{F}_2(t)$
50	0.2	1.50×10^{-2}	5.20×10^{-4}	1.76×10^{-2}	5.60×10^{-4}
	0.4	1.54×10^{-2}	7.90×10^{-4}	2.01×10^{-2}	9.00×10^{-4}
	0.6	6.64×10^{-3}	8.50×10^{-4}	1.05×10^{-2}	8.10×10^{-4}
	0.8	8.14×10^{-3}	1.11×10^{-3}	6.72×10^{-3}	9.80×10^{-4}
	1	2.84×10^{-2}	1.99×10^{-3}	2.98×10^{-2}	2.04×10^{-3}
	1.2	5.04×10^{-2}	3.94×10^{-3}	5.60×10^{-2}	4.49×10^{-3}
	1.4	7.40×10^{-2}	7.08×10^{-3}	8.36×10^{-2}	8.52×10^{-3}
	1.6	9.73×10^{-2}	1.13×10^{-2}	1.11×10^{-1}	1.41×10^{-2}
100	0.2	1.43×10^{-2}	3.20×10^{-4}	1.68×10^{-2}	4.10×10^{-4}
	0.4	1.49×10^{-2}	4.10×10^{-4}	1.78×10^{-2}	5.80×10^{-4}
	0.6	5.92×10^{-3}	3.10×10^{-4}	7.60×10^{-3}	4.20×10^{-4}
	0.8	7.86×10^{-3}	4.40×10^{-4}	6.57×10^{-3}	4.00×10^{-4}
	1	2.81×10^{-2}	1.30×10^{-3}	2.74×10^{-2}	1.16×10^{-3}
	1.2	4.61×10^{-2}	2.62×10^{-3}	5.01×10^{-2}	2.95×10^{-3}
	1.4	6.03×10^{-2}	4.04×10^{-3}	7.45×10^{-2}	5.98×10^{-3}
	1.6	7.03×10^{-2}	5.26×10^{-3}	1.08×10^{-1}	1.22×10^{-2}
200	0.2	9.50×10^{-3}	1.10×10^{-4}	1.26×10^{-2}	1.80×10^{-4}
	0.4	1.19×10^{-2}	2.00×10^{-4}	1.15×10^{-2}	1.70×10^{-4}
	0.6	5.31×10^{-3}	1.20×10^{-4}	6.48×10^{-3}	1.70×10^{-4}
	0.8	7.02×10^{-3}	2.70×10^{-4}	5.09×10^{-3}	1.90×10^{-4}
	1	2.70×10^{-2}	9.80×10^{-4}	2.31×10^{-2}	7.10×10^{-4}
	1.2	3.76×10^{-2}	1.54×10^{-3}	4.49×10^{-2}	2.19×10^{-3}
	1.4	5.20×10^{-2}	2.80×10^{-3}	6.11×10^{-2}	3.87×10^{-3}
	1.6	6.80×10^{-2}	4.69×10^{-3}	9.75×10^{-2}	9.70×10^{-3}
500	0.2	7.77×10^{-3}	6.00×10^{-5}	1.03×10^{-2}	1.10×10^{-4}
	0.4	9.48×10^{-3}	1.00×10^{-4}	7.70×10^{-3}	7.00×10^{-5}
	0.6	4.51×10^{-3}	5.00×10^{-5}	5.82×10^{-3}	6.00×10^{-5}
	0.8	6.62×10^{-3}	1.30×10^{-4}	5.00×10^{-3}	1.00×10^{-4}
	1	1.82×10^{-2}	3.70×10^{-4}	1.99×10^{-2}	4.40×10^{-4}
	1.2	3.31×10^{-2}	1.13×10^{-3}	3.97×10^{-2}	1.61×10^{-3}
	1.4	4.73×10^{-2}	2.25×10^{-3}	5.98×10^{-2}	3.59×10^{-3}
	1.6	6.48×10^{-2}	4.21×10^{-3}	8.97×10^{-2}	8.06×10^{-3}

TABLE 2
Estimates of recurrent cumulative incidence functions
with $a(C_i) = C_i \sim U(0,5)$.

n	t	Abs.bias $\hat{F}_1(t)$	MSE $\hat{F}_1(t)$	Abs.bias $\hat{F}_2(t)$	MSE $\hat{F}_2(t)$
50	0.2	5.16×10^{-2}	3.24×10^{-3}	5.56×10^{-2}	3.74×10^{-3}
	0.4	7.33×10^{-2}	6.45×10^{-3}	8.08×10^{-2}	7.75×10^{-3}
	0.6	7.84×10^{-2}	7.70×10^{-3}	8.63×10^{-2}	9.10×10^{-3}
	0.8	7.29×10^{-2}	7.23×10^{-3}	7.96×10^{-2}	8.30×10^{-3}
	1	6.12×10^{-2}	5.93×10^{-3}	6.49×10^{-2}	6.48×10^{-3}
	1.2	4.44×10^{-2}	4.40×10^{-3}	4.65×10^{-2}	4.72×10^{-3}
	1.4	2.57×10^{-2}	3.32×10^{-3}	2.42×10^{-2}	3.35×10^{-3}
	1.6	5.76×10^{-3}	2.83×10^{-3}	8.50×10^{-4}	3.02×10^{-3}
100	0.2	4.60×10^{-2}	2.28×10^{-3}	5.13×10^{-2}	2.78×10^{-3}
	0.4	6.70×10^{-2}	4.82×10^{-3}	6.87×10^{-2}	4.96×10^{-3}
	0.6	6.19×10^{-2}	4.15×10^{-3}	7.74×10^{-2}	6.38×10^{-3}
	0.8	6.16×10^{-2}	4.27×10^{-3}	6.92×10^{-2}	5.23×10^{-3}
	1	5.82×10^{-2}	4.23×10^{-3}	5.72×10^{-2}	3.91×10^{-3}
	1.2	4.00×10^{-2}	2.45×10^{-3}	4.41×10^{-2}	2.90×10^{-3}
	1.4	2.49×10^{-2}	1.71×10^{-3}	2.11×10^{-2}	1.49×10^{-3}
	1.6	4.15×10^{-3}	1.07×10^{-3}	4.40×10^{-4}	1.33×10^{-3}
200	0.2	4.47×10^{-2}	2.05×10^{-3}	4.91×10^{-2}	2.47×10^{-3}
	0.4	5.89×10^{-2}	3.53×10^{-3}	6.78×10^{-2}	4.67×10^{-3}
	0.6	5.67×10^{-2}	3.27×10^{-3}	6.99×10^{-2}	4.98×10^{-3}
	0.8	5.17×10^{-2}	2.75×10^{-3}	6.17×10^{-2}	3.92×10^{-3}
	1	3.61×10^{-2}	1.40×10^{-3}	5.56×10^{-2}	3.31×10^{-3}
	1.2	2.91×10^{-2}	1.04×10^{-3}	4.03×10^{-2}	1.96×10^{-3}
	1.4	1.64×10^{-2}	5.70×10^{-4}	1.89×10^{-2}	7.60×10^{-4}
	1.6	3.25×10^{-3}	4.60×10^{-4}	6.00×10^{-5}	5.70×10^{-4}
500	0.2	3.42×10^{-2}	1.18×10^{-3}	4.33×10^{-2}	1.88×10^{-3}
	0.4	5.38×10^{-2}	2.90×10^{-3}	6.18×10^{-2}	3.84×10^{-3}
	0.6	5.10×10^{-2}	2.62×10^{-3}	5.81×10^{-2}	3.38×10^{-3}
	0.8	4.93×10^{-2}	2.45×10^{-3}	5.91×10^{-2}	3.53×10^{-3}
	1	3.28×10^{-2}	1.09×10^{-3}	5.09×10^{-2}	2.66×10^{-3}
	1.2	1.91×10^{-2}	3.90×10^{-4}	3.54×10^{-2}	1.35×10^{-3}
	1.4	1.24×10^{-2}	2.20×10^{-4}	1.28×10^{-2}	2.70×10^{-4}
	1.6	2.09×10^{-3}	4.00×10^{-5}	3.00×10^{-5}	2.10×10^{-4}

TABLE 3
Estimates of recurrent cumulative incidence functions with $a(C_i) = 1$ and $C_i \sim U(0,3)$.

n	t	Abs.bias $\hat{F}_1(t)$	MSE $\hat{F}_1(t)$	Abs.bias $\hat{F}_2(t)$	MSE $\hat{F}_2(t)$
50	0.2	9.27×10^{-3}	2.20×10^{-4}	1.11×10^{-2}	2.40×10^{-4}
	0.4	6.54×10^{-3}	3.10×10^{-4}	9.47×10^{-3}	3.40×10^{-4}
	0.6	4.23×10^{-3}	4.30×10^{-4}	2.41×10^{-3}	3.90×10^{-4}
	0.8	2.04×10^{-2}	9.80×10^{-4}	2.12×10^{-2}	1.00×10^{-3}
	1	4.15×10^{-2}	2.39×10^{-3}	4.54×10^{-2}	2.76×10^{-3}
	1.2	6.41×10^{-2}	4.93×10^{-3}	7.22×10^{-2}	6.07×10^{-3}
	1.4	8.79×10^{-2}	8.71×10^{-3}	1.00×10^{-1}	1.10×10^{-2}
	1.6	1.11×10^{-1}	1.36×10^{-2}	1.28×10^{-1}	1.76×10^{-2}
100	0.2	8.84×10^{-3}	1.30×10^{-4}	1.04×10^{-2}	1.70×10^{-4}
	0.4	5.27×10^{-3}	1.10×10^{-4}	7.67×10^{-3}	2.00×10^{-4}
	0.6	3.57×10^{-3}	1.90×10^{-4}	1.67×10^{-3}	1.50×10^{-4}
	0.8	2.02×10^{-2}	6.50×10^{-4}	2.04×10^{-2}	6.20×10^{-4}
	1	4.10×10^{-2}	1.99×10^{-3}	4.40×10^{-2}	2.21×10^{-3}
	1.2	5.36×10^{-2}	3.10×10^{-3}	6.95×10^{-2}	5.16×10^{-3}
	1.4	7.60×10^{-2}	6.03×10^{-3}	9.47×10^{-2}	9.28×10^{-3}
	1.6	9.36×10^{-2}	8.99×10^{-3}	1.10×10^{-1}	1.23×10^{-2}
200	0.2	7.43×10^{-3}	7.00×10^{-5}	8.42×10^{-3}	9.00×10^{-5}
	0.4	4.79×10^{-3}	6.00×10^{-5}	7.45×10^{-3}	1.00×10^{-4}
	0.6	3.23×10^{-3}	1.00×10^{-4}	9.60×10^{-4}	6.00×10^{-5}
	0.8	1.96×10^{-2}	5.10×10^{-4}	2.03×10^{-2}	5.20×10^{-4}
	1	4.04×10^{-2}	1.78×10^{-3}	4.23×10^{-2}	1.92×10^{-3}
	1.2	5.08×10^{-2}	2.63×10^{-3}	5.47×10^{-2}	3.08×10^{-3}
	1.4	6.38×10^{-2}	4.10×10^{-3}	8.35×10^{-2}	7.07×10^{-3}
	1.6	8.55×10^{-2}	7.35×10^{-3}	1.06×10^{-1}	1.13×10^{-2}
500	0.2	5.30×10^{-3}	3.00×10^{-5}	8.01×10^{-3}	7.00×10^{-5}
	0.4	1.75×10^{-3}	1.00×10^{-5}	6.38×10^{-3}	5.00×10^{-5}
	0.6	3.00×10^{-3}	5.00×10^{-5}	8.30×10^{-4}	3.00×10^{-5}
	0.8	1.95×10^{-2}	4.30×10^{-4}	2.01×10^{-2}	4.40×10^{-4}
	1	4.01×10^{-2}	1.68×10^{-3}	3.99×10^{-2}	1.63×10^{-3}
	1.2	4.55×10^{-2}	2.08×10^{-3}	5.17×10^{-2}	2.69×10^{-3}
	1.4	6.32×10^{-2}	4.01×10^{-3}	8.04×10^{-2}	6.47×10^{-3}
	1.6	8.28×10^{-2}	6.87×10^{-3}	1.06×10^{-1}	1.12×10^{-2}

TABLE 4
Estimates of recurrent cumulative incidence functions with $a(C_i) = 1$ and $C_i \sim U(0,5)$.

n	t	Abs.bias $\hat{F}_1(t)$	MSE $\hat{F}_1(t)$	Abs.bias $\hat{F}_2(t)$	MSE $\hat{F}_2(t)$
50	0.2	3.54×10^{-2}	1.55×10^{-3}	3.86×10^{-2}	1.82×10^{-3}
	0.4	4.83×10^{-2}	2.89×10^{-3}	5.37×10^{-2}	3.54×10^{-3}
	0.6	4.80×10^{-2}	3.16×10^{-3}	5.33×10^{-2}	3.76×10^{-3}
	0.8	3.91×10^{-2}	2.60×10^{-3}	4.29×10^{-2}	2.96×10^{-3}
	1	2.50×10^{-2}	1.89×10^{-3}	2.57×10^{-2}	2.03×10^{-3}
	1.2	6.72×10^{-3}	1.49×10^{-3}	5.21×10^{-3}	1.60×10^{-3}
	1.4	1.30×10^{-2}	1.79×10^{-3}	1.83×10^{-2}	2.07×10^{-3}
	1.6	3.37×10^{-2}	2.88×10^{-3}	4.24×10^{-2}	3.75×10^{-3}
100	0.2	3.43×10^{-2}	1.29×10^{-3}	3.85×10^{-2}	1.62×10^{-3}
	0.4	4.25×10^{-2}	1.97×10^{-3}	5.34×10^{-2}	3.13×10^{-3}
	0.6	4.25×10^{-2}	2.04×10^{-3}	5.30×10^{-2}	3.20×10^{-3}
	0.8	3.64×10^{-2}	1.69×10^{-3}	4.08×10^{-2}	2.08×10^{-3}
	1	2.01×10^{-2}	8.10×10^{-4}	2.31×10^{-2}	1.06×10^{-3}
	1.2	5.72×10^{-3}	5.40×10^{-4}	4.86×10^{-3}	1.33×10^{-3}
	1.4	1.23×10^{-2}	7.20×10^{-4}	1.65×10^{-2}	9.30×10^{-4}
	1.6	3.00×10^{-2}	1.64×10^{-3}	3.87×10^{-2}	2.18×10^{-3}
200	0.2	3.17×10^{-2}	1.04×10^{-3}	3.57×10^{-2}	1.31×10^{-3}
	0.4	3.70×10^{-2}	1.40×10^{-3}	4.30×10^{-2}	1.89×10^{-3}
	0.6	3.67×10^{-2}	1.40×10^{-3}	4.15×10^{-2}	1.78×10^{-3}
	0.8	3.02×10^{-2}	1.00×10^{-3}	3.79×10^{-2}	1.57×10^{-3}
	1	1.88×10^{-2}	5.10×10^{-4}	1.45×10^{-2}	3.20×10^{-4}
	1.2	3.92×10^{-3}	2.50×10^{-4}	2.36×10^{-3}	2.80×10^{-4}
	1.4	1.21×10^{-2}	4.30×10^{-4}	1.50×10^{-2}	4.50×10^{-4}
	1.6	2.15×10^{-2}	7.00×10^{-4}	2.23×10^{-2}	6.40×10^{-4}
500	0.2	2.10×10^{-2}	4.40×10^{-4}	3.18×10^{-2}	1.02×10^{-3}
	0.4	3.27×10^{-2}	1.07×10^{-3}	3.95×10^{-2}	1.56×10^{-3}
	0.6	2.73×10^{-2}	7.50×10^{-4}	3.87×10^{-2}	1.51×10^{-3}
	0.8	2.80×10^{-2}	8.10×10^{-4}	3.44×10^{-2}	1.22×10^{-3}
	1	1.40×10^{-2}	2.20×10^{-4}	6.14×10^{-3}	5.00×10^{-5}
	1.2	5.10×10^{-4}	5.00×10^{-5}	1.98×10^{-3}	1.10×10^{-4}
	1.4	1.02×10^{-2}	1.90×10^{-4}	1.06×10^{-2}	1.90×10^{-4}
	1.6	1.37×10^{-2}	2.50×10^{-4}	1.95×10^{-2}	4.30×10^{-4}

TABLE 5
Estimates of recurrent cumulative incidence functions with $a(C_i) = C = 3$.

n	t	Abs.bias $\hat{F}_1(t)$	MSE $\hat{F}_1(t)$	Abs.bias $\hat{F}_2(t)$	MSE $\hat{F}_2(t)$
50	0.2	3.65×10^{-2}	1.59×10^{-3}	4.10×10^{-2}	1.98×10^{-3}
	0.4	5.17×10^{-2}	3.18×10^{-3}	5.87×10^{-2}	4.05×10^{-3}
	0.6	5.19×10^{-2}	3.39×10^{-3}	5.96×10^{-2}	4.41×10^{-3}
	0.8	4.26×10^{-2}	2.67×10^{-3}	5.04×10^{-2}	3.62×10^{-3}
	1	2.82×10^{-2}	1.82×10^{-3}	3.37×10^{-2}	2.38×10^{-3}
	1.2	9.69×10^{-3}	1.21×10^{-3}	1.15×10^{-2}	1.55×10^{-3}
	1.4	1.06×10^{-2}	1.36×10^{-3}	1.30×10^{-2}	1.71×10^{-3}
	1.6	3.17×10^{-2}	2.36×10^{-3}	3.79×10^{-2}	3.06×10^{-3}
100	0.2	3.56×10^{-2}	1.40×10^{-3}	3.83×10^{-2}	1.56×10^{-3}
	0.4	5.05×10^{-2}	2.79×10^{-3}	5.79×10^{-2}	3.65×10^{-3}
	0.6	5.06×10^{-2}	2.90×10^{-3}	5.94×10^{-2}	3.92×10^{-3}
	0.8	4.14×10^{-2}	2.16×10^{-3}	4.92×10^{-2}	2.90×10^{-3}
	1	2.66×10^{-2}	1.21×10^{-3}	3.20×10^{-2}	1.61×10^{-3}
	1.2	7.82×10^{-3}	6.20×10^{-4}	1.01×10^{-2}	7.70×10^{-4}
	1.4	9.72×10^{-3}	6.10×10^{-4}	1.29×10^{-2}	8.20×10^{-4}
	1.6	2.86×10^{-2}	1.30×10^{-3}	3.74×10^{-2}	2.09×10^{-3}
200	0.2	2.98×10^{-2}	9.10×10^{-4}	3.53×10^{-2}	1.27×10^{-3}
	0.4	4.93×10^{-2}	2.54×10^{-3}	5.06×10^{-2}	2.62×10^{-3}
	0.6	4.92×10^{-2}	2.57×10^{-3}	5.35×10^{-2}	2.97×10^{-3}
	0.8	4.09×10^{-2}	1.87×10^{-3}	4.44×10^{-2}	2.11×10^{-3}
	1	2.62×10^{-2}	9.40×10^{-4}	2.88×10^{-2}	1.03×10^{-3}
	1.2	6.74×10^{-3}	2.90×10^{-4}	6.10×10^{-3}	2.50×10^{-4}
	1.4	4.84×10^{-3}	2.70×10^{-4}	2.94×10^{-3}	2.00×10^{-4}
	1.6	2.34×10^{-2}	7.80×10^{-4}	2.67×10^{-2}	9.10×10^{-4}
500	0.2	2.77×10^{-2}	7.70×10^{-4}	3.45×10^{-2}	1.20×10^{-3}
	0.4	4.17×10^{-2}	1.75×10^{-3}	4.75×10^{-2}	2.27×10^{-3}
	0.6	4.24×10^{-2}	1.81×10^{-3}	4.64×10^{-2}	2.16×10^{-3}
	0.8	3.70×10^{-2}	1.40×10^{-3}	3.79×10^{-2}	1.46×10^{-3}
	1	2.29×10^{-2}	5.70×10^{-4}	2.50×10^{-2}	6.60×10^{-4}
	1.2	4.85×10^{-3}	8.00×10^{-5}	2.29×10^{-3}	5.00×10^{-5}
	1.4	2.35×10^{-3}	5.00×10^{-5}	1.90×10^{-4}	5.00×10^{-5}
	1.6	2.23×10^{-2}	5.50×10^{-4}	2.34×10^{-2}	6.00×10^{-4}

TABLE 6
Estimates of recurrent cumulative incidence functions with $a(C_i) = C = 5$.

n	t	Abs.bias $\hat{F}_1(t)$	MSE $\hat{F}_1(t)$	Abs.bias $\hat{F}_2(t)$	MSE $\hat{F}_2(t)$
50	0.2	1.05×10^{-1}	1.16×10^{-2}	1.09×10^{-1}	1.24×10^{-2}
	0.4	1.57×10^{-1}	2.57×10^{-2}	1.66×10^{-1}	2.86×10^{-2}
	0.6	1.82×10^{-1}	3.44×10^{-2}	1.91×10^{-1}	3.78×10^{-2}
	0.8	1.88×10^{-1}	3.72×10^{-2}	1.96×10^{-1}	4.01×10^{-2}
	1	1.85×10^{-1}	3.63×10^{-2}	1.91×10^{-1}	3.86×10^{-2}
	1.2	1.76×10^{-1}	3.33×10^{-2}	1.78×10^{-1}	3.40×10^{-2}
	1.4	1.63×10^{-1}	2.90×10^{-2}	1.61×10^{-1}	2.84×10^{-2}
	1.6	1.48×10^{-1}	2.43×10^{-2}	1.42×10^{-1}	2.28×10^{-2}
100	0.2	1.04×10^{-1}	1.11×10^{-2}	1.09×10^{-1}	1.21×10^{-2}
	0.4	1.56×10^{-1}	2.49×10^{-2}	1.64×10^{-1}	2.75×10^{-2}
	0.6	1.81×10^{-1}	3.33×10^{-2}	1.89×10^{-1}	3.65×10^{-2}
	0.8	1.88×10^{-1}	3.60×10^{-2}	1.95×10^{-1}	3.89×10^{-2}
	1	1.85×10^{-1}	3.52×10^{-2}	1.89×10^{-1}	3.67×10^{-2}
	1.2	1.75×10^{-1}	3.18×10^{-2}	1.77×10^{-1}	3.22×10^{-2}
	1.4	1.62×10^{-1}	2.73×10^{-2}	1.60×10^{-1}	2.68×10^{-2}
	1.6	1.47×10^{-1}	2.28×10^{-2}	1.41×10^{-1}	2.13×10^{-2}
200	0.2	1.03×10^{-1}	1.07×10^{-2}	1.05×10^{-1}	1.12×10^{-2}
	0.4	1.49×10^{-1}	2.23×10^{-2}	1.56×10^{-1}	2.43×10^{-2}
	0.6	1.79×10^{-1}	3.24×10^{-2}	1.77×10^{-1}	3.14×10^{-2}
	0.8	1.80×10^{-1}	3.26×10^{-2}	1.84×10^{-1}	3.41×10^{-2}
	1	1.70×10^{-1}	2.91×10^{-2}	1.78×10^{-1}	3.20×10^{-2}
	1.2	1.61×10^{-1}	2.63×10^{-2}	1.62×10^{-1}	2.64×10^{-2}
	1.4	1.52×10^{-1}	2.35×10^{-2}	1.57×10^{-1}	2.51×10^{-2}
	1.6	1.37×10^{-1}	1.90×10^{-2}	1.34×10^{-1}	1.83×10^{-2}
500	0.2	9.90×10^{-2}	9.81×10^{-3}	9.90×10^{-2}	9.81×10^{-3}
	0.4	1.40×10^{-1}	1.95×10^{-2}	1.46×10^{-1}	2.15×10^{-2}
	0.6	1.60×10^{-1}	2.57×10^{-2}	1.66×10^{-1}	2.76×10^{-2}
	0.8	1.68×10^{-1}	2.84×10^{-2}	1.75×10^{-1}	3.05×10^{-2}
	1	1.62×10^{-1}	2.62×10^{-2}	1.66×10^{-1}	2.77×10^{-2}
	1.2	1.51×10^{-1}	2.29×10^{-2}	1.53×10^{-1}	2.36×10^{-2}
	1.4	1.41×10^{-1}	1.99×10^{-2}	1.40×10^{-1}	1.95×10^{-2}
	1.6	1.24×10^{-1}	1.55×10^{-2}	1.20×10^{-1}	1.43×10^{-2}

6. An Application

We now apply the proposed methods to the Kidney catheter data (Lawless, 2011). The data set is available in *survival* package of R software (https://vincentarel bundock.github .io/Rdatasets/csv/survival/kidney.csv). The complete analysis of the data is not carried out here. Our aim here is to illustrate the procedure using real-life data. These data show the recurrence time to infection at the point of insertion of the catheter for 38 kidney patients using portable dialysis equipment with four different causes. Catheters may be removed for reasons other than infection, in which case the observation is censored. Data on the first and second recurrences of infection are given. Kidney disease type glomerulonephritis (GN), acute nephritis (AN), polycystic kidney diseases (PKD) and other diseases (OTHER) are four different causes. Table 7 shows the kidney catheter data. The variable id represents the individual, and recurrence time is the time at which the infection occurred. The variable status shows 1 if the infection happens and 0 if the patient is censored. The variable disease shows four different types of kidney disease.

The tiny structures that do the work in our kidneys are called nephrons. These are the basic functional units of blood filtration and urine production. Each of our kidneys contains about one million nephrons. Each nephron has a small blood vessel that brings in unfiltered blood, a glomerulus that filters the blood, a tubule that caries away filtered waste materials in the urine and a small blood vessel that returns filtered blood to the body. Nephritis is inflammation of the kidneys that may involve the glomeruli, tubules or interstitial tissue surrounding the glomeruli and tubules. Diseases that injure glomeruli are called glomerular diseases. There are two types of glomeruli nephritis - acute and chronic. The acute form develops suddenly. It may be caused by infections such as strep throat. Polycystic kidney disease (PKD) is an inherited kidney disorder. It causes fluid-filled cysts to form in the kidneys. PKD may impair kidney function and eventually cause kidney failure. One of the leading causes of kidney failure is PKD. Kidney disease usually leads to kidney failure to some degree which depends on the type of disease.

We compute estimates of recurrent cumulative incidence functions (RCIF) corresponding to four causes, and the results are presented in Figure 1. The red, blue, black and green curves represent the estimators of RCIF corresponding to the causes 'GN', 'AN', 'PKD' and 'OTHER', respectively. The individuals under study have different censoring times, and we have chosen the weight function $a(C_i) = C_i$ for computing the estimates.

In the early period, three causes, GN, AN and OTHER, are competing with each other and the third cause PKD, seems to have lesser risk as the cumulative incidence function is small. Initially, the cause AN is dominating up to 17 days. After that, the cause OTHER is dominating up to 51 days. The cause of GN having control over the remaining causes up to 245 days. Finally, the fourth cause, OTHER, is dominating after 245 days. Thus the probability of occurrence of recurrent infection in kidney patients due to the fourth cause, OTHER, is more dominant than the other three causes after

TABLE 7
Kidney catheter data.

id	recurrent time	status	disease	id	recurrent time	status	disease
1	8	1	Other	20	15	1	Other
1	16	1	Other	20	108	0	Other
2	23	1	GN	21	152	1	PKD
2	13	0	GN	21	562	1	PKD
3	22	1	Other	22	402	1	Other
3	28	1	Other	22	24	0	Other
4	447	1	Other	23	13	1	AN
4	318	1	Other	23	66	1	AN
5	30	1	Other	24	39	1	AN
5	12	1	Other	24	46	0	AN
6	24	1	Other	25	12	1	AN
6	245	1	Other	25	40	1	AN
7	7	1	GN	26	113	0	AN
7	9	1	GN	26	201	1	AN
8	511	1	GN	27	132	1	GN
8	30	1	GN	27	156	1	GN
9	53	1	AN	28	34	1	AN
9	196	1	AN	28	30	1	AN
10	15	1	GN	29	2	1	GN
10	154	1	GN	29	25	1	GN
11	7	1	AN	30	130	1	GN
11	333	1	AN	30	26	1	GN
12	141	1	Other	31	27	1	AN
12	8	0	Other	31	58	1	AN
13	96	1	AN	32	5	0	AN
13	38	1	AN	32	43	1	AN
14	149	0	AN	33	152	1	PKD
14	70	0	AN	33	30	1	PKD
15	536	1	Other	34	190	1	GN
15	25	0	Other	34	5	0	GN
16	17	1	AN	35	119	1	Other
16	4	0	AN	35	8	1	Other
17	185	1	Other	36	54	0	Other
17	177	1	Other	36	16	0	Other
18	292	1	Other	37	6	0	PKD
18	114	1	Other	37	78	1	PKD
19	22	0	GN	38	63	1	PKD
19	159	0	GN	38	8	0	PKD

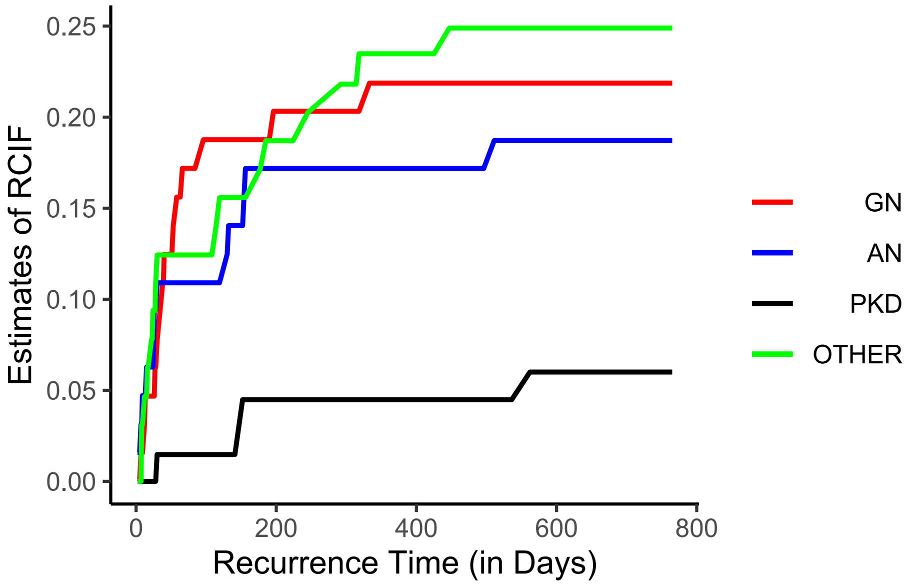

Figure 1 – Estimates of the recurrent cumulative incidence functions (RCIF) for GN, AN, PKD and OTHER.

245 days. The third cause, PKD, does not have much influence on the recurrence time.

7. Conclusion

In this article, we have developed non-parametric estimators of cumulative incidence functions of recurrent events experiencing competing risks. Simulation results established that the absolute bias and mean squared error of the estimators decrease as sample size increases. The proposed method was applied to a real-life data set. The existence of frailty assumption is used for ensuring the i.i.d. structure of the recurrence time within the subject, and the distribution of frailty is not used in this proposed methodology. The work in this direction will be reported elsewhere.

Acknowledgements

The authors are grateful to the referees and editor for their constructive comments. The first author thanks INSPIRE, Department of Science and Technology, Government of India, for providing financial support. The second author thanks KSCSTE, Government of Kerala, for providing financial support.

APPENDIX

A. PROOFS

PROOF. Theorem 1
We use an empirical process approach to prove the theorem. Write,

$$\begin{aligned}
\hat{\Lambda}_l(t) - \Lambda_l(t) &= \left\{ \int_0^t \frac{d\hat{G}_{al}(u)}{\hat{H}_a(u)} - \frac{dG_{al}(u)}{H_a(u)} \right\} \\
&= \int_0^t \left\{ \frac{1}{\hat{H}_a(u)} - \frac{1}{H_a(u)} \right\} dG_{al}(u) + \int_0^t \frac{1}{H_a(u)} d\left\{ \hat{G}_{al}(u) - G_{al}(u) \right\} \\
&\quad + \int_0^t \left\{ \frac{1}{\hat{H}_a(u)} - \frac{1}{H_a(u)} \right\} d\left\{ \hat{G}_{al}(u) - G_{al}(u) \right\}. \qquad (19)
\end{aligned}$$

The third term in Eq. (19) can be shown to be of order of $o_p(n^{-\frac{1}{2}})$ and is asymptotically negligible. Approximating the first term of Eq. (19) using techniques similar to those of Breslow and Crowley (1974), we obtain

$$\begin{aligned}
\hat{\Lambda}_l(t) - \Lambda_l(t) &= \int_0^t \left\{ \frac{H_a(u) - \hat{H}_a(u)}{H_a(u)\hat{H}_a(u)} \right\} dG_{al}(u) + \int_0^t \{H_a(u)\}^{-1} d\left\{ \hat{G}_{al}(u) - G_{al}(u) \right\} \\
&\quad + o_p(n^{-\frac{1}{2}}) \\
&= \int_0^t \{H_a(u)\}^{-2} \left\{ H_a(u) - \hat{H}_a(u) \right\} dG_{al}(u) + \\
&\qquad \int_0^t \{H_a(u)\}^{-1} d\left\{ \hat{G}_{al}(u) - G_{al}(u) \right\} + o_p(n^{-\frac{1}{2}}) \\
&= -\int_0^t \{H_a(u)\}^{-2} \hat{H}_a(u) dG_{al}(u) + \int_0^t \{H_a(u)\}^{-1} d\hat{G}_{al}(u) + o_p(n^{-\frac{1}{2}}) \\
&= -\frac{1}{n} \sum_{i=1}^n \psi_{il}(t) + o_p(n^{-\frac{1}{2}}), \qquad (20)
\end{aligned}$$

where

$$\begin{aligned}
\psi_{il}(t) &= \int_0^t \{H_a(u)\}^{-2} \left\{ \frac{a_i}{m_i^*} \sum_{j=1}^{m_i^*} I(y_{ij} \geq u) \right\} dG_{al}(u) \\
&\quad - \frac{a_i I(m_{il} \geq 2)}{m_{il}} \sum_{j=1}^{m_{il}} \frac{I(y_{ij} < t, J_{ij} = l)}{H_a(y_{ij})}, \; l = 1, 2, \ldots, k. \qquad (21)
\end{aligned}$$

Now applies integration by parts on the first term of Eq. (21) and deletes higher order terms which are asymptotically negligible. Then we get $E[\psi_{il}(t)] = 0$.

Define $W_l(t) = \sqrt{n}\{\hat{\Lambda}_l(t) - \Lambda_l(t)\}$. From Eq. (20), we obtain

$$\begin{aligned} W_l(t) &= \sqrt{n}\{\hat{\Lambda}_l(t) - \Lambda_l(t)\} \\ &= -n^{-\frac{1}{2}} \sum_{i=1}^{n} \psi_{il}(t) + o_p(n^{-\frac{1}{2}}). \end{aligned} \tag{22}$$

Note that

$$\hat{F}_l(t) - F_l(t) = \int_0^t S(u) d\left\{\hat{\Lambda}_l(u) - \Lambda_l(u)\right\} + \int_0^t \left\{\hat{S}(u) - S(u)\right\} d\hat{\Lambda}_l(u). \tag{23}$$

Note that $\Lambda(t) = \sum_{l=1}^{k} \Lambda_l(t)$, $S(t) = \exp -(\Lambda(t))$ and $\hat{S}(t) = \exp -(\hat{\Lambda}(t))$. Then,

$$\int_0^t \{\hat{S}(u) - S(u)\} d\hat{\Lambda}_l(u) = \int_0^t (-S(u))\left(1 - e^{-\sum_{l=1}^{k} \hat{\Lambda}_l(u) + \sum_{l=1}^{k} \Lambda_l(u)}\right) d\hat{\Lambda}_l(u) \tag{24}$$

Using Taylor series expansion, we obtain Eq. (24) as

$$\begin{aligned} \int_0^t \{\hat{S}(u) - S(u)\} d\hat{\Lambda}_l(u) &= \int_0^t (-S(u))\left(1 - \left[1 + -\sum_{l=1}^{k} \hat{\Lambda}_l(u) + \sum_{l=1}^{k} \Lambda_l(u) + o_p(n)\right]\right) \\ &\qquad \times d\hat{\Lambda}_l(u) \\ &= -\int_0^t (S(u)\left(\sum_{l=1}^{k} \hat{\Lambda}_l(u) - \sum_{l=1}^{k} \Lambda_l(u)\right) d\hat{\Lambda}_l(u). \end{aligned} \tag{25}$$

It follows that,

$$\sqrt{n}\{\hat{F}_l(t) - F_l(t)\} \approx \int_0^t S(u) dW_l(u) - \sum_{l=1}^{k} \int_0^t S(u) W_l(u) d\Lambda_l(u). \tag{26}$$

Now the second term in the expression Eq. (26) can be written as

$$\begin{aligned} \int_0^t S(u) W_l(u) d\Lambda_l(u) &= \int_0^t W_l(u) S(u) d\Lambda_l(u) \\ &= \int_0^t W_l(u) dF_l(u). \end{aligned} \tag{27}$$

Then, by using integration by parts, Eq. (27) becomes,

$$\begin{aligned}\int_0^t W_l(u)dF_l(u) &= F_l(t)W_l(t) - \int_0^t F_l(u)dW_l(u)\\ &= \int_0^t \{F_l(t)-F_l(u)\}\, dW_l(u)\\ &= \int_0^t F_l^c(t,u)dW_l(u), \end{aligned} \tag{28}$$

where $F_l^c(t,u) = F_l(t) - F_l(u)$. Using Eq. (22), the identity Eq. (27) becomes

$$\begin{aligned}\int_0^t S(u)W_l(u)d\Lambda_l(u) &= \int_0^t F_l^c(t,u)dW_l(u)\\ &= -\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\int_0^t F_l^c(t,u)\psi_{il}(u) + o_p(n^{-\frac{1}{2}}). \end{aligned} \tag{29}$$

Substituting Eq. (29) in Eq. (26) we get,

$$\begin{aligned}\sqrt{n}\{\hat{F}_l(t)-F_l(t)\} &\approx \int_0^t S(u)dW_l(u) - \sum_{l=1}^{k}\int_0^t F_l^c dW_l(u)\\ &= -\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\int_0^t S(u)\psi_{il}(u) + \frac{1}{\sqrt{n}}\sum_{l=1}^{k}\sum_{i=1}^{n}\int_0^t F_l^c(t,u)\psi_{il}(u)\\ &\qquad + o_p(n^{-\frac{1}{2}})\\ &= -\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\left[\sum_{l=1}^{k}\int_0^t F_l^c(t,u)\psi_{il}(u) + \int_0^t S(u)\psi_{il}(u)\right] + o_p(n^{-\frac{1}{2}})\\ &= -\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\left[\int_0^t [\sum_{l=1}^{k} F_l^c(t,u) + S(u)]\psi_{il}(u)\right] + o_p(n^{-\frac{1}{2}})\\ &= -\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\epsilon_{li}(t) + o_p(n^{-\frac{1}{2}}). \end{aligned} \tag{30}$$

where $\epsilon_{li}(t) = \int_0^t [\sum_{l=1}^{k} F_l^c(t,u) + S(u)]\psi_{il}(u)$.

Hence, $\sqrt{n}\{\hat{F}_l(t) - F_l(t)\}$ has an asymptotically i.i.d. representation of Eq. (30). Then, $E(\epsilon_{li}(t)) = 0$, since $\psi_{il}(t) = 0$. The random variables $F_l^c(t,u)\psi_{il}(t)$ and $S(t)\psi_{il}(t)$ are uniformly bounded in $0 \le t \le t^*$ when a_i is bounded, $i = 1,2,\ldots,n$ and $l = 1,2,\ldots,k$. Then the random variable $\epsilon_{li}(t)$ is also uniformly bounded in $0 \le t \le t^*$. Thus, by multivariate central limit theorem, the finite-dimensional distributions of a random process $\sqrt{n}\{\hat{F}_l(t)-F_l(t)\}$ converge weakly to mean zero Gaussian processes with variance-covariance function

$$\sigma_l(t_1,t_2)=E[\epsilon_{li}(t_1)\epsilon_{li}(t_2)], \quad l=1,2,\ldots,k. \tag{31}$$

Using the techniques of Billingsley (2013), we can derive

$$E\left[\left(\hat{H}_a(t)-H_a(t_1)\right)\left(\hat{H}_a(t_2)-H_a(t)\right)\right]\leq \text{constant.}(t-t_1)(t_2-t), \tag{32}$$

$$E\left[\left(\hat{G}_a(t)-G_a(t_1)\right)\left(\hat{G}_a(t_2)-G_a(t)\right)\right]\leq \text{constant.}(t-t_1)(t_2-t) \tag{33}$$

and

$$E\left[\left(\hat{G}_{al}(t)-G_{al}(t_1)\right)\left(\hat{G}_{al}(t_2)-G_{al}(t)\right)\right]\leq \text{constant.}(t-t_1)(t_2-t), \quad l=1,2,..,k \tag{34}$$

for $t_1 \leq t \leq t_2$. The tightness of sequences follows from Theorem 13.5 of Billingsley (2013) and arguments similar to those of Breslow and Crowley (1974). Thus we complete the proof. □

PROOF. Theorem 3
We have

$$\hat{F}_l(t)-F_l(t)=\int_0^t S(u)d\left\{\hat{\Lambda}_l(u)-\Lambda_l(u)\right\}+\int_0^t \left\{\hat{S}(u)-S(u)\right\}d\hat{\Lambda}_l(u). \tag{35}$$

Now

$$\begin{aligned}\sup_{0<t<t^*}\left|\hat{F}_l(t)-F_l(t)\right| \leq & \sup_{0<t<t^*}\left|\int_0^t S(u)d\left(\hat{\Lambda}_l(u)-\Lambda_l(u)\right)\right| \\ & +\sup_{0<t<t^*}\left|\int_0^t \left(\hat{S}(u)-S(u)\right)d\hat{\Lambda}_l(u)\right|.\end{aligned} \tag{36}$$

We know that

$$\hat{S}(t)-S(t)=\frac{1}{n}\sum_{i=1}^{n}S(t)\phi_i(t)+o_p(n^{-\frac{1}{2}}), \tag{37}$$

where

$$\begin{aligned}\phi_i(t)=&\int_0^t \{H_a(u)\}^{-2}\{\frac{a_i}{m_i^*}\sum_{j=1}^{m_i^*}I(y_{ij}\geq u)\}dG_a(u)\\ &\frac{a_iI(m_i\geq 2)}{m_i^*}\sum_{j=1}^{m_i^*}\frac{1}{H_a(y_{ij})}I(y_{ij}<t)\end{aligned}$$

and

$$\hat{\Lambda}_l(t) - \Lambda_l(t) = -\frac{1}{n}\sum_{i=1}^{n}\psi_{il}(t) + o_p(n^{-\frac{1}{2}}). \tag{38}$$

Then, as $n \to \infty$

$$\sup_{0<t<t^*}\left|\hat{S}(t) - S(t)\right| \xrightarrow{a.s} 0 \tag{39}$$

and

$$\sup_{0<t<t^*}\left|\hat{\Lambda}_l(t) - \Lambda_l(t)\right| \xrightarrow{a.s} 0. \tag{40}$$

Thus, we can conclude that as $n \to \infty$

$$\sup_{0<t<t^*}\left|\hat{F}_l(t) - F_l(t)\right| \xrightarrow{a.s} 0. \tag{41}$$

which implies that $\hat{F}_l(t)$ is strongly consistent for $F_l(t)$. □

REFERENCES

P. K. ANDERSEN, J. ANGST, H. RAVN (2019). *Modeling marginal features in studies of recurrent events in the presence of a terminal event*. Lifetime Data Analysis, 25, no. 4, pp. 681–695.

J. BEYERSMANN, A. LATOUCHE, A. BUCHHOLZ, M. SCHUMACHER (2009). *Simulating competing risks data in survival analysis*. Statistics in Medicine, 28, no. 6, pp. 956–971.

P. BILLINGSLEY (2013). *Convergence of Probability Measures*. John Wiley & Sons, New York.

O. BOUAZIZ, F. COMTE, A. GUILLOUX (2013). *Nonparametric estimation of the intensity function of a recurrent event process*. Statistica Sinica, 23, no. 2, pp. 635–665.

N. BRESLOW, J. CROWLEY (1974). *A large sample study of the life table and product limit estimates under random censorship*. The Annals of Statistics, 2, no. 3, pp. 437–453.

X. CHE, J. ANGUS (2016). *A new joint model of recurrent event data with the additive hazards model for the terminal event time*. Metrika, 79, no. 7, pp. 763–787.

R. J. COOK, J. LAWLESS (2007). *The Statistical Analysis of Recurrent Events*. Springer Science & Business Media, New York.

R. J. COOK, J. F. LAWLESS, C. NADEAU (1996). *Robust tests for treatment comparisons based on recurrent event responses*. Biometrics, 52, no. 2, pp. 557–571.

M. J. CROWDER (2001). *Classical Competing Risks*. Chapman and Hall/CRC, London.

J. Y. DAUXOIS, S. SENCEY (2009). *Non-parametric tests for recurrent events under competing risks*. Scandinavian Journal of Statistics, 36, no. 4, pp. 649–670.

H. DONG, L. L. ROBISON, W. M. LEISENRING, L. J. MARTIN, G. T. ARMSTRONG, Y. YASUI (2015). *Estimating the burden of recurrent events in the presence of competing risks: the method of mean cumulative count*. American Journal of Epidemiology, 181, no. 7, pp. 532–540.

S. GEFFRAY (2013). *Modeling and inferential thoughts for consecutive gap times observed with death and censoring*. Journal of the Iranian Statistical Society, 12, no. 1, pp. 71–112.

D. GHOSH (2004). *Accelerated rates regression models for recurrent failure time data*. Lifetime Data Analysis, 10, no. 3, pp. 247–261.

N. A. GOUSKOVA, F. C. LIN, J. P. FINE (2017). *Nonparametric analysis of competing risks data with event category missing at random*. Biometrics, 73, no. 1, pp. 104–113.

M. HAN, L. SUN, Y. LIU, J. ZHU (2018). *Joint analysis of recurrent event data with additive–multiplicative hazards model for the terminal event time*. Metrika, 81, no. 5, pp. 523–547.

J. F. LAWLESS (2011). *Statistical Models and Methods for Lifetime Data*. John Wiley & Sons, New York.

J. F. LAWLESS, C. NADEAU (1995). *Some simple robust methods for the analysis of recurrent events*. Technometrics, 37, no. 2, pp. 158–168.

D. LIN, L. WEI, Z. YING (1998). *Accelerated failure time models for counting processes*. Biometrika, 85, no. 3, pp. 605–618.

D. Y. LIN, L. J. WEI, I. YANG, Z. YING (2000). *Semiparametric regression for the mean and rate functions of recurrent events*. Journal of the Royal Statistical Society-Series B, 62, no. 4, pp. 711–730.

H. MA, L. PENG, Z. ZHANG, H. J. LAI (2018). *Generalized accelerated recurrence time model for multivariate recurrent event data with missing event type*. Biometrics, 74, no. 3, pp. 954–965.

P. MAITRA, L. D. AMORIM, J. CAI (2020). *Multiplicative rates model for recurrent events in case-cohort studies*. Lifetime Data Analysis, 26, no. 1, pp. 134–157.

J. NING, M. H. RAHBAR, S. CHOI, J. PIAO, C. HONG, D. J. DEL JUNCO, E. RAHBAR, E. E. FOX, J. B. HOLCOMB, M.-C. WANG (2017). *Estimating the ratio of multivariate recurrent event rates with application to a blood transfusion study*. Statistical Methods in Medical Research, 26, no. 4, pp. 1969–1981.

F. OSMANI, E. HAJIZADEH, A. A. RASEKHI (2018). *Association between multiple recurrent events with multivariate modeling: a retrospective cohort study*. Journal of Research in Health Sciences, 18, no. 4, p. e00433.

M. S. PEPE, J. CAI (1993). *Some graphical displays and marginal regression analyses for recurrent failure times and time dependent covariates*. Journal of the American Statistical Association, 88, no. 423, pp. 811–820.

R. L. PRENTICE, B. J. WILLIAMS, A. V. PETERSON (1981). *On the regression analysis of multivariate failure time data*. Biometrika, 68, no. 2, pp. 373–379.

L. QU, L. SUN, L. LIU (2017). *Joint modeling of recurrent and terminal events using additive models*. Statistics and its Interface, 10, no. 4, pp. 699–710.

P. G. SANKARAN, P. ANISHA (2011). *Shared frailty model for recurrent event data with multiple causes*. Journal of Applied Statistics, 38, no. 12, pp. 2859–2868.

P. G. SANKARAN, P. ANISHA (2012). *Additive hazards models for gap time data with multiple causes*. Statistics & Probability Letters, 82, no. 7, pp. 1454–1462.

R. STOCKER, A. ADEKPEDJOU (2020). *A general class of additive semiparametric models for recurrent event data*. Journal of Statistical Planning and Inference, 205, pp. 231–244.

L. L. TAYLOR, E. A. PENA (2014). *Nonparametric estimation with recurrent competing risks data*. Lifetime Data Analysis, 20, no. 4, pp. 514–537.

M. C. WANG, S. H. CHANG (1999). *Nonparametric estimation of a recurrent survival function*. Journal of the American Statistical Association, 94, no. 445, pp. 146–153.

L. J. WEI, D. Y. LIN, L. WEISSFELD (1989). *Regression analysis of multivariate incomplete failure time data by modeling marginal distributions*. Journal of the American Statistical Association, 84, no. 408, pp. 1065–1073.

H. YU, Y.-J. CHENG, C.-Y. WANG (2018). *Methods for multivariate recurrent event data with measurement error and informative censoring*. Biometrics, 74, no. 3, pp. 966–976.

D. ZENG, J. CAI (2010). *A semiparametric additive rate model for recurrent events with an informative terminal event*. Biometrika, 97, no. 3, pp. 699–712.