



JOHNS HOPKINS  
BLOOMBERG  
SCHOOL of PUBLIC HEALTH

---

Johns Hopkins University, Dept. of Biostatistics Working Papers

---

10-20-2003

# Nonparametric Estimation of the Bivariate Recurrence Time Distribution

Chiung-Yu Huang

*Division of Biostatistics, School of Public Health, University of Minnesota, cyhuang@biostat.umn.edu*

Mei-Cheng Wang

*Johns Hopkins Bloomberg School of Public Health, mcwang@jhsph.edu*

---

## Suggested Citation

Huang, Chiung-Yu and Wang, Mei-Cheng, "Nonparametric Estimation of the Bivariate Recurrence Time Distribution" (October 2003). *Johns Hopkins University, Dept. of Biostatistics Working Papers*. Working Paper 26.  
<http://biostats.bepress.com/jhubiostat/paper26>

This working paper is hosted by The Berkeley Electronic Press (bepress) and may not be commercially reproduced without the permission of the copyright holder.

Copyright © 2011 by the authors

# Nonparametric Estimation of the Bivariate Recurrence Time Distribution

**Chiung-Yu Huang**

Division of Biostatistics, School of Public Health, University of Minnesota,

420 Delaware Street S.E., Minneapolis, Minnesota 55455, U.S.A.

email: cyhuang@biostat.umn.edu

**and**

**Mei-Cheng Wang**

Department of Biostatistics, Johns Hopkins University,

615 N. Wolfe Street, Baltimore, Maryland 21205, U.S.A.

**SUMMARY.** This paper considers statistical models in which two different types of events, such as the diagnosis of a disease and the remission of the disease, occur alternately over time and are observed subject to right censoring. We propose nonparametric estimators for the joint distribution of bivariate recurrence times and the marginal distribution of the first recurrence time. In general, the marginal distribution of the second recurrence time cannot be estimated due to an identifiability problem, but a conditional distribution of the second recurrence time can be estimated nonparametrically. In literature, statistical methods have been developed to estimate the joint distribution of bivariate recurrence times based on data of the first pair of censored bivariate recurrence times. These methods are inefficient in the current model because recurrence times of higher orders are not used. Asymptotic properties of the estimators are established. Numerical studies demonstrate the estimator performs well with practical sample sizes. We apply the proposed method to a Denmark psychiatric case register data set for illustration of the methods and theory.

**KEY WORDS:** Alternating renewal process; Recurrent events; Frailty.

# 1 INTRODUCTION

## 1.1 BACKGROUND

Recurrent event data arise in longitudinal studies where each study subject may experience multiple events during follow-up. In analyzing recurrent event data, two types of time scale are considered in the literature: the time since entering the study and the time since the last event. For the situation where the time since study entry is of interest, a variety of statistical methods has been developed, among them the methods proposed by Andersen and Gill (1982), Prentice, Williams, and Peterson (1981), Pepe and Cai (1993), Lin, Wei, Yang, and Ying (2000), and Wang, Qin, and Chiang (2001). These methods consider individual's multiple events as the realization of a counting process, and formulate their model based on either the intensity function or the occurrence rate function of the underlying event process.

When the study interest is placed on the times between consecutive events, i.e. gap times, the stochastic ordering structure of recurrent events generates challenges for statistical analysis, such as induced dependent censoring and sampling bias, and consequently it impedes the development of statistical methods. In recent years, various statistical methods have been proposed for the analysis of gap times: when the events are of the same type, Pena, Strawderman, and Hollander (2002) proved that the generalized Kaplan-Meier estimator has the nonparametric maximal likelihood estimator (NPMLE) property with the iid assumption on gap time distributions; Wang and Chang (1999) relaxed the independence assumption on gap times and extended the risk set methods for the estimation of univariate recurrence time distribution. When the events are of different types, various nonparametric methods, such as Visser (1996), Huang and Louis (1998), Wang and Wells (1998), and Lin, Sun, and Ying (1999), as well as semiparametric methods, such as Huang (1999), Chang (2000), and Lin (2000), have been developed in the literature.

In many applications, study subjects experience two different types of events alternately over time. For examples, a patient with the schizophrenia disease could be repeatedly admitted into and discharged from a hospital; in a reliability study, a mechanical system may alternate between periods of use and repair; in the study of the South Verona psychiatric case register data a patient is defined, at each time point, to be in either the stage of a PCR-break, if he/she has not had any contact with a mental health service for 90 days, or in an PCR episode of care (PCR-EC). When evaluating the efficacy of an intervention, researchers may want to study whether the treatment shortens the illness stage as well as whether it prolongs the duration of the disease-relapse stage. In the reliability example, the duration in service and the downtime during system repair are both meaningful in the comparison of two mechanical systems.

The estimation of the bivariate distribution of recurrence times not only plays an important role in estimating the degree of association within bivariate recurrent events, but also serves as a basis of model building and testing. In the literature, statistical methods have been developed for estimating the multivariate distribution when events are of different types, as well as the univariate distribution when all the events are of the same type. These methods can be applied to the current bivariate recurrence time data but they are subject to serious limitations: the former can be used for the first pair of recurrence times and such an approach loses efficiency because bivariate times of higher orders are not used in the estimation; the latter approach can be applied to estimate the distribution of the sum of the bivariate recurrence times - however, noting that the sum does not serve as an appropriate outcome variable especially when the two recurrent events have opposite scientific meanings. Thus, naive applications of the existing methods could lead to either waste of the data information or failure to address the scientific problems.

## 1.2 MODEL AND DATA

Suppose a risk population of interest consists of subjects who experience a certain event (i.e., the initial event) in their lifetime. Assume a group of subjects were randomly sampled from the risk population and recruited into a follow-up study. Suppose the follow-up starts at the initial event and each subject's status alternated between two states over time. Denote by random variables  $V$  and  $W$  the durations of the two states, and by  $N$  the collection of the two-state recurrence times on a study subject, i.e.,  $N = \{(V_1, W_1), (V_2, W_2), \dots\}$ . Assume subjects are sampled independently, but correlation among recurrence times from the same subject is allowed. To accommodate such a characteristic, the following assumptions are adopted:

**Assumption 1.** There exists a latent variable  $Z$  so that, conditioning on  $Z$ , the bivariate random vectors  $(V_j, W_j)$ ,  $j = 1, 2, \dots$ , are identically and independently distributed (iid).

In other words, the bivariate random vectors  $(V_j, W_j)$ ,  $j = 1, 2, \dots$ , form an alternating renewal process. Note that the recurrence times  $V_j$  and  $W_j$  is allowed be correlated even conditioning on  $Z$ . The latent variable  $Z$  characterizes the association between bivariate vectors  $\{(V_j, W_j)\}$  within a subject. In this paper both the distribution of  $Z$  and the correlation between  $(V_j, W_j)$  on the individual level are not further modelled. Define the joint probability density function of  $(V_j, W_j)$  to be  $F_{VW}(v, w)$ , then

$$F_{VW}(v, w) = \int P(V \leq v, W \leq w \mid Z = z) dP_Z(z),$$

where  $P_Z(z)$  is the probability distribution function of  $Z$ . The observation on the bivariate recurrent event process  $\{(V_j, W_j) : j = 1, 2, \dots\}$  is subject to right censoring  $C$ , where  $C$  has a survival function  $G(\cdot)$ . We make the usual assumption of independent censorship:

**Assumption 2.**  $C$  is independent of  $(N, Z)$ .

Let  $m$  denote the index of censored bivariate recurrence time, that is,  $m$  satisfies

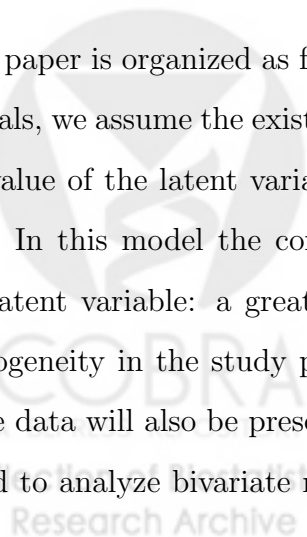
$$\sum_{j=1}^{m-1} (V_j + W_j) \leq C \text{ and } \sum_{j=1}^m (V_j + W_j) > C.$$

Clearly the index  $m$  is a random variable. Note that the last bivariate recurrence times  $(V_m, W_m)$  are subject to censoring with the censoring time  $C^* = C - \sum_{j=1}^{m-1} (V_j + W_j)$ , where  $V_m$  may or may not be completely observed, but  $W_m$  is always censored. The observed data are

$$\{(V_1, W_1), \dots, (V_{m-1}, W_{m-1}), (V_m^+, W_m^+)\},$$

where  $V_m^+ = \min(V_m, C^*)$  and  $W_m^+ = \min(W_m, \max(C^* - V_m, 0))$ . In the observed data it is possible that the last  $V$ -observation is uncensored and only the last  $W$  is censored, i.e., in this case,  $V_m^+ = V_m$ . It's easy to check that conditioning on  $m$ , the observed bivariate times  $\{(V_1, W_1), \dots, (V_{m-1}, W_{m-1})\}$  are identically but not independently distributed. Furthermore, the recurrence time  $V_m + W_m$  tends to be longer than earlier recurrence times due to intercepted sampling. Because the observation time of  $V_j + W_j$  does not exceed  $C$ , the joint distribution function  $F_{VW}(v, w)$  is identifiable only in the region  $\{(v, w) : v + w \leq \tau_c\}$ , where  $\tau_c$  is the maximal support of  $C$  defined by  $\tau_c = \sup\{t : G(t) > 0\}$ . The marginal distribution function of  $V$  and  $W$  cannot be obtained directly from the joint distribution function, unless  $\tau_c > \tau_V + \tau_W$ , where  $\tau_V$  and  $\tau_W$  are the maximal support of  $V$  and  $W$ , respectively.

This paper is organized as follows. In section 2, to characterize the heterogeneity among individuals, we assume the existence of a subject-specific latent variable so that, conditioning on the value of the latent variable, the bivariate recurrence times are iid bivariate random vectors. In this model the correlation between bivariate random vectors is characterized by the latent variable: a greater variability in the latent variable indicates a higher level of heterogeneity in the study population. A brief review of existing methods for bivariate gap time data will also be presented in section 2. In section 3 nonparametric estimators are proposed to analyze bivariate recurrence time data by combining techniques for univariate



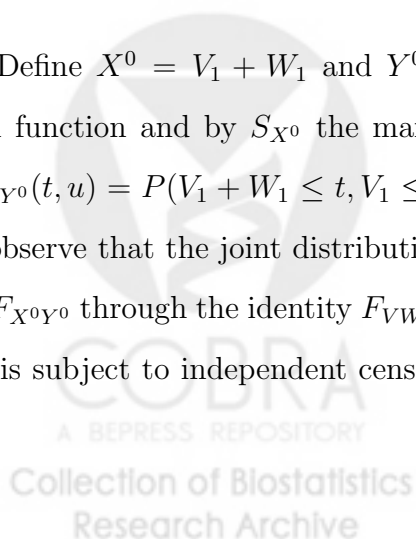
recurrent event data and techniques for bivariate gap time data. Large sample properties of the proposed estimators will be presented in the same section. In Section 4, simulation studies with practical sample sizes and the results of a data analysis using proposed methods will be reported. Conclusion remarks and discussions will be presented in the final section.

## 2 NONPARAMETRIC ESTIMATIONS OF $F_{VW}$ AND $F_V$

### 2.1 A BRIEF REVIEW OF SHL-METHOD

In a bivariate failure time model, nonparametric methods have been proposed (Wang and Wells, 1998; Huang and Louis, 1998; Lin et al., 1999) to estimate the joint distribution  $F_{VW}(v, w)$  based on censored bivariate failure times. Nevertheless, in the analysis of the alternating bivariate event data, their methods apply only to the first pair of bivariate times and therefore are expected to be inefficient. In this section a nonparametric estimator of Stute (1993) and Huang and Louis (1998), termed as the SHL-estimator, is briefly reviewed. As will be seen in Section 2.3, the SHL-estimator together with the technique for univariate recurrence times (Wang and Chang, 1999), termed as the WC-estimator, can be adopted as the base for the development of nonparametric methods for the estimation of  $F_{VW}(v, w)$  based on bivariate recurrence times with censoring.

Define  $X^0 = V_1 + W_1$  and  $Y^0 = (V_1, W_1)$ , and denote by  $F_{X^0Y^0}$  their joint distribution function and by  $S_{X^0}$  the marginal survival function of  $X^0$ ; that is, for  $u = (u_1, u_2)$ ,  $F_{X^0Y^0}(t, u) = P(V_1 + W_1 \leq t, V_1 \leq u_1, W_1 \leq u_2)$  and  $S_{X^0}(t) = F_{X^0Y^0}(t, (\infty, \infty))$ . It is easy to observe that the joint distribution function of  $(V_1, W_1)$ , and hence  $(V, W)$ , is determined by  $F_{X^0Y^0}$  through the identity  $F_{VW}(v, w) = F_{X^0Y^0}(v+w, (v, w))$ . Suppose the total duration  $X^0$  is subject to independent censoring time  $C$  and the value of  $Y^0$  is completely observed



only when  $X^0$  is uncensored. Further we define

$$X = \min(X^0, C), Y = Y^0 I(X^0 \leq C), \text{ and } \Delta = I(X^0 \leq C),$$

and denote by  $S_X$  the marginal survival function of  $X$ . Let  $Y^0 \leq u$  be defined componentwise and let  $F(t, u) = P(X^0 \leq t, Y^0 \leq u, \Delta = 1)$ .

Under the independent censorship assumption of  $(V_1, W_1)$  and  $C$ , it is easy to verify that  $F(dt, u) = F_{X^0 Y^0}(dt, u)G(t-)$  and  $S_X(t-) = S_{X^0}(t-)G(t-)$ . The bivariate distribution function  $F_{X^0 Y^0}$  can be expressed as

$$F_{X^0 Y^0}(t, u) = \int_0^t S_{X^0}(s-) \frac{F_{X^0 Y^0}(ds, u)}{S_{X^0}(s-)} = \int_0^t S_{X^0}(s-) \frac{F(ds, u)}{S_X(s-)} . \quad (1)$$

This representation was adopted by Huang and Louis (1998) for estimating the joint distribution of survival time and mark variables. When applied to the bivariate recurrence time data,  $Y^0 = (V_1, W_1)$  is considered as a bivariate mark vector of the recurrence time  $X^0 = V_1 + W_1$ , where  $X^0$  is subject to random censorship. Clearly, in this setting,  $(V_1, W_1)$  is observed when  $X^0$  is uncensored. In the construction of the estimator, the survival function  $S_{X^0}$  is estimated by the Kaplan-Meier estimator, and  $F$  and  $S_X$  are estimated by their corresponding empirical measures using only the first pair of possibly censored recurrence times. The estimator is identical to the estimator proposed in Example 1.1 by Stute (1993). We term this estimator the SHL-estimator. While the estimator studied by Wang and Wells (1998) is path-dependent and the one proposed by Lin et al. (1999) may give negative mass points, the SHL-estimator is path-independent and has the desirable monotonicity property. When bivariate recurrent event data in the current setting are observed, the SHL-estimator is inefficient because the second and higher order bivariate recurrence times are not used in the estimator. In the next section we propose an estimator of  $F_{VW}$  that makes better use of the bivariate recurrence time data.



## 2.2 THE PROPOSED ESTIMATORS

Let the subscript  $i$  be the index for a subject,  $i = 1, 2, \dots, n$ . Denote the underlying alternating recurrent event process of the  $i^{\text{th}}$  subject by  $N_i = \{(V_{i1}, W_{i1}), (V_{i2}, W_{i2}), \dots\}$ . Denote by  $m_i$  the number of completely observed bivariate recurrence times for subject  $i$ . We denote

$$X_{ij}^0 = V_{ij} + W_{ij}, \quad Y_{ij}^0 = (V_{ij}, W_{ij}),$$

$X_{ij} = \min(X_{ij}^0, C_i)$ ,  $Y_{ij} = Y_{ij}^0 I(X_{ij}^0 \leq C_i)$ , and  $\Delta_{ij} = I(X_{ij}^0 \leq C_i)$ . We further define the functions  $F_a(t, u) = E[a_i I(X_{i1} \leq t, Y_{i1} \leq u, \Delta_{i1} = 1)]$  and  $R_a(t) = E[a_i I(X_{i1} \geq t)]$ , where  $u = (u_1, u_2)$  is a vector of real numbers and  $a_i = a(C_i)$  is a non-negative function of  $C_i$  with  $E(a_i^2) < \infty$ . It is easy to verify that, under Assumptions 1 and 2,  $F_a(ds, u)/R_a(s) = F(ds, u)/S_X(s-)$ , and it follows (1) that

$$F_{X^0 Y^0}(t, u) = \int_0^t S_{X^0}(s-) \frac{F_a(ds, u)}{R_a(s-)} . \quad (2)$$

Since  $F_a$  and  $R_a$  are expectations of observed random variables, they can be estimated by the corresponding empirical measures. We define  $m_i^* = m_i - 1$  for  $m_i \geq 2$  and  $m_i^* = 1$  for  $m_i = 1$ . For  $u = (u_1, u_2)$ , let

$$\begin{aligned} \hat{F}_a(t, u) &= \frac{1}{n} \sum_{i=1}^n \frac{a_i I(m_i \geq 2)}{m_i^*} \sum_{j=1}^{m_i^*} I(X_{ij} \leq t, Y_{ij} \leq u) \\ &= \frac{1}{n} \sum_{i=1}^n \frac{a_i I(m_i \geq 2)}{m_i^*} \sum_{j=1}^{m_i^*} I(V_{ij} + W_{ij} \leq t, V_{ij} \leq u_1, W_{ij} \leq u_2), \end{aligned}$$

and

$$\hat{R}_a(t) = \frac{1}{n} \sum_{i=1}^n \frac{a_i}{m_i^*} \sum_{j=1}^{m_i^*} I(X_{ij} \geq t) = \frac{1}{n} \sum_{i=1}^n \frac{a_i}{m_i^*} \sum_{j=1}^{m_i^*} I(V_{ij} + W_{ij} \geq t).$$

$\hat{F}_a(t, u)$  and  $\hat{R}_a(t)$  are both moment-type estimators of  $F_a(t, u)$  and  $R_a(t)$ . These estimators are constructed using the exchangeability of the complete bivariate observations. To avoid

sampling bias, the last censored pairs of bivariate gap times are not used in either estimator unless  $m_i = 1$ . Next we define

$$\hat{H}_a(t, u) = \sum_{i=1}^n \frac{a_i \mathbb{I}(m_i \geq 2)}{m_i^*} \sum_{j=1}^{m_i^*} \mathbb{I}(X_{ij} = t, Y_{ij} \leq u)$$

and

$$\hat{\Lambda}_a(t, u) = \int_0^t \frac{\hat{F}_a(ds, u)}{\hat{R}_a(s)}.$$

For convenience we use  $\infty$  to represent  $(\infty, \infty)$  whenever necessary. It is easy to check that the nonparametric estimator proposed by Wang and Chang (1999) for the recurrent survival function  $S_{X^0}$  is equivalent to

$$\hat{S}_{X^0}(t) = \prod_{t_k^* \leq t} \left( 1 - \frac{\hat{H}_a(t_k^*, \infty)}{\hat{R}_a(t_k^*)} \right), \quad (3)$$

where  $t_1^*, t_2^*, \dots, t_K^*$  are the distinct and uncensored recurrence times from  $\{X_{ij}, j = 1, \dots, m_i^*, i = 1, \dots, n\}$ .

Following representations (2) and (3), a natural estimator for  $F_{X^0 Y^0}$  is given by

$$\hat{F}_{X^0 Y^0}(t, u) = \sum_{t_k^* \leq t} \prod_{j < k} \left( 1 - \frac{\hat{H}_a(t_j^*, \infty)}{\hat{R}_a(t_j^*)} \right) \frac{\hat{H}_a(t_k^*, u)}{\hat{R}_a(t_k^*)} \quad (4)$$

It follows from (4) that the bivariate recurrence time function can be estimated by:

$$\begin{aligned} \hat{F}_{VW}(v, w) &= \hat{F}_{X^0 Y^0}(v + w, (v, w)) \\ &= \sum_{t_k^* \leq v+w} \prod_{j < k} \left( 1 - \frac{\hat{H}_a(t_j^*, \infty)}{\hat{R}_a(t_j^*)} \right) \frac{\hat{H}_a(t_k^*, (v, w))}{\hat{R}_a(t_k^*)} \quad v + w \leq \tau_c, \end{aligned}$$

where  $\tau_c$  is the maximal support of  $C$ .  $F_{VW}$  is identifiable on the domain  $\{(v, w) : v + w \leq \tau_c\}$ .

It can be shown that the proposed estimator  $\hat{F}_{VW}$  has the desired monotonicity property and a smaller variance than the SHL-estimator based on the first pairs of gap time data.

Because  $F_{VW}(v, w)$  is identifiable only for  $v + w \leq \tau_c$ , in general, the marginal recurrence functions of  $V$  (and  $W$ ) cannot be estimated directly by  $F_{VW}(v, \infty)$  (or  $F_{VW}(\infty, w)$ ), but

techniques of the WC-estimator can be applied to the observed recurrence times,  $\{V_{i1}, \dots, V_{im_i^*}, i = 1, \dots, n\}$ , for the estimation of  $S_V$ . In the bivariate case, it is possible that, within the last censored pair,  $V_{im_i}$  is observed and its corresponding gap time  $W_{im_i}$  is censored. To avoid sampling bias, as long as  $m_i \geq 2$ ,  $V_{im_i}$  is not used in the proposed estimation even if it is not censored. To be specific, the marginal recurrent survival function  $S_V$  is identifiable on  $\{V \leq \tau_c\}$  and can be estimated by

$$\hat{S}_V(t) = \prod_{v^* \leq t} \left\{ 1 - \frac{\hat{H}_V(v^*)}{\hat{R}_V(v^*)} \right\},$$

where  $v_1^*, v_2^*, \dots, v_L^*$  are the distinct and uncensored recurrence times from  $\{V_{ij}, j = 1, \dots, m_i^*, i = 1, \dots, n\}$ , and where

$$\hat{H}_V(t) = \sum_{i=1}^n \frac{a_i I(m_i \geq 2)}{m_i^*} \sum_{j=1}^{m_i^*} I(V_{ij} = t)$$

and

$$\hat{R}_V(t) = \sum_{i=1}^n \frac{a_i}{m_i^*} \sum_{j=1}^{m_i^*} I(V_{ij} \geq t).$$

The recurrence function of  $W$  is generally not estimable due to induced dependent censoring (Gelber, Gelman, and Goldhirsch, 1989); however, it is possible to estimate the conditional distribution function

$$F_{W|V}(w | v) = P(W \leq w | V \leq v) = F_{VW}(v, w) / F_V(v),$$

for  $v + w \leq \tau_c$ . An estimator of  $F_{W|V}$  is given by  $\hat{F}_{VW}(v, w) / (1 - \hat{S}_V(v))$ . Estimation of such a conditional distribution function can be used to, say, detect possible correlation between recurrence times  $V$  and  $W$ .

### 3 ASYMPTOTIC PROPERTIES

Denote by  $\tau$  the maximal support of  $R_a = E[a_1 I(X_1 \geq t)]$ ; by definition  $\tau$  is smaller than the maximal support of  $C$ , i.e.  $\tau_c$ . Let  $L$  be any number smaller than  $\tau$ . To simplify

the discussion, we assume that  $G(t)$  is absolutely continuous on  $[0, L]$  and  $F_{X^0Y^0}(t, u)$  is absolutely continuous on  $\Omega = \{ (t, (u_1, u_2)) : 0 \leq u_1 + u_2 \leq t \leq L \}$ .

Define  $\Lambda_a(t, u) = \int_0^t F_a(ds, u)/R_a(s)$ . It can be verified that the survival function of  $X^0$  equals the product integral of  $1 - \Lambda_a(t, \infty)$ , i.e.,

$$S_{X^0}(t) = \prod_{[0,t]} \{1 - \Lambda_a(ds, \infty)\}. \quad (5)$$

Let  $\mathcal{D}(\Omega)$  denote the space of bivariate right-continuous functions on  $\Omega$  with left-hand limits. Following (2) and (5) we are able to define a mapping from  $\mathcal{D}(\Omega)$  to itself by  $\Phi : \Lambda_a \rightarrow F_{X^0Y^0}$ , where

$$F_{X^0Y^0}(t, u) = \Phi(\Lambda_a)(t, u) = \int_{[0,t]} \prod_{[0,s]} \{1 - \Lambda_a(ds, \infty)\} \Lambda_a(ds, u).$$

The mapping can be shown to be compactly differentiable with respect to the supremum norm at a given  $\Lambda_a$  with derivative

$$\begin{aligned} \{d\Phi(\Lambda_a) \cdot h\}(t, u) &= \int_0^t F_{X^0Y^0}(s, u) - F_{X^0Y^0}(t, u)h(ds, \infty) \\ &\quad + \int_0^t 1 - F_{X^0Y^0}(s, \infty)h(ds, \infty), \end{aligned}$$

where  $h \in \mathcal{D}(\Omega)$ . Define  $\hat{\Lambda}_a$  to be the estimator of  $\Lambda_a$  with  $R_a$  and  $F_a$  replaced by  $\hat{R}_a$  and  $\hat{F}_a$ , respectively. It is easy to check that the WC-estimator for the recurrent survival function  $S_{X^0}$  defined in (3) is equivalent to the product integral of  $1 - \hat{\Lambda}_a(t, \infty)$ , that is,

$$\hat{S}_{X^0}(t) = \prod_{[0,t]} \{1 - \hat{\Lambda}_a(ds, \infty)\}.$$

Moreover, the proposed nonparametric estimator of  $F_{X^0Y^0}$  equals  $\hat{F}_{X^0Y^0} = \Phi(\hat{\Lambda}_a)$ . To study the large sample properties of the proposed estimator, it is sufficient to study the asymptotic properties of  $\hat{\Lambda}_a$  and then apply the functional delta method to the mapping  $\Phi$ .

The weak convergence properties of  $\hat{R}_a$  and  $\hat{F}_a$  are expected because they are both sample-mean type estimators of  $R_a$  and  $F_a$ , respectively. A sketch of the proof of the weak

convergence properties for the two empirical processes,  $\hat{R}_a$  and  $\hat{F}_a$ , is given in Appendix. The large-sample properties of  $\hat{\Lambda}_a$  are given in Theorem 1 and the asymptotic properties of the proposed estimator  $\hat{F}_{X^0Y^0}$  are explored in Theorem 2, both with proofs given in the Appendix.

**Theorem 1.** Assume  $a(c)$  is a bounded function on  $[0, L]$  and define  $a_i = a(C_i)$ . For any  $L < \tau$  and  $(t, u) = (t, (u_1, u_2)) \in \Omega$ , the stochastic process  $\sqrt{n} (\hat{\Lambda}_a(t, u) - \Lambda_a(t, u))$  has an asymptotically iid representation

$$\sqrt{n} (\hat{\Lambda}_a(t, u) - \Lambda_a(t, u)) = \frac{1}{\sqrt{n}} \sum_{i=1}^n \psi_i(t, u) + o_p(1)$$

which converges weakly to a Gaussian process  $\mathcal{U}(t, u)$  with mean 0 and variance-covariance function  $E[\psi_1(t_1, u_1)\psi_1(t_2, u_2)]$ . The function  $\psi_i$  is defined in Appendix.

The variance-covariance function,  $E[\psi_i(t_1, u_1)\psi_i(t_2, u_2)]$ , of the limiting distribution  $\mathcal{U}(t, u)$  can be consistently estimated by  $n^{-1} \sum_{i=1}^n \hat{\psi}_i(t_1, u_1)\hat{\psi}_i(t_2, u_2)$ , where  $\hat{\psi}_i$  is the corresponding estimator of  $\psi_i$  with  $(F_a, R_a)$  replaced by  $(\hat{F}_a, \hat{R}_a)$ .

**Theorem 2.** Assume  $a$  is a bounded function on  $[0, L]$  and define  $a_i = a(C_i)$ . For any  $L < \tau$  and  $(t, u) \in \Omega$ , the stochastic process  $\sqrt{n}\{\hat{F}_{X^0Y^0}(t, u) - F_{X^0Y^0}(t, u)\}$  has an asymptotically iid representation

$$\sqrt{n}(\hat{F}_{X^0Y^0}(t, u) - F_{X^0Y^0}(t, u)) = \frac{1}{\sqrt{n}} \sum_{i=1}^n \phi_i(t, u) + o_p(1)$$

which converges weakly to a mean zero Gaussian process with the variance-covariance function  $E[\phi_1(t_1, u_1)\phi_1(t_2, u_2)]$ , where  $(t_j, u_j) \in \Omega$ ,  $j = 1, 2$ , and the function  $\phi_i$  is defined in Appendix.

The variance-covariance function,  $E[\phi_1(t_1, u_1)\phi_1(t_2, u_2)]$ , can be consistently estimated by  $n^{-1} \sum_{i=1}^n \hat{\phi}_i(t_1, u_1)\hat{\phi}_i(t_2, u_2)$ , where  $\hat{\phi}_i$  is the corresponding estimator of  $\phi_i$  with  $(F_a, R_a)$  replaced by  $(\hat{F}_a, \hat{R}_a)$ . It can be further shown that the asymptotic variance of  $\hat{S}_{X^0}$  given in

Wang and Chang (1999) is identical to  $E[\phi_1(t_1, \infty)\phi_1(t_2, \infty)]$ . Moreover, Theorem 2 provides a more comprehensive formula for the asymptotic covariance structure of the SHL-estimator by forcing  $m_i^* = 1$  and  $a_i = 1$  for all  $i = 1, \dots, n$ .

It follows from Theorem 2 that, for  $0 \leq v+w \leq L$ ,  $\sqrt{n}(\hat{F}_{VW}(v, w) - F_{VW}(v, w))$  converges weakly to a mean zero Gaussian process with variance-covariance function

$$E[\phi_1(v_1 + w_1, (v_1, w_1))\phi_1(v_2 + w_2, (v_2, w_2))]$$

for  $0 \leq v_j + w_j \leq L$ ,  $j = 1, 2$ .

Finally, the large sample properties of  $\hat{S}_V$  can be studied in a way similar to WC-estimator. Define the functions  $H_V(t) = E[a_i I(V_{i1} \leq t)I(V_{i1} \leq C_i)]$ ,  $R_V(t) = E[a_i I(V_{i1} \geq t)I(C_i \geq t)]$ , and

$$\xi_i(t) = \frac{a_i I(m_i \geq 2)}{m_i^*} \sum_{j=1}^{m_i^*} \frac{I(V_{ij} \leq t)}{R_V(V_{ij})} - \int_{[0,t]} \frac{a_i}{m_i^*} \sum_{j=1}^{m_i^*} I(V_{ij} \geq s) \frac{H_V(ds)}{R_V(s)^2}.$$

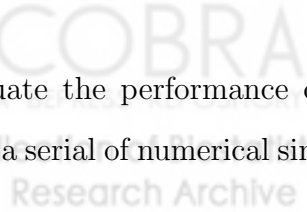
**Theorem 3.** Assume  $a$  is a bounded function on  $[0, L]$  and define  $a_i = a(C_i)$ . For any  $t < \tau$ , the stochastic process  $\sqrt{n}\{\hat{S}_V(t) - S_V(t)\}$  which converges weakly to a mean zero Gaussian process with the variance-covariance function  $S_V(t_1)S_V(t_2)E[\xi_1(t_1)\xi_1(t_2)]$ .

The proof of Theorem 3 closely follows Theorem 1 in Wang and Chang (1999), and thus is omitted in this paper.

## 4 SIMULATIONS AND DATA ANALYSIS

### 4.1 MONTE-CARLO SIMULATIONS

To evaluate the performance of the proposed estimator under moderate sample size, we conduct a serial of numerical simulation studies. In each simulation study the latent variable,



$Z$ , is generated from either a uniform(0,2) distribution or an exponential distribution with mean 1. Note that the defined population under  $Z \sim \text{uniform}(0,2)$  is more homogenous than those under  $Z \sim \text{exp}(1)$  because uniform(0,2) have smaller variance. Given the value of the latent variable,  $Z = z$ , the iid bivariate recurrence times are generated from Clayton's multivariate failure time distribution (Clayton, 1978; Oakes 1982) with joint survivorship function

$$S_{VW}(v, w | z) = \left( S_1(v | z)^{1-\theta} + S_2(w | z)^{1-\theta} - 1 \right)^{\frac{1}{1-\theta}}, \quad \theta \geq 1,$$

where

$$S_1(s | z) = Pr(V > s | z) = \exp(-e^z s^2),$$

and

$$S_2(t | z) = Pr(W > t | z) = \exp(-e^{-z} t^{1.5}).$$

Thus, on the individual level, the degree of association between the bivariate recurrence times is determined by the value of  $\theta$ .

Eight sets of simulations were carried out. We set  $\theta$  to be 3 and 9 so that the corresponding Kendall's  $\tau$  coefficient of concordance for the bivariate recurrence times are 0.5 and 0.8, which indicate mild and strong association among bivariate recurrence times on the individual level, respectively. The observation of bivariate recurrence processes is terminated by the censoring time  $C$ . We set  $C$  to follow a uniform(0,15) distribution and a uniform(0,8) distribution, so that the former scenario has longer censoring times allowing for the occurrence of more bivariate recurrent events than the latter scenario.

In each simulation study 1000 samples are generated, each with 200 subjects. In the scenarios where  $C \sim \text{uniform}(0, 8)$ , the proportion of subjects having at least one pair of recurrence time is approximately 70% when  $Z \sim \text{uniform}(0, 1)$  and 68% when  $Z \sim \text{exp}(1)$ ; the average number of observed bivariate recurrence times (censored or uncensored),  $m_i$ , is approximately 2.5 under both distributions of  $Z$ . In the scenarios when  $C \sim \text{uniform}(0, 15)$ ,

the proportion of having at least one complete pair of recurrence time is approximately 84% when  $Z \sim \text{uniform}(0,1)$  and is 81% when  $Z \sim \text{exp}(1)$ , while the average  $m_i$  is approximately 4.0 under either distributions of  $Z$ .

Three different estimation methods were used to estimate  $F_{VW}$ : the first two estimators are the proposed estimator with different weight functions  $a(c) = c$  and  $a(c) = 1$ , and the last estimator is the SHL-estimator applied to the first pair of bivariate recurrence times. Tables 1 through 4 summarize the simulation results for these three estimators of  $F_{VW}(v, w)$  at 16 selected grid points  $(v, w)$ , where  $v$  takes values 0.5, 0.7, 1, and 1.5, and  $w$  takes values 1, 2, 3, and 4. It can be observed in the tables that all three estimation methods work reasonably well, i.e., the averages of the estimates over 1000 simulations are very close to their true values. In consideration of efficiency, the proposed estimator, with censoring time as the weight, outperforms the other two estimators when more bivariate recurrent events are observed. In the case when fewer pairs of recurrence times are observed, the proposed estimator with identical weights is still a better estimator than the SHL-estimator in the sense that it has smaller standard errors. It is also illustrated in the tables that the standard deviations of all three estimators are smaller when the defined population is more homogeneous.

Finally, in Section 2.2, the proposed nonparametric estimator of the marginal survival function  $S_V$  does not use the gap time  $V_{im_i}$  from the last pair of recurrence times in order to avoid sampling bias. To further illustrate this point we conducted a simulation study under a similar setting as above, with large sample size 1000, uniform latent variable distribution on  $[0, 2]$ , moderate association  $\theta = 3$ , and uniform censoring distribution on  $[0, 15]$ . Figure 1 shows the estimated survival curve using the proposed method with and without uncensored  $V_{im_i}$ . It is observed that the proposed estimator of  $S_V$ , with  $V_{im_i}$  excluded, is very close to the true survival curve, while the estimator using the last uncensored  $V_{im_i}$  overestimates the survival probability. It is known that, due to intercept sampling, the combined length of



last censored bivariate recurrence times,  $V_{im_i} + W_{im_i}$ , tends to be longer than uncensored ones, and the bias applies to both  $V_{i,m_i}$  and  $W_{i,m_i}$ . Thus, as described in Section 2.2, the last uncensored  $V_{im_i}$  should be disregarded in order to construct consistent estimator for the marginal survival function.

## 4.2 DATA ANALYSIS

South Verona, Italy, is an urban area with a population of about 75,000. Following the provisions of the Italian psychiatric reform of 1978, all admissions to mental hospitals were stopped and in 1979 a well-integrated South-Verona Community Mental Health Service (CMHS) was established to provide psychiatric care to defined geographical areas. The CMHS is run by the Section of Psychiatry, Department of Medicine and Public Health, and it includes a Community Mental Health Center providing day care, rehabilitation and home visits, a psychiatric unit in a general hospital with 15 beds, sheltered apartments, outpatient services, and 24-hour crisis intervention service. The CMHS covers 80% of those receiving psychiatric care, and only a small proportion of service is provided by private hospitals.

The South-Verona Psychiatric Case Register (PCR) has collected information from all psychiatric services in South-Verona since 1978 (Tansella, 1991). A subsample of South-Verona PCR data including 336 incident cases of schizophrenia and related disorders (ICD10 codes: F20 through F29; F84) from 1981 to 1995 is used to illustrate the proposed estimator. An incident case is defined as the first-ever contact in the register, with one of the health care facilities reporting to the PCR.

The definition of PCR episodes of care (PCR-EC) was used in Sturt et al. (1982) and Tansella (1995) to evaluate the usage of mental health services. Under their definition, a psychiatric patient, at each time point after the first contact in PCR, is always in one of the two states: PCR-EC or PCR-break. A PCR-break is a break between consecutive mental

health service usages greater than 90 days. A PCR-EC starts with a psychiatric contact and ends when a break between contacts greater than 90 days occurs. Estimation of the joint distribution of PCR-EC and PCR-break is an important tool for investigating the usage pattern of mental health services.

Table 5 summarizes the number of bivariate recurrence times, censored or uncensored, from these 336 psychiatric patients. The number of bivariate recurrence times ranges from 1 to 18. The majority (43.2%) of the study population has only one incomplete pair of PCR-EC and PCR-break. In total there are 1035 bivariate recurrence times from all patients, and the SHL-estimator only uses about one third (336/1035) of the data information. Table 6 shows the estimates of the cumulative joint distribution function and the corresponding bootstrap standard errors, using the proposed estimator with identity weight, at 36 selected bivariate time points. About 50% of the bivariate recurrence times have a PCR-EC less than or equal to 12 months and a PCR-break less than or equal to 48 months.

It would be of interest to study the marginal distribution of PCR-EC. The estimated marginal recurrent survival function is given in Figure 5. The median recurrence time of PCR-EC is estimated to be 81 days or 2.7 months, and 80% of the recurrence times are within 317 days or 10.4 months, suggesting that the distribution of PCR-EC is heavily right skewed. The conditional distribution of PCR-break given PCR-EC can be used to investigate the association between PCR-EC and PCR-break. In Figure 5 the 30%, 60%, and 90% quantiles of the marginal distribution of PCR-EC are estimated to be 25, 123, and 707 days, respectively. We estimate the distribution function of PCR-break condition on the three PCR-EC subgroups: PCR-EC  $\leq$  25 days, PCR-EC between 26 and 123 days, and PCR-EC between 124 and 707 days. Figure 3 shows the three cumulative conditional distribution functions of PCR-break by PCR-EC subgroups. It is observed that, given longer PCR-EC, the chance of opening a new episode of care within a certain time period of break is higher; therefore a negative association between the two recurrence times is present. This

is what we expected because patients receiving a longer duration of PCR-EC tend to be sicker ones, and therefore have a greater chance to have a new episode of care while they are not receiving health care.

## 5 DISCUSSIONS

In this paper, we present a nonparametric estimator for the distribution function of the bivariate recurrence times by combining techniques for univariate recurrence time data and techniques for bivariate gap time data. The proposed estimator does not require assumptions on the distribution of the latent variable  $Z$  in Assumption 1, and the use of weight  $a_i = a(C_i)$  in the estimator could improve the efficiency of the estimator. The choice of the optimal weight, however, does not seem to have a closed-form expression and could vary for different values of (bivariate) time points. Based on our experience, assigning the weight function  $a_i = C_i$  produces satisfactory efficiency results when three or more pairs of recurrent events can be potentially observed.

The proposed estimator relies on the conditionally iid assumption on the bivariate recurrence times. When the iid assumption fails to hold the uncensored bivariate recurrence times are not exchangeable, and, as the result, the proposed method is not valid. For the purpose of model checking, it is essential to generalize the trend analysis for univariate recurrence times, proposed by Wang and Chen (2000), to the bivariate case.

It is worth mentioning that in the section of data analysis we investigated the association between PCR-EC and PCR-break in the South-Verona PCR data through conditional distribution functions; however, the degree of association is not quantified using this approach. Subject to unidentifiability of the distribution function, a global measurement for the association can not be obtained directly from the estimated bivariate distribution function. It would be desirable to develop a general measurement for the association between

the bivariate recurrence times.

## REFERENCES

- Andersen, P. K., and Gill, R. D. (1982), “Cox’s Regression Model for Counting Processes: A Large Sample Study,” *The Annals of Statistics*, **10**, 1100–1120.
- Billingsley, P. (1999). *Convergence of Probability Measures*. Second edition. New York: John Wiley and Sons.
- Breslow, N.E. and Crowley, J. (1974). A large sample study of the life table and product limit estimates under random censorship. *The Annals of Statistics* **2**, 437–453.
- Chang, S.-H. (2000). A two-sample comparison for multiple ordered event data. *Biometrics* **56**, 183–189.
- Clayton, D.G. (1978). A model for association in bivariate lifetables and its application in epidemiological studies of family tendency in chronic disease incidence. *Biometrika* **65**, 141–151.
- Gelber, R. D., Gelman, R. S. and Goldhirsch, A. (1989). A quality-of-life oriented endpoint for comparing therapies. *Biometrics*, **45**, 781-795.
- Huang, Y.(1999). The two-sample problem with induced dependent censorship. *Biometrics* **55**, 1108–1113.
- Huang, Y. and Louis, T.A. (1998). Nonparametric estimation of the joint distribution of survival time and mark variable. *Biometrika* **85**, 785–798.
- Lin, D.Y. (2000). Linear regression analysis of censored medical costs. *Biostatistics* **1**, 35–47.

- Lin, D.Y., Sun, W., and Ying, Z. (1999). Nonparametric estimation of gap time distributions for serial events with censored data. *Biometrika* **86**, 59–70.
- Lin, D. Y., Wei, L. J., Yang, I., and Ying, Z. (2000), “Semiparametric Regression for the Mean and Rate Functions of Recurrent Events,” *Journal of the Royal Statistical Society B*, **62**, 711–730.
- Oakes, D. (1989). Bivariate survival models induced by frailties. *Journal of the American Statistical Association* **84**, 487–493.
- Pena, E.A., Strawderman, R.L., and Hollander, M. (2001). Nonparametric estimation with recurrent event data. *Journal of the American Statistical Association* **96**, 1299–1315.
- Pepe, M. S., and Cai, J. (1993), “Some Graphical Displays and Marginal Regression Analyses for Recurrent Failure Times and Time Dependent Covariates,” *Journal of the American Statistical Association*, **88**, 811-820.
- Prentice R. L., Williams, B. J., and Peterson, A. V. (1981), “On the Regression Analysis of Multivariate Failure Time Data,” *Biometrika*, **68**, 373-379.
- Sturt, E., Wykes, T., and Creer, C. (1982). Demographic, social and clinical characteristics of the sample. In Long-Term Community Care: Experience in a London Borough (ed. J. K. Wing), pp. 5-14. Psychological Medicine (Monograph Supplement 2). Cambridge: Cambridge University Press.
- Stute, W. (1993). Consistent estimation under random censorship when covariates are present. *Journal of Multivariate Analysis* **45**, 89–103.
- Tansella, M. (1991). *Community-Based Psychiatry. Long-Term Patterns of Care in South-Verona. Psychological Medicine (Monograph Supplement 19)*. Cambridge: Cambridge University Press.

- Tansella, M., Micciolo, R., Biggeri, A., Bisoffi, G., and Balestrieri, M. (1995). Episode of Care for First-Ever Psychiatric Patients. A Long-Term Case-Register Evaluation in a Mainly Urban Area. *British Journal of Psychiatry* **167**, 220–227.
- van der Vaart, A.W. (1998). *Asymptotic Statistics*. Cambridge: Cambridge University Press.
- Visser, M. (1996). Nonparametric estimation of the bivariate survival function with application to vertically transmitted AIDS. *Biometrika* **83**, 507–518.
- Wang, M.-C. and Chang, S.-H. (1999). Nonparametric estimation of a recurrent survival function. *Journal of the American Statistical Association* **94**, 146–153.
- Wang, M.-C, Qin, J., and Chiang, C.-T. (2001), “Analyzing Recurrent Event Data With Informative Censoring,” *Journal of the American Statistical Association*, **96**, 1057–1065.
- Wang, W.-J. and Wells, M.T. (1998). Nonparametric estimation of successive duration times under dependent censoring. *Biometrika* **85**, 561–572.
- Wang, M.-C. and Chen, Y.Q. (2000). Nonparametric and semiparametric trend analysis for stratified recurrence time data. *Biometrics* **56**, 789–794.

## Appendix

In order to study the large sample properties of  $\hat{\Lambda}$ , it is helpful to establish the weak convergence of the two empirical processes  $\hat{F}_a$  and  $\hat{R}_a$ . The empirical processes  $\sqrt{n} \{ \hat{F}_a(t, u) - F_a(t, u) \}$  and  $\sqrt{n} \{ \hat{R}_a(t) - R_a(t) \}$  are sums of a sequence of normalized iid processes given by

$$\sqrt{n} \{ \hat{F}_a(t, u) - F_a(t, u) \} = \frac{1}{\sqrt{n}} \sum_{i=1}^n \left\{ \frac{a_i I(m_i \geq 2)}{m_i^*} \sum_{j=1}^{m_i^*} I(x_{ij} \leq t, y_{ij} \leq u) - F_a(t, u) \right\}$$

$$\sqrt{n} \left\{ \hat{R}_a(t) - R_a(t) \right\} = \frac{1}{\sqrt{n}} \sum_{i=1}^n \left\{ \frac{a_i}{m_i^*} \sum_{j=1}^{m_i^*} I(x_{ij} \geq t) - R_a(t) \right\}.$$

Finite-dimensional weak convergence for these two stochastic processes is expected from the Central Limit Theorem. Additionally, the tightness of the sequences of distributions induced by  $\sqrt{n}\{\hat{R}_a(t) - R_a(t)\}$  and  $\sqrt{n}\{\hat{F}_a(t, u) - F_a(t, u)\}$  follows from the two inequalities:

$$\begin{aligned} n^2 E \left[ \left\{ \hat{R}_a(t) - R_a(t) - \hat{R}_a(t') + R_a(t') \right\}^2 \left\{ \hat{R}_a(t'') - R_a(t'') - \hat{R}_a(t) + R_a(t) \right\}^2 \right] \\ \leq \text{constant} \times (R_a(t) - R_a(t'))(R_a(t'') - R_a(t)), \end{aligned}$$

and

$$\begin{aligned} n^2 E \left[ \left\{ \hat{F}_a(t, u) - F_a(t, u) - \hat{F}_a(t', u') + F_a(t', u') \right\}^2 \times \right. \\ \left. \left\{ \hat{F}_a(t'', u'') - F_a(t'', u'') - \hat{F}_a(t, u) + F_a(t, u) \right\}^2 \right] \\ \leq \text{constant} \times (F_a(t, u) - F_a(t', u'))(F_a(t'', u'') - F_a(t, u)), \end{aligned}$$

for  $0 \leq t' \leq t \leq t'' \leq L$ ,  $0 \leq u'_1 \leq u_1 \leq u''_1 \leq L$ , and  $0 \leq u'_2 \leq u_2 \leq u''_2 \leq L$  (Billingsley, 1999). Consequently,  $\sqrt{n} \left( \hat{F}_a(t, u) - F_a(t, u), \hat{R}_a(t) - R_a(t) \right)$  induces a tight sequence of distributions on  $\mathcal{D}(\Omega) \times \mathcal{D}^-([0, L])$ , where  $\mathcal{D}^-([0, L])$  is the space of left-continuous functions on  $[0, L]$  with right-hand limits. Therefore  $\sqrt{n}(\hat{F}_a(t, u) - F_a(t, u), \hat{R}_a(t) - R_a(t))$  converges weakly to a zero mean bivariate Gaussian process  $(\mathcal{F}_a(t, u), \mathcal{R}_a(t))$ .

### Proof of Theorem 1.

It is easy to show that the mapping from  $(F_a, R_a)$  to  $\Lambda_a$ , defined by

$$\Lambda_a(t, u) = \int_0^t \frac{F_a(ds, u)}{R_a(s)},$$

is compactly differentiable. Following the weak convergence of  $(F_a, R_a)$  and applying the functional delta method (Chapter 20, van der Vaart 1998), one can obtain the asymptotically iid representation for  $\sqrt{n}(\hat{\Lambda}_a(t, u) - \Lambda_a(t, u))$  as

$$\begin{aligned}
& \sqrt{n}\{\hat{\Lambda}_a(t, u) - \Lambda_a(t, u)\} \\
&= \sqrt{n} \left\{ \int_{[0,t]} \frac{\hat{F}_a(ds, u)}{\hat{R}_a(s)} - \int_{[0,t]} \frac{F_a(ds, u)}{R_a(s)} \right\} \\
&= \int_{[0,t]} \frac{\sqrt{n}\{\hat{F}_a(ds, u) - F_a(ds, u)\}}{R_a(s)} - \int_{[0,t]} \frac{\sqrt{n}\{\hat{R}_a(s) - R_a(s)\}}{R_a(s)^2} F_a(ds, u) + o_p(1). \\
&= \int_{[0,t]} \frac{\sqrt{n}\hat{F}_a(ds, u)}{R_a(s)} - \int_{[0,t]} \frac{\sqrt{n}\hat{R}_a(s)}{R_a(s)^2} F_a(ds, u) + o_p(1) \\
&= \frac{1}{\sqrt{n}} \sum_{i=1}^n \psi_i(t, u) + o_p(1),
\end{aligned}$$

where the iid random variables  $\psi_i(t, u)$  are defined as

$$\psi_i(t, u) = \frac{a_i I(m_i \geq 2)}{m_i^*} \sum_{j=1}^{m_i^*} \frac{I(x_{ij} \leq t, y_{ij} \leq u)}{R_a(x_{ij})} - \int_{[0,t]} \frac{a_i}{m_i^*} \sum_{j=1}^{m_i^*} I(x_{ij} \geq s) \frac{F_a(ds, u)}{R_a(s)^2}.$$

The finite dimensional weak convergence of  $\sqrt{n}\{\hat{\Lambda}_a(t, u) - \Lambda_a(t, u)\}$  follows from the Central Limit Theorem, and, in a manner similar to the arguments in Breslow and Crowley (1974), tightness follows from the weak convergence of  $\sqrt{n}\{\hat{F}_a(t, u) - F_a(t, u), \hat{R}_a(t) - R_a(t)\}$ . This completes the proof of Theorem 4.1.

### Proof of Theorem 4.2.

In Theorem 4.1 we have established the large sample properties of  $\hat{\Lambda}_a$ . Since the mapping  $\Phi : \Lambda_a \rightarrow F_{X^0 Y^0}$  is continuous and compactly differentiable with respect to the supremum



norm at a given  $\Lambda_a$ , we apply the functional delta method to  $\sqrt{n}\{\Phi(\hat{\Lambda}_a)(t, u) - \Phi(\Lambda_a)(t, u)\}$  and derive its asymptotically iid representation as the following:

$$\begin{aligned}
\sqrt{n}\{\hat{F}_{X^0Y^0}(t, u) - F_{X^0Y^0}(t, u)\} &= \sqrt{n}\{\Phi(\hat{\Lambda}_a) - \Phi(\Lambda_a)\} \\
&= d\Phi_{\Lambda_a}\left(\sqrt{n}\{\hat{\Lambda}_a - \Lambda_a\}\right)(t, u) + o_p(1) \\
&= \int_{[0,t]} \{F_{X^0Y^0}(s, u) - F_{X^0Y^0}(t, u)\} \sqrt{n}\{\hat{\Lambda}_a(ds, \infty) - \Lambda_a(ds, \infty)\} \\
&\quad + \int_{[0,t]} S_{X^0}(s) \sqrt{n}\{\hat{\Lambda}_a(ds, u) - \Lambda_a(ds, u)\} + o_p(1) \\
&= \frac{1}{\sqrt{n}} \sum_{i=1}^n \left\{ \int_{[0,t]} F_{X^0Y^0}(s, u) \psi_i(ds, \infty) + \int_{[0,t]} S_{X^0}(s) \psi_i(ds, u) \right. \\
&\quad \left. - F_{X^0Y^0}(t, u) \psi_i(t, u) \right\} + o_p(1) \\
&= \frac{1}{\sqrt{n}} \sum_{i=1}^n \phi_i(t, u) + o_p(1),
\end{aligned}$$

where the iid random variables  $\phi_i$  are defined as

$$\phi_i(t, u) = \int_{[0,t]} F_{X^0Y^0}(s, u) \psi_i(ds, \infty) + \int_{[0,t]} S_{X^0}(s) \psi_i(ds, u) - F_{X^0Y^0}(t, u) \psi_i(t, u).$$

The finite-dimensional weak convergence of  $\sqrt{n}\{\hat{F}_{X^0Y^0}(t, u) - F_{X^0Y^0}(t, u)\}$  follows from the central limit theorem, and its tightness follows from arguments similar to those of Breslow and Crowley (1974) and from the weak convergence of  $\sqrt{n}\{\hat{\Lambda}(t, u) - \Lambda(t, u)\}$ . This completes the proof of weak convergence in Theorem 4.2.

Table 1: *Simulation Summary Statistics:  $Z \sim \text{uniform}(0, 2)$  and  $\theta = 3$*

		$C \sim \text{uniform}(0, 8)$				$C \sim \text{uniform}(0, 15)$			
		w=1	w=2	w=3	w=4	w=1	w=2	w=3	w=4
0.5	(a)	0.218	0.374	0.450	0.486	0.218	0.374	0.450	0.486
	(b)	-0.3, 16	-0.0, 23	-0.4, 29	-0.9, 34	-0.0, 12	-0.6, 18	-0.3, 21	-0.2, 25
	(c)	-0.4, 17	0.1, 25	-0.1, 30	-0.7, 34	-0.4, 14	-0.8, 21	-0.4, 25	-0.2, 28
	(d)	0.4, 23	0.5, 31	1.1, 36	0.4, 40	0.2, 23	0.3, 31	0.5, 37	0.5, 38
0.7	(a)	0.281	0.496	0.607	0.664	0.281	0.496	0.607	0.664
	(b)	-0.7, 18	0.2, 26	0.4, 31	-0.4, 34	0.1, 13	-0.0, 20	0.1, 22	-0.5, 24
	(c)	-0.9, 19	0.5, 27	0.8, 32	0.0, 34	-0.3, 16	-0.1, 23	0.0, 26	-0.5, 26
	(d)	-0.4, 25	0.4, 34	2.0, 38	1.0, 38	-0.1, 25	0.6, 33	0.5, 38	-0.1, 37
1	(a)	0.322	0.593	0.735	0.809	0.322	0.593	0.735	0.809
	(b)	-0.8, 19	0.1, 28	0.5, 32	0.9, 31	0.2, 14	0.0, 22	0.5, 23	-0.2, 21
	(c)	-0.9, 20	0.4, 30	0.9, 33	1.3, 31	-0.2, 17	-0.1, 25	0.2, 26	-0.2, 23
	(d)	-0.3, 26	0.2, 35	1.6, 38	1.6, 34	0.1, 26	0.6, 35	0.0, 37	-0.9, 32
2	(a)	0.333	0.642	0.814	0.903	0.333	0.642	0.814	0.903
	(b)	-0.7, 19	0.4, 29	0.6, 34	0.8, 27	0.3, 14	0.1, 23	0.7, 25	-0.2, 19
	(c)	-0.8, 20	0.6, 30	0.8, 34	0.8, 27	-0.2, 16	0.0, 25	0.2, 27	-0.1, 20
	(d)	-0.1, 26	0.4, 35	1.0, 38	0.8, 28	0.2, 26	0.7, 36	0.0, 37	-0.7, 24

(a) True value of  $F_{VW}(v, w)$ ; (b)–(d) Monte-Carlo bias  $\times 10^3$  and standard error  $\times 10^3$  of the proposed estimator with weight  $c$ , the proposed estimator with identity weight, and the SHL-estimator applied to the first pairs of bivariate recurrence times.

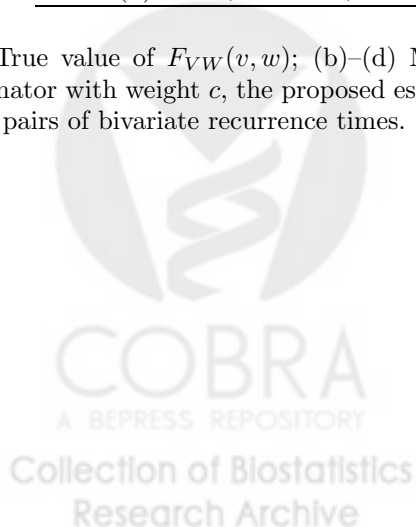


Table 2: Simulation Summary Statistics:  $Z \sim \text{uniform}(0, 2)$  and  $\theta = 9$

		$C \sim \text{uniform}(0, 8)$				$C \sim \text{uniform}(0, 15)$			
		w=1	w=2	w=3	w=4	w=1	w=2	w=3	w=4
v=									
0.5	(a)	0.254	0.397	0.465	0.495	0.254	0.397	0.465	0.495
	(b)	-0.8, 25	-0.1, 30	0.8, 33	-0.1, 35	0.0, 18	0.1, 22	0.8, 24	0.5, 25
	(c)	-1, 26	-0.3, 31	0.6, 34	-0.4, 35	0.3, 21	0.4, 26	0.9, 27	0.5, 27
	(d)	-1.9, 32	-0.1, 38	0.8, 39	-0.5, 40	1.1, 31	2.3, 36	2.3, 37	1.5, 36
0.7	(a)	0.307	0.511	0.614	0.668	0.307	0.511	0.614	0.668
	(b)	-0.7, 28	0.0, 33	0.7, 34	-0.0, 35	0.4, 20	0.5, 23	0.9, 24	0.7, 24
	(c)	-0.8, 29	-0.1, 34	0.4, 34	-0.2, 35	0.8, 23	0.9, 27	1.1, 27	0.9, 26
	(d)	-1.4, 35	-0.4, 40	0.1, 39	-0.6, 40	1.2, 33	2.4, 37	2.4, 36	1.4, 34
1	(a)	0.332	0.600	0.738	0.811	0.332	0.600	0.738	0.811
	(b)	0.0, 30	0.5, 35	1.4, 33	1.1, 32	0.3, 22	0.4, 25	1.1, 23	1.0, 21
	(c)	0.0, 31	0.9, 36	1.5, 33	1.1, 32	0.6, 24	0.7, 27	1.2, 25	0.9, 23
	(d)	-0.6, 36	1.0, 41	1.9, 37	1.6, 36	0.6, 34	0.9, 37	1.0, 33	-0.2, 30
2	(a)	0.333	0.642	0.814	0.903	0.333	0.642	0.814	0.903
	(b)	0.0, 30	0.3, 36	1.4, 32	1.8, 29	0.3, 22	0.3, 26	0.9, 23	0.6, 19
	(c)	0.0, 31	0.6, 36	1.6, 31	1.9, 28	0.6, 24	0.6, 28	1.0, 25	0.5, 21
	(d)	-0.6, 36	0.0, 40	1.2, 33	1.6, 29	0.7, 34	1.0, 37	1.7, 31	-0.1, 25

(a) True value of  $F_{VW}(v, w)$ ; (b)–(d) Monte-Carlo bias  $\times 10^3$  and standard error  $\times 10^3$  of the proposed estimator with weight  $c$ , the proposed estimator with identity weight, and the SHL-estimator applied to the first pairs of bivariate recurrence times.

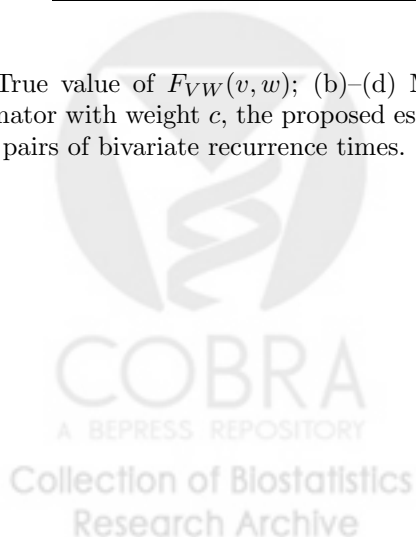


Table 3: *Simulation Summary Statistics:  $Z \sim \exp(1)$  and  $\theta = 3$*

		$C \sim \text{uniform}(0, 8)$				$C \sim \text{uniform}(0, 15)$			
		w=0.5	w=1	w=2	w=4	w=0.5	w=1	w=2	w=4
v=	(a)	0.206	0.318	0.369	0.394	0.206	0.318	0.369	0.394
	(b)	2.6, 22	3.9, 28	1.6, 31	4.7, 33	2.1, 18	2.8, 21	0.4, 24	3.5, 25
	(c)	2.9, 24	4.1, 30	1.7, 33	4.8, 34	1.8, 20	2.1, 24	-0.3, 26	2.7, 27
	(d)	2.4, 30	4.1, 36	1.2, 38	4.2, 40	1.5, 30	2.1, 36	-0.4, 37	2.7, 38
0.5	(a)	0.283	0.452	0.525	0.560	0.283	0.452	0.525	0.560
	(b)	6.2, 26	5.2, 32	4.1, 34	7.7, 36	6.0, 21	4.1, 24	3.3, 26	6.6, 27
	(c)	6.7, 27	5.9, 33	4.5, 35	8.0, 36	5.9, 23	3.7, 27	3.0, 28	6.4, 28
	(d)	5.6, 34	5.4, 39	3.3, 41	6.5, 42	5.7, 34	3.3, 38	2.6, 39	5.9, 39
0.7	(a)	0.343	0.588	0.689	0.736	0.343	0.588	0.689	0.736
	(b)	6.6, 28	0.7, 34	-1.7, 34	1.8, 35	6.5, 23	-0.6, 27	-2.3, 27	0.9, 27
	(c)	7.0, 30	1.4, 36	-1.3, 35	2.1, 35	6.3, 25	-0.7, 29	-2.3, 29	0.9, 28
	(d)	5.7, 37	0.9, 39	-2.5, 38	0.6, 38	6.0, 36	-1.1, 39	-2.6, 37	0.8, 35
1	(a)	0.362	0.669	0.811	0.872	0.362	0.669	0.811	0.872
	(b)	6.0, 30	-1.5, 36	-4.3, 34	-0.3, 32	5.7, 24	-2.9, 30	-5.0, 29	-1.0, 25
	(c)	6.4, 32	-0.8, 37	-3.9, 34	0.1, 32	5.3, 26	-3.0, 30	-4.9, 28	-0.8, 25
	(d)	5.2, 38	-0.6, 40	-4.0, 36	-0.2, 32	5.6, 36	-3.3, 37	-4.7, 32	-0.6, 27
2	(a)	0.362	0.669	0.811	0.872	0.362	0.669	0.811	0.872
	(b)	6.0, 30	-1.5, 36	-4.3, 34	-0.3, 32	5.7, 24	-2.9, 30	-5.0, 29	-1.0, 25
	(c)	6.4, 32	-0.8, 37	-3.9, 34	0.1, 32	5.3, 26	-3.0, 30	-4.9, 28	-0.8, 25
	(d)	5.2, 38	-0.6, 40	-4.0, 36	-0.2, 32	5.6, 36	-3.3, 37	-4.7, 32	-0.6, 27

(a) True value of  $F_{VW}(v, w)$ ; (b)–(d) Monte-Carlo bias  $\times 10^3$  and standard error  $\times 10^3$  of the proposed estimator with weight  $c$ , the proposed estimator with identity weight, and the SHL-estimator applied to the first pairs of bivariate recurrence times.

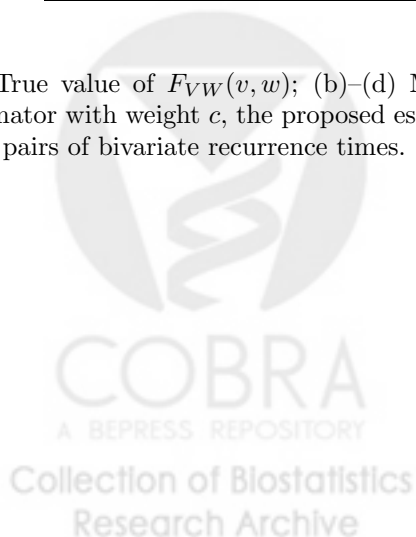
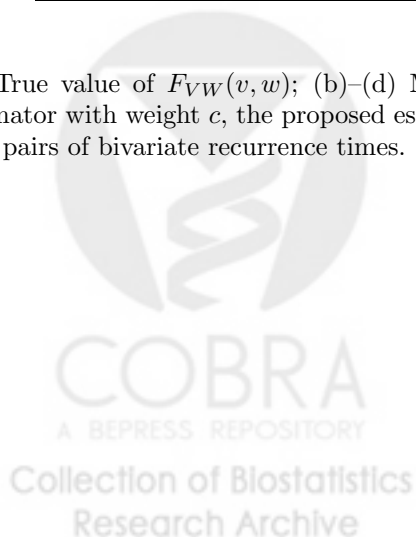


Table 4: *Simulation Summary Statistics:  $Z \sim \exp(1)$  and  $\theta = 9$ .*

		$C \sim \text{uniform}(0, 8)$				$C \sim \text{uniform}(0, 15)$			
		w=0.5	w=1	w=2	w=4	w=0.5	w=1	w=2	w=4
v=	(a)	0.248	0.337	0.380	0.406	0.248	0.337	0.380	0.406
	(b)	-2.5, 23	-1.2, 29	-1.6, 31	-3.3, 34	-2.2, 19	-1.6, 22	-2.0, 25	-3.0, 26
	(c)	-1.8, 24	-0.6, 29	-1.0, 32	-2.6, 34	-2.7, 21	-2.1, 25	-2.6, 27	-3.6, 28
	(d)	-1.8, 31	-3.0, 36	-0.6, 38	-1.9, 41	-2.3, 30	-1.1, 33	-1.5, 36	-2.6, 37
0.5	(a)	0.322	0.467	0.535	0.572	0.322	0.467	0.535	0.572
	(b)	0.5, 27	1.6, 33	-1.0, 34	-3.3, 36	1.1, 21	2.0, 25	-0.8, 27	-2.3, 27
	(c)	1.2, 28	2.3, 33	-0.3, 35	-2.6, 37	0.6, 23	1.6, 27	-1.2, 28	-2.6, 29
	(d)	0.8, 35	1.9, 38	-0.1, 40	-2.1, 42	1.0, 33	2.8, 36	0.0, 37	-1.7, 38
0.7	(a)	0.365	0.596	0.692	0.740	0.365	0.596	0.692	0.740
	(b)	1.7, 23	1.2, 27	-1.1, 27	-1.1, 27	1.3, 29	1.6, 35	-1.5, 35	-2.1, 35
	(c)	1.3, 24	0.9, 28	-1.2, 29	-1.1, 28	2.1, 30	2.1, 35	-0.9, 36	-1.5, 35
	(d)	1.7, 37	2.0, 39	-0.7, 38	-1.3, 38	0.4, 33	0.7, 36	-1.4, 36	-1.8, 36
1	(a)	0.365	0.668	0.807	0.872	0.365	0.668	0.807	0.872
	(b)	2.6, 29	0.2, 36	0.7, 34	-0.3, 33	3.0, 23	-0.9, 29	1.2, 29	1.3, 26
	(c)	3.3, 30	0.5, 36	1.0, 33	-0.1, 32	2.5, 25	-1.1, 29	1.3, 28	1.4, 26
	(d)	3.1, 37	0.7, 39	0.6, 34	-0.7, 32	1.7, 33	-0.6, 34	1.7, 32	1.6, 28
2	(a)	0.365	0.668	0.807	0.872	0.365	0.668	0.807	0.872
	(b)	2.6, 29	0.2, 36	0.7, 34	-0.3, 33	3.0, 23	-0.9, 29	1.2, 29	1.3, 26
	(c)	3.3, 30	0.5, 36	1.0, 33	-0.1, 32	2.5, 25	-1.1, 29	1.3, 28	1.4, 26
	(d)	3.1, 37	0.7, 39	0.6, 34	-0.7, 32	1.7, 33	-0.6, 34	1.7, 32	1.6, 28

(a) True value of  $F_{VW}(v, w)$ ; (b)–(d) Monte-Carlo bias  $\times 10^3$  and standard error  $\times 10^3$  of the proposed estimator with weight  $c$ , the proposed estimator with identity weight, and the SHL-estimator applied to the first pairs of bivariate recurrence times.



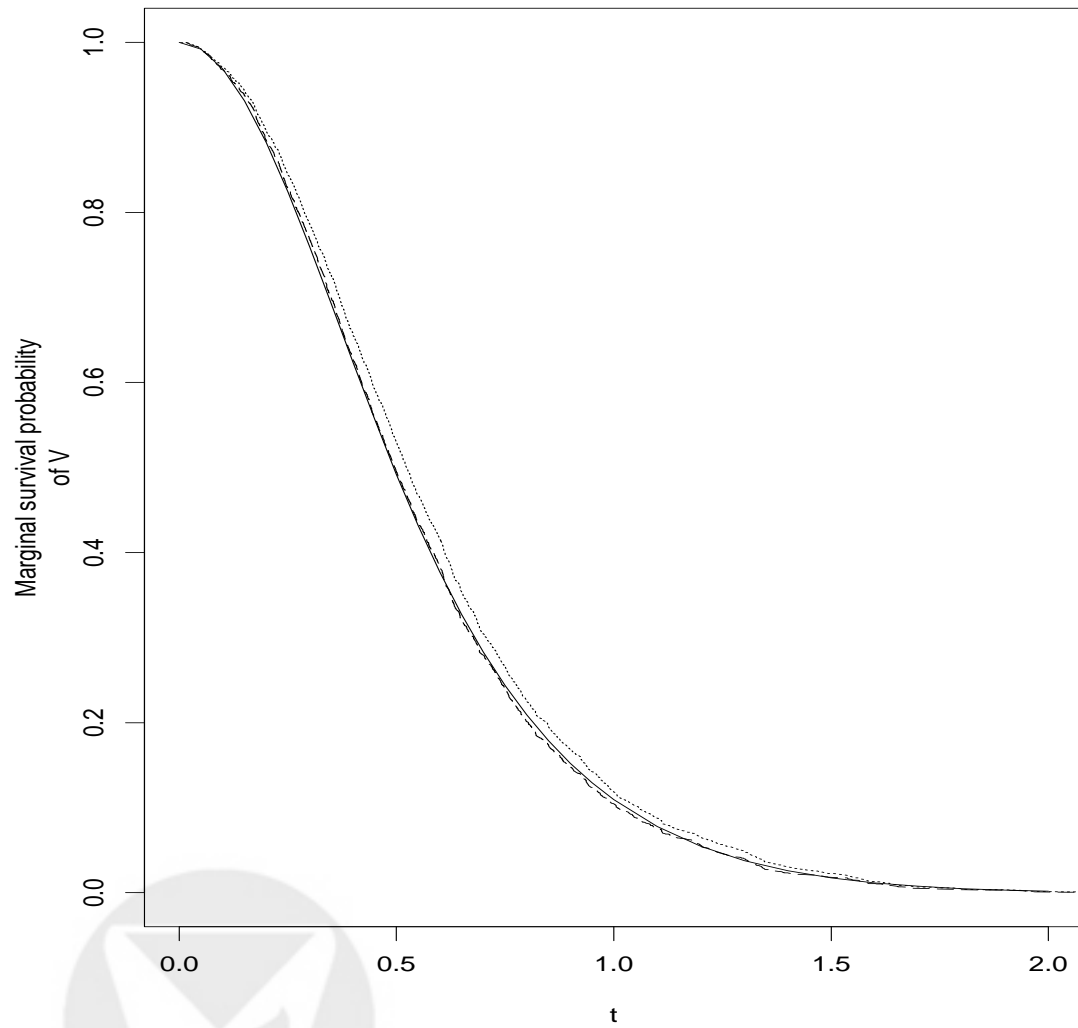


Figure 1: Estimated marginal survival function  $S_V$  with  $n=1000$ . —, true ; - - -, without uncensored  $V_{im_i}$  ; ⋯, with uncensored  $V_{im_i}$

Table 5: *Summary of number of bivariate recurrence times (censored or uncensored)*

		No. of bivariate recurrence times						
		1	2	3	4	5	6	$\geq 7$
No. patients	336	145	53	36	27	20	16	39
(%)	100	43.2	15.8	10.7	8	6	4.8	11.6

Table 6: *Summary of estimates for cumulative joint distribution function*

PCR-EC (months)	PCR-break (months)					
	4	6	12	24	48	60
1	0.05, 0.008	0.07, 0.009	0.11, 0.011	0.15, 0.014	0.20, 0.018	0.25, 0.024
3	0.07, 0.010	0.13, 0.012	0.19, 0.015	0.26, 0.020	0.31, 0.022	0.39, 0.027
6	0.10, 0.012	0.18, 0.015	0.27, 0.018	0.36, 0.022	0.43, 0.025	0.52, 0.030
12	0.12, 0.014	0.22, 0.017	0.33, 0.020	0.43, 0.023	0.50, 0.025	0.60, 0.031
24	0.14, 0.015	0.26, 0.019	0.38, 0.023	0.49, 0.025	0.56, 0.027	0.67, 0.032
48	0.16, 0.016	0.29, 0.022	0.42, 0.025	0.53, 0.027	0.62, 0.029	0.72, 0.032

Empirical averages, bootstrap standard errors for proposed estimator with identity weight function.

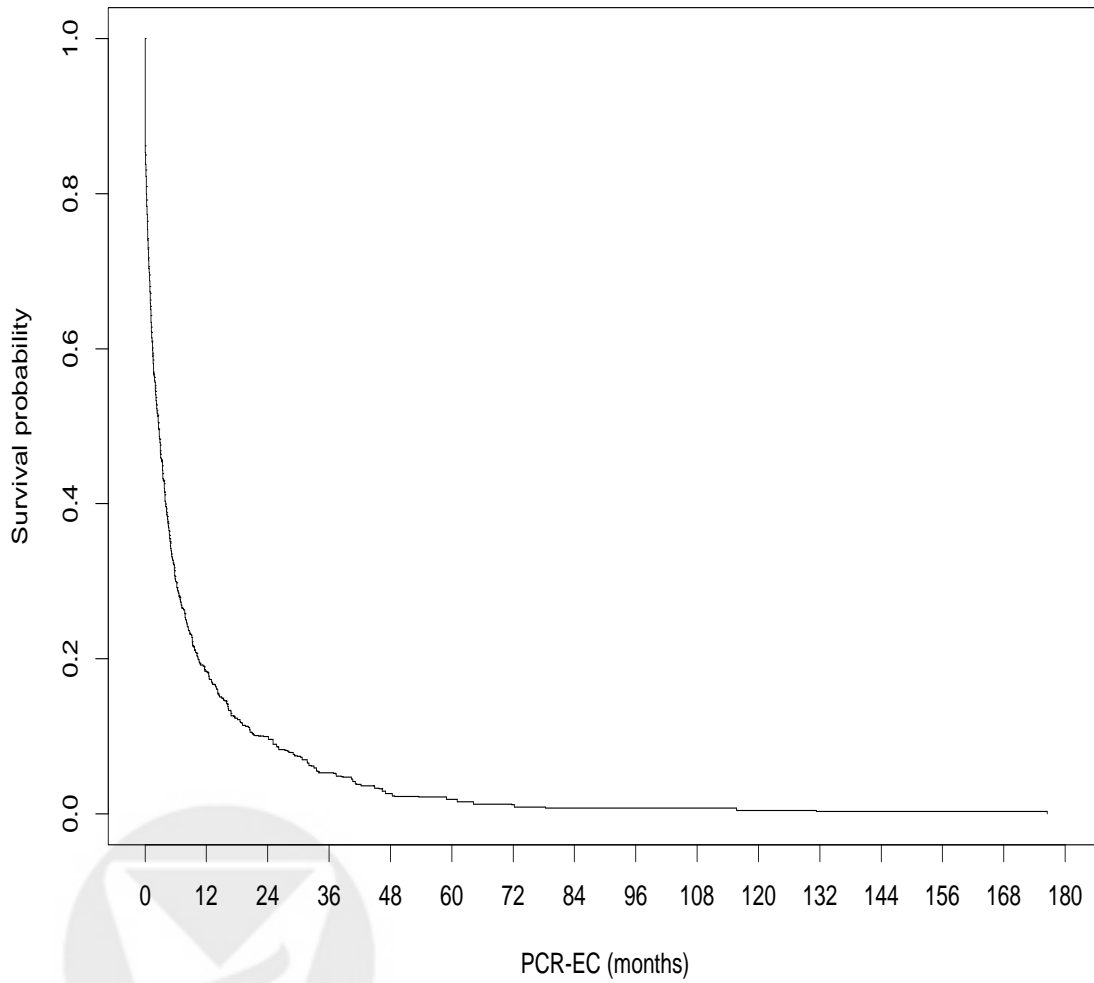
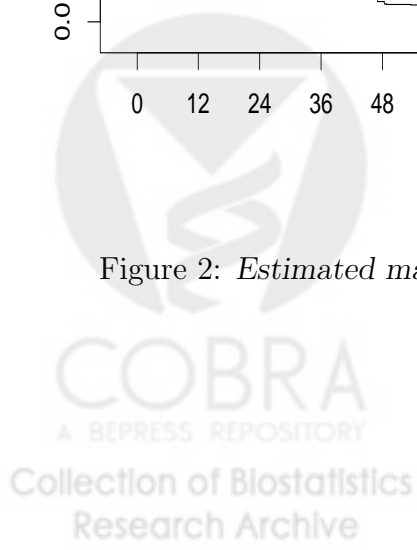


Figure 2: *Estimated marginal recurrent survival function of PCR-EC*





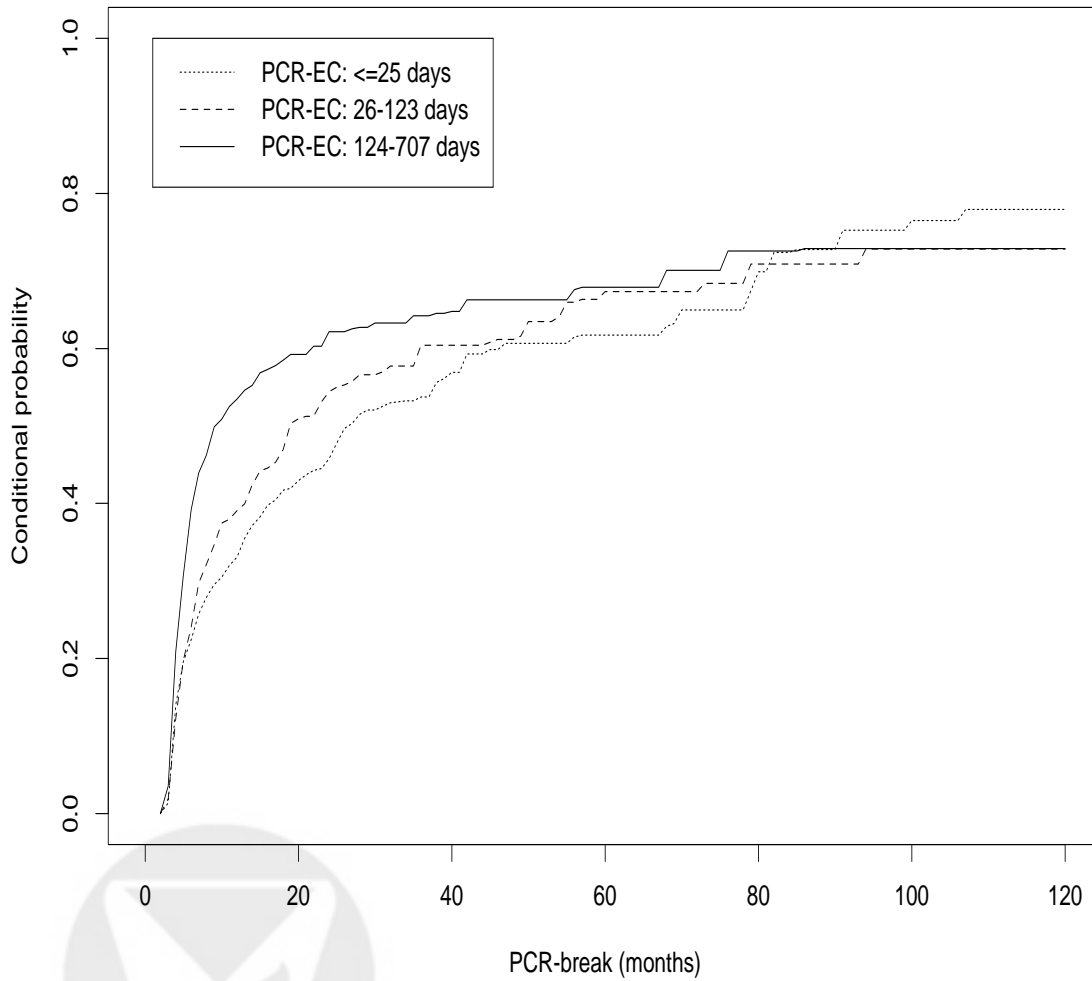


Figure 3: *Estimated cumulative conditional distribution function of PCR-break given different PCR-EC subgroups*